

The logo for NEPS (National Educational Panel Study) features the acronym 'NEPS' in a bold, blue, sans-serif font. To the left of the text is a stylized orange bracket shape that partially encloses the letters.

NEPS

National Educational Panel Study

NEPS Technical Report

**Report on the Coding of Verbatim Answers
About Courses Outside School
From the Perspective of Arts Education**

Jannis Burkhard, Gregor Lampel

The logo for LifBi (Leibniz Institute for Educational Trajectories) consists of the letters 'LifBi' in a bold, black, sans-serif font. A vertical blue bar is positioned to the left of the 'i', and a vertical pink bar is positioned to the left of the 'B'.

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; June 26, 2023

Report on the Coding of Verbatim Answers About Courses Outside School From the Perspective of Arts Education

Jannis Burkhard¹, Gregor Lampel²

¹ *DIPF | Leibniz Institute for Research and Information in Education*

² *Leibniz Institute for Educational Trajectories (LIfBi)*

E-mail address of lead author:

j.burkhard@dipf.de

Bibliographic data:

Burkhard, Jannis & Lampel, Gregor. (2023). *Report on the coding of verbatim answers about courses outside school from the perspective of arts education* (NEPS Technical Report, 2023). Leibniz Institute for Educational Trajectories, National Educational Panel Study.

Acknowledgements:

We thank Matilda Massa Gahein-Sama, Karoline Kairies and Jingyu Zhang for their assistance in creating the data and this report.

Report on the Coding of Verbatim Answers About Courses Outside School From the Perspective of Arts Education

Abstract

This report describes the procedure of coding verbatim answers about attending courses outside school from the perspective of arts education. We first briefly describe the original variables. Then, we outline the coding guidelines that were used and explain the process of coding. Subsequently, we give a brief overview of the created variables. We conclude with recommendations on how to work with the data. The results of this paper refer to the future NEPS-SUF release 13.0.0.

Keywords

NEPS SC3, arts education, courses outside school, coding, verbatim answers

Data

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

1. Original Variables

We used verbatim answers of the scale “courses outside school” from starting cohort 3 for coding. Students are asked whether they attend courses at music schools, adult education centers and youth art schools. They are also asked to give more detail with an open text answer, if they specify yes. Then they are given the opportunity to specify any other courses they attended in two more open text boxes (“what” and “where”). We used data from Scientific Use File (SUF) Version 11.0.0, so data are available until wave 11. Table 1 provides an overview of the original variables.

Table 1: Overview of original variables

Item	Variable	Coding	Waves
<i>Have you attended any out-of-school courses this or last school year (excluding sport)? If so, what exactly did you do?</i>			
classes at the music school (e.g. instrumental or singing classes)	t27111a	1 = yes, 2 = no	1-5, 7-11
	t27111v_O	open	4-5, 7-11
a course at the adult education center (VHS)	t27111b	1 = yes, 2 = no	1-5, 7-11
	t27111w_O	open	4-5, 7-11
a course at the youth art school	t27111c	1 = yes, 2 = no	1-5, 7-11
	t27111x_O	open	4-5, 7-11
<i>Have you attended any other out-of-school courses this or last school year (excluding sport)? If so, what exactly did you do and where?</i>			
what	t27111u_O	open	3-5, 7-11
where	t27111d_O	open	3-5, 7-11

2. Coding Guidelines & Process of Coding

The first draft of the coding guidelines was developed deductively. It contained categories 1-12 and missing code -55 (s. table 2). However, category 12 was not yet labelled “Other / arts education” but only “other”. An initial sample of 101 text answers was randomly selected to test the guidelines. Interrater agreement of this sample coding can be found in appendix (table 6) and was rather poor (Landis & Koch, 1977). Analyzing the answers that were coded differently by the two coders, it became obvious that categories “other” and “not determinable” were not yet clearly distinct. So, the coding guidelines were revised, which led to the final guidelines that can be found in tables 2 (English translation) and 5 (German original). In particular, category “other” was divided into “other / arts education” and “other / non arts education”. Additionally, categories 14 16 were developed inductively and “craft” was explicitly added to category “visual arts”.

