NEPS
National Educational Panel Study

Forschungsdaten

FDZ-LIfBi

Datenmanual

NEPScomp

Version 2.0.0

LIfBi
LEIBNIZ-INSTITUT FÜR
BILDUNGSVERLÄUFE

Research Data Documentation

Die *NEPS Research Data Documentation Series* versammelt Materialien, die der Unterstützung bei der Arbeit mit den Daten des Nationalen Bildungspanels dienen.

Kontakt

E-Mail: fdz@lifbi.de
Webseite: https://www.lifbi.de/FDZ

Zitation des Dokuments

FDZ-LIfBi. (2026). *Datenmanual NEPScomp, Version 2.0.0*. Bamberg, Leibniz-Institut für Bildungsverläufe, National Educational Panel Study.

Würdigung

Das vorliegende Handbuch zum Datenbestand „NEPScomp“ wurde von Mitarbeitenden des Forschungsdatenzentrums des Leibniz-Instituts für Bildungsverläufe (FDZ-LIfBi) erstellt.

Ein besonderer Dank für die geleistesten Beiträge geht an:

Daniel Fuß
Tobias Koberg
Katja Vogel

Dieses Dokument ist unter CC BY 4.0 lizenziert, für Details siehe:
https://creativecommons.org/licenses/by/4.0/

Leibniz-Institut für Bildungsverläufe (LIfBi)
Wilhelmsplatz 3, 96047 Bamberg
Direktorin: Prof. Dr. Cordula Artelt
Administrativer Direktor: Dr. Stefan Echinger

Bamberg, 25. Oktober 2026

Inhaltsverzeichnis

1 Einleitung **1**

2 Nationales Bildungspanel **2**

3 NEPScomp **6**
3.1 Erzeugung von NEPScomp . . . 7
3.2 Datenstruktur von NEPScomp . . . 8
3.3 Definition fehlender Werte in NEPScomp . . . 11
3.4 Umformatierte Datensätze in NEPScomp . . . 13
3.5 Zusammenfassung . . . 15

4 Datendokumentation **16**
4.1 Dokumentation zu NEPScomp . . . 16
4.2 Weiterführende Dokumentation . . . 16
4.3 Publikationen . . . 18
4.4 Serviceangebote . . . 19
4.5 Weiterentwicklung . . . 20

5 Datennutzung **22**
5.1 Zugang zu den Daten . . . 22
5.2 Zitation der Daten . . . 24
5.3 Empfehlungen . . . 25
5.4 Kontakt zum Forschungsdatenzentrum des LIfBi . . . 27

A Appendix **29**
A.1 Beispiel für Merging-Syntax in SPSS . . . 29
A.2 Beispiel für Merging-Syntax in R . . . 30
A.3 Kompetenzdaten . . . 31
A.4 Zeitkonstante Variablen . . . 42
A.5 Release Notes . . . 54

1 Einleitung

NEPScomp ist ein zusätzliches Angebot für Studierende und alle, die neu in die Arbeit mit den komplexen Forschungsdaten des Nationalen Bildungspanels (NEPS, National Educational Panel Study) einsteigen. Es ist insbesondere für den Einsatz in der universitären Lehre konzipiert. Das „comp“ im Titel – Kurzform für „compendium“ – verweist auf den Anspruch dieses Datenprodukts: Durch eine Reihe von Vereinfachungen und Kürzungen im Vergleich zu den regulären Scientific-Use-Files des NEPS soll ein kompakter und leicht zu analysierender Datenbestand zur Verfügung stehen, der inhaltlich sinnvolle Befunde zu einer Vielzahl von Fragestellungen ermöglicht.