Using the final guidelines another sample of 10% (n = 502) of all answers was coded and interrater agreement analyzed. Results of the interrater agreement can be found in the appendix (table 7). Agreement had improved substantially, so it could have been considered to continue with single coding each answer. But in order to achieve better reliability and validity it was decided to proceed with double coding. After all, every answer was coded by

two coders separately. The codes were validated automatically for those cases that both coders agreed upon. The answers that were coded differently by the two coders were reviewed and validated either by a third person or discussed and validated in a dialogue.

The coders received both content-related training on how to use the guidelines and technical training on how to use the coding tool. We used the coding tool CODI developed by Leibniz Institute for Educational Trajectories (LifBi). CODI enables the coding of open entries in a standardized setting, in which the coders are supported by suggestion algorithms. The algorithm applies a text and similarity search by which the words (word stem, excluding stop words) of the open text data are compared with the labels of the categories, the text descriptions of the categories and predefined keywords. This result is the basis for a ranking, which is presented to the coders as a descending score. The algorithm never learns from the other coder when multiple codings are made, which keeps the setting comparable. Coders, on the other hand, have the option of making corrections to already taken decisions in their own learning process. Coders select their choices by mouse click and have the option to leave notes and comments to the supervisor to reason the decision. This makes it easier to reach a decision in the case of disagreement between several coders. The validation process also takes place in CODI.

Table 2: Final Coding Guidelines (English translation)

Code	Definition	Encoding Rule	Example
1	Course in music	Answers of this type indicate that the target person learns to produce or cognitively reflect music	“Gitarrenunterricht”, “Jazzband”, “Orchester”
2	Course in dance	Answers of this type indicate that the target person learns to dance or cognitively reflect dance	“Bauchtanzkurs”, “Tanzschule”, “Gesellschaftstanzkurs”, “Volkstanz”
3	Course in theater	Answers of this type indicate that the target person participates in scenic, predominantly spoken performances or learns how to reflect them cognitively	“Theater AG”, “Ferienkurs: Theater”, “Schauspielklasse”
4	Course in visual arts / crafts	Answers of this type indicate that the target person learns to produce or cognitively reflect objects of visual arts or craft	“Malkurs”, “Zeichnen”, “Nähkurs”, “Klöppeln”, “Oregami”
5	Course in literature	Answers of this type indicate that the target person learns to produce or cognitively reflect belletristic texts	“Kreatives Schreiben”, “Schriftsteller AG”, “Schülerzeitung”, “Buch-AG”

6	Course in film/photography	Answers of this type indicate that the target person learns to produce or cognitively reflect moving or non-moving images	"Fotokurs", "Kamerakurs"
7	Course in media design	Answers of this type indicate that the target person learns to produce aesthetic-expressive objects using (mostly digital) media or cognitively reflect those	"Schulradio", "Poer-Point", "Bürgerradio", "Game Disine"
8	Course in languages	Answers of this type indicate that the target person learns a language	"Arabisch lernen", "Cambridge Certificate", "Italienisch"
9	Course in technology	Answers of this type indicate that the target person learns about the usage of technology without aesthetic-expressive aspirations being paramount	"Excel" "Mädchen und Technik" "Roboter AG"
10	Course in cultural heritage / museum	Answers of this type indicate that the target person attends a course that covers cultural heritage or that takes place at a museum	"im Naturkundemuseum"
11	Course in circus	Answers of this type indicate that the target person learns to take part in or cognitively reflect artistic (circus) performances	"Akrobatik", "Zauberei"
12	Other / arts education	Answers of this type contain information on arts education but cannot be assigned any of the other categories	"Folklore", "Traditionspflege", "Varieté"
13	Other / non arts education	Answers of this type contain information on other courses that are not clearly part of arts education	"Reitkurs", "Survival", "Erste Hilfe", "AG Kochen und Backen", "Fahrschule", "Obstbaumschnittkurs" "Retorikkurs"
14	Course in social commitment	Answers of this type indicate that the target person attends a course which leads to active engagement in society	"Jugendfeuerwehr", "Tierschutzjugend XXX" "Alt und Jung"
15	Course in religion	Answers of this type indicate that the target person attends a course in which religious contents are taught	"Konfirmandenunterricht", "Mädchenkreis", "Christenlehre", "Ministranten"
16	Course in school related areas / private tutoring	Answers of this type indicate that the target person receives education in school relevant contents	"Mathenachhilfe", "Naturwissenschaftliches Praktikum", "Mathe Abivorbereitungskurs",

			“Physik AG”, “Tipps zum Vortragen”
-55	Not determinable	Meaningless answers	“Es tut mir leid, ich hatte keine Konzentration mehr”, “Nichts”, “Bei meinem Nachbarn”

3. Overview of Created Variables

Because it was possible to assign up to 3 categories to each answer, every original variable resulted in two or three new variables. Table 3 shows the names of the newly generated variables with the corresponding number of observations across all waves. Most answers received one category only. 3 categories were only assigned to courses in music schools and other courses, which is why no third variable was created for courses in adult education centers and youth art schools.

Table 4 contains information on how often categories were assigned per wave. It should be noted that the music category was assigned to 55.5% of all answers and thus is by far the most assigned category. Other categories, in contrast, were assigned much less.

Table 3: Overview of new variables

Original Variable	New Variable 1 st category	New Variable 2 nd category	New Variable 3 rd category
t27111v_O Music school	t27111v_g4 n = 5672	t27111v_g5 n = 51	t27111v_g6 n = 14
t27111w_O Adult education center	t27111w_g4 n = 636	t27111w_g5 n = 29	- -
t27111x_O Youth art school	t27111x_g4 n = 209	t27111x_g5 n = 8	- -
t27111u_O & t27111d_O Other courses (what & where)	t27111u_g4 n = 5003	t27111u_g5 n = 296	t27111u_g6 n = 37
Sum	11520	384	51

Note: The value of -55 is excluded from the calculation of the number of cases.

Table 4: Sum of assigned categories per wave

	W3	W4	W5	W7	W8	W9	W10	W11	Sum (per row)	in %
Music	829	2053	1624	1099	777	627	120	10	7139	55.4
Dance	119	85	123	80	35	36	7	4	489	3.8
Theater	62	70	62	38	37	22	6	0	297	2.3
Visual arts / craft	127	128	106	76	42	43	7	1	530	4.1
Literature	12	7	9	8	3	3	0	0	42	0.3
Film / Photography	10	14	16	15	7	7	3	0	72	0.6
Media design	5	12	6	8	5	7	1	0	44	0.3
Languages	35	62	83	62	32	31	10	0	315	2.4
Technology	58	57	73	37	19	12	1	1	258	2.0
Museum	4	1	1	0	0	0	0	0	6	0.0
Circus	9	5	7	2	0	0	0	0	23	0.2
Other / arts education	6	4	2	1	3	4	1	0	21	0.2
Other / non arts education	448	291	307	294	188	176	39	5	1748	13.6
Social commitment	80	65	73	61	20	25	3	0	327	2.5
Religion	75	130	65	40	18	14	2	0	344	2.7
School related / private tutoring	60	64	58	37	26	37	14	4	300	2.3
Not determinable	153	206	171	215	99	47	22	10	923	7.2
Sum (per column)	2092	3254	2786	2073	1311	1091	236	35	12,878	100.0
in %	16.2	25.3	21.6	16.1	10.2	8.5	1.8	0.3	100.0	

4. Recommendations for working with the data

The data/variables which resulted from the coding will be published with SUF version 13.0.0. If you want to work with the data, consider the following:

- Newly generated variables work like dummy variables
One text answer can contain multiple courses of the same category, which will not show in the new variables. E.g. the answer “drums, violin, orchestra” would only receive the category “music” one time, not three times. Do not try to build a sum of all courses attended.
- Content-wise bias
Note that “music” is the category that was assigned by far most often when interpreting potential results.
- Dropout in later waves
Be aware of the low case numbers in later waves. In addition to the panel mortality inherent to all panel studies, the scale “courses outside school” is presented in a computer assisted web interview (CAWI) for target persons who are individually followed up. Many interviewees decide to not fill in the online questionnaire after having the personal interview, which contributes to the low sample sizes in later waves. Please refer to the general NEPS documentation for further information.