Hintergrund der Entwicklung von NEPScomp ist die zunehmende Nachfrage nach qualitativ hochwertigen und umfassend dokumentierten Paneldaten, die für die Ausbildung in der empirisch-statistischen Sozial- und Lebensverlaufsforschung sowie für studentische Abschlussarbeiten geeignet sind. Primäres Ziel ist es, interessierten Personen den Einstieg in die Arbeit mit den Daten des Nationalen Bildungspanels zu erleichtern und für den weiteren Verlauf einer wissenschaftlichen Karriere an das ausführlichere Datenmaterial der regulären Scientific-Use-Files heranzuführen. NEPScomp ist daher so gestaltet, dass ein Wechsel zum vollständigen Datenangebot der einzelnen NEPS-Startkohorten jederzeit möglich ist.

Das vorliegende Datenmanual enthält alle wesentlichen Informationen für den Umgang mit NEPScomp. Es ist das zentrale Dokument, um die Besonderheiten der Datenstruktur und die wichtigsten Empfehlungen für die Arbeit mit den Daten kennenzulernen. Nach dieser Einleitung gibt das **zweite Kapitel** zunächst einen kurzen Überblick zum Nationalen Bildungspanel und dessen Datenangebot. Im **dritten Kapitel** werden die Schritte zur Erzeugung von NEPScomp und insbesondere die Unterschiede zu den regulären Scientific-Use-Files erläutert. Das **vierte Kapitel** verweist auf das umfangreiche Angebot an Dokumentations- und Informationsmaterialien zu NEPScomp und zu den einzelnen Startkohorten. Hinweise zum Zugang zu den NEPScomp-Daten und zur Datenzitation sowie Empfehlungen zur Verwendung dieser Daten finden sich im **fünften Kapitel**. Das **sechste Kapitel** widmet sich der Weiterentwicklung von NEPScomp, die vor allem im Austausch mit den Datennutzenden erfolgen soll. Das Forschungsdatenzentrum des LIfBi hat hierfür einen speziellen Bereich auf der Online-Plattform „Forum4MICA“ eingerichtet, um einerseits Fragen zu beantworten und andererseits Feedback zu sammeln.

NEPScomp ist als ein dynamisches Datenprodukt konzipiert. Es ist ausdrücklich intendiert, das Angebot sowohl inhaltlich als auch strukturell weiter an die Bedarfe der Zielgruppen anzupassen. Aus diesem Grund sei bereits an dieser Stelle herzlich zur Mitwirkung an der Weiterentwicklung aufgerufen. Das Forschungsdatenzentrum des LIfBi nimmt jederzeit Hinweise auf Probleme, Anregungen zu Verbesserungen und sonstige Kommentare zu NEPScomp entgegen.

2 Nationales Bildungspanel

Das Nationale Bildungspanel (NEPS) ist die größte Langzeit-Bildungsstudie in Deutschland. Seit 2009 befragt und testet das NEPS über 70.000 Teilnehmende, um Bildungsprozesse und Kompetenzentwicklungen in ihren Bedingungen, Ergebnissen und Folgen über den gesamten Lebensverlauf beschreiben und erklären zu können. Zunächst deckten sechs sogenannte Startkohorten – bezogen auf das aktuelle Alter bzw. die jeweilige Bildungsetappe der Zielpersonen zum Start der Erhebungen – die Zeitspanne von der frühen Kindheit bis ins hohe Erwachsenenalter ab. Seit dem Jahr 2022 läuft eine weitere Startkohorte, die mehr als zehn Jahre nach dem erstmaligen Einsatz von Fünfklässler:innen im Rahmen der Startkohorte 3 erneut mit repräsentativ ausgewählten Schüler:innen der fünften Jahrgangsstufe startete. Abbildung 1 veranschaulicht das komplexe Multi-Kohorten-Sequenz-Design des NEPS über einen Zeitraum von 15 Jahren mit in der Regel jährlichen Erhebungen innerhalb der sieben Startkohorten (sortiert nach Alter der Zielpersonen auf der Y-Achse). Die unterschiedlichen Farben der Kästchen im Verlauf einer jeden Startkohorte machen den Anspruch des Designs deutlich – die Begleitung der Zielpersonen über mindestens zwei bedeutsame Bildungsübergänge hinweg.

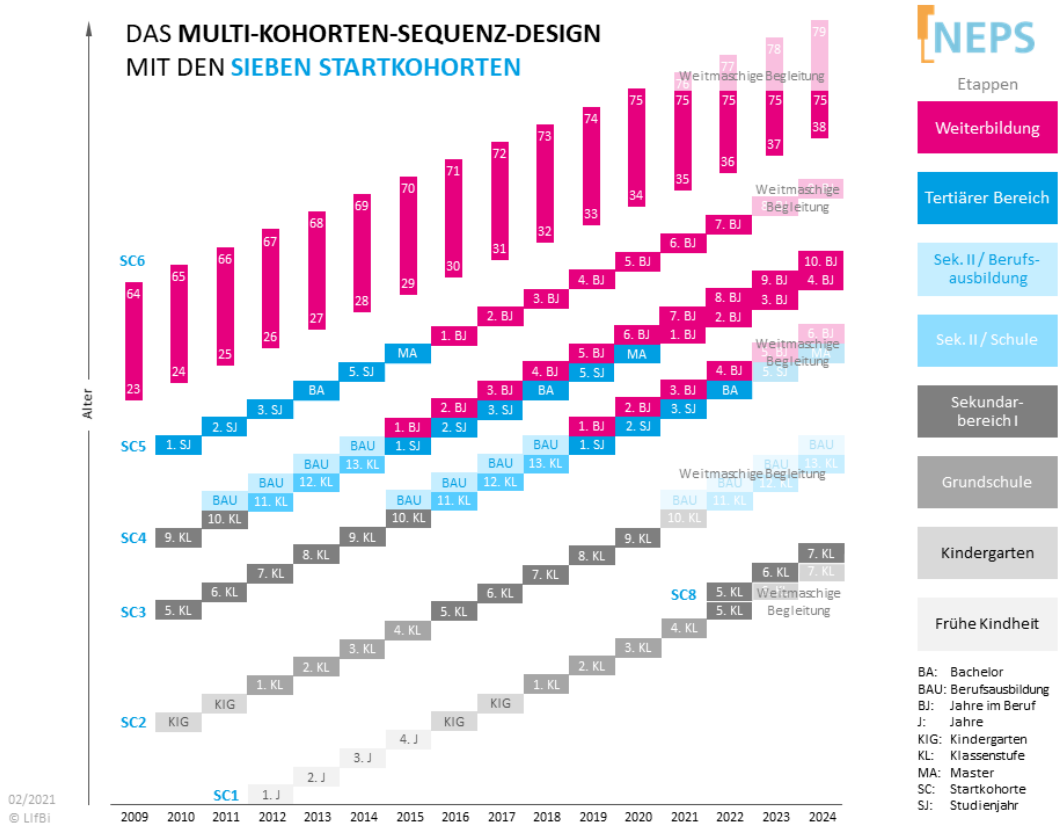

Abbildung 1: Multi-Kohorten-Sequenz-Design des NEPS

Die im Rahmen des NEPS gesammelten Informationen stammen nicht nur von den Zielpersonen selbst, sondern darüber hinaus von etwa 50.000 Personen aus dem direkten Umfeld. Zu diesen Kontextpersonen zählen insbesondere Eltern und pädagogisches Fachpersonal. Die nachfolgende Auflistung gibt einen Überblick über alle NEPS-Startkohorten und Merkmale der zum Veröffentlichungszeitpunkt des Manuals aktuell verfügbaren Scientific-Use-Files (SUF).

Startkohorte 1 – Neugeborene

Zielpersonen	6-8 Monate alte Kleinkinder bei Panelstart
Kontextpersonen	Elternteil (i. d. R. Mutter), Tagesbetreuung, Kita-Gruppenbetreuung, Kita