5. Literature

- Blossfeld, H.-P., & Roßbach, H.-G. (2019). *Education as a lifelong process: The German National Educational Panel Study (NEPS)* (Bd. Edition ZfE (2nd ed.)). Springer VS.
- Klein, D. (2019). Implementing a General Framework for Assessing Interrater Agreement. *The Stata Journal: Promoting communications on statistics and Stata* 18, 871–901.
- Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1), 159–174.
- NEPS-Network. (2022). *National Educational Panel Study, Scientific Use File of Starting Cohort Grade 5*. Leibniz Institute for Educational Trajectories (LifBi). Bamberg.

6. Appendix

Table 1: Final Coding Guidelines (German original)

Code	Definition	Codierregel	Ankerbeispiele
1	Kurs im Bereich Musik	Angaben dieses Typs lassen darauf schließen, dass die Zielperson lernt, Musik selbst zu produzieren oder diese kognitiv zu reflektieren	„Gitarrenunterricht“, „Jazzband“, „Orchester“
2	Kurs im Bereich Tanz	Angaben dieses Typs lassen darauf schließen, dass die Zielperson lernt selbst zu tanzen oder Tanz kognitiv zu reflektieren.	„Bauchtanzkurs“, „Tanzschule“, „Gesellschaftstanzkurs“, „Volkstanz“
3	Kurs im Bereich Sprechtheater/ Schauspiel	Angaben dieses Typs lassen darauf schließen, dass die Zielperson selbst an szenischen, überwiegend gesprochenen Aufführungen beteiligt ist oder lernt, diese kognitiv zu reflektieren.	„Theater AG“, „Ferienkurs: Theater“, „Schauspielklasse“
4	Kurs im Bereich Bildende Kunst / (Kunst-)Handwerk	Angaben dieses Typs lassen darauf schließen, dass die Zielperson lernt, visuell gestaltete Kunstobjekte zu produzieren oder diese kognitiv zu reflektieren.	„Malkurs“, „Zeichnen“, „Nähkurs“, „Klöppeln“, „Oregami“
5	Kurs im Bereich Literatur	Angaben dieses Typs lassen darauf schließen, dass die Zielperson lernt, belletristische Texte selbst zu verfassen oder diese kognitiv zu reflektieren.	„Kreatives Schreiben“, „Schriftsteller AG“, „Schülerzeitung“, „Buch-AG“
6	Kurs im Bereich Film/Fotografie	Angaben dieses Typs lassen darauf schließen, dass die Zielperson lernt, bewegte oder nicht bewegte Bilder zu produzieren oder diese kognitiv zu reflektieren.	„Fotokurs“, „Kamerakurs“
7	Kurs im Bereich Mediengestaltung	Angaben dieses Typs lassen darauf schließen, dass die Zielperson lernt, mithilfe von (meist digitalen) Medien ästhetisch-expressive Objekte zu erschaffen oder diese kognitiv zu reflektieren.	„Schulradio“, „Poer-Point“, „Bürgerradio“, „Game Disine“
8	Kurs im Bereich Sprachlernen	Angaben dieses Typs lassen darauf schließen, dass die Zielperson in dem Kurs eine Sprache erlernt. Sprachkurse, die nicht explizit als Nachhilfe angegeben werden, sollen nur als 8 kodiert werden (ohne Doppelkodierung)	„Arabisch lernen“, „Cambridge Certificate“, „Italienisch“

9	Kurs im Bereich Technik	Angaben dieses Typs lassen darauf schließen, dass die Zielperson sich in der Verwendung von Technik weiterbildet, ohne dass ästhetisch-expressive Ansprüche im Vordergrund stehen.	„Excel“ „Mädchen und Technik“ „Roboter AG“
10	Kurs im Bereich Museum	Angaben dieses Typs lassen darauf schließen, dass die Zielperson einen Kurs besucht, der kulturelles Erbe behandelt oder der an einem Museum stattfindet.	„im Naturkundemuseum“
11	Kurs im Bereich Zirkus	Angaben dieses Typs lassen darauf schließen, dass die Zielperson lernt, selbst zirkensische (artistische) Darbietungen aufzuführen oder diese kognitiv zu reflektieren	„Akrobatik“, „Zauberei“
12	Sonstige – KuBi	Angaben dieses Typs enthalten Informationen zu kultureller Bildung, kann aber den anderen Kategorien nicht eindeutig zugeordnet werden	„Folklore“, „Traditionspflege“, „Varieté“
13	Sonstige – nicht KuBi	Angaben dieses Typs enthalten Informationen zu sonstigen Kursen, die aber nicht eindeutig in den Bereich kultureller Bildung fallen	„Reitkurs“, „Survival“, „Erste Hilfe“, „AG Kochen und Backen“, „Fahrschule“, „Obstbaumschnittkurs“ „Retorikkurs“
14	Kurs im Bereich Gesellschaftliches Engagement	Angaben dieses Typs lassen darauf schließen, dass die Zielperson sich mit einem Kurs gesellschaftlich engagiert	„Jugendfeuerwehr“, „Tierschutzjugend XXX“ „Alt und Jung“
15	Kurs im Bereich Religion	Angaben dieses Typs lassen darauf schließen, dass die Zielperson einen Kurs zu religiösen Inhalten besucht	„Konfirmandenunterricht“ „Mädchenkreis“ „Christenlehre“ „Ministranten“
16	Kurs im Bereich Nachhilfe / schulbezogene Inhalte	Angaben dieses Typs lassen darauf schließen, dass die Zielperson sich in schulischen Inhalten weiterbildet	„Mathenachhilfe“, „Naturwissenschaftliches Praktikum“ „Mathe Abivorbereitungskurs“ „Physik AG“ „Tipps zum Vortragen“
-55	Nicht ermittelbar	Inhaltlose Angaben	„Es tut mir leid, ich hatte keine Konzentration mehr“ „Nichts“ „Bei meinem Nachbarn“

Table 2: Interrater agreement of sample coding with initial guidelines

Number of subjects = 101
 Ratings per subject = 2
 Number of rating categories = 13

	Coef.	Std. Error	t	p> t 	[95 % Conf. Interval]	
Percent Agreement	.5743	.0494	11.61	0.000	.4762	.6724
Brennan and Prediger	.5388	.0536	10.06	0.000	.4325	.6451
Cohen/Conger's Kappa	.5150	.0551	9.35	0.000	.4057	.6242
Scott/Fleiss' Pi	.4881	.0646	7.56	0.000	.3599	.6162
Gwet's AC	.5426	.0528	10.27	0.000	.4377	.6474
Krippendorff's Alpha	.4906	.0646	7.60	0.000	.3625	.6187

*Calculated using the ado kappaetc (Klein, 2019)

Table 7: Interrater agreement with final coding guidelines

Number of subjects = 502
 Ratings per subject = 2
 Number of rating categories = 43

	Coef.	Std. Error	t	p> t 	[95 % Conf. Interval]	
Percent Agreement	.8486	.0160	52.99	.000	.8171	.8801
Brennan and Prediger	.8450	.0164	51.54	.000	.8128	.8772
Cohen/Conger's Kappa	.8237	.0184	44.85	.000	.7876	.8598
Scott/Fleiss' Pi	.8236	.0184	44.79	.000	.7875	.8598
Gwet's AC	.8454	.0164	51.70	.000	.8133	.8776
Krippendorff's Alpha	.8238	.0184	44.80	.000	.7877	.8600

*Calculated using the ado kappaetc (Klein, 2019)

Table 3: Sums of assigned categories per original variable

	Music school (t27111v_O)	Adult education center (t27111w_O)	Youth art school (t27111x_O)	Other courses (t27111u_O & t27111d_O)
Music	5491	52	10	1586
Dance	41	42	7	399
Theater	46	6	21	224
Visual art / craft	16	84	142	288
Literature	1	5	3	33
Film / Photography	6	26	1	39
Media design	2	13	7	22
Languages	10	119	0	186
Technology	4	88	2	164
Museum	0	0	1	5
Circus	1	3	0	19
Other / arts education	0	1	2	18
Other / non arts education	76	177	14	1481
Social commitment	15	4	2	306
Religion	12	1	2	329
School related / private tutoring	16	44	3	237
Not determinable	96	88	64	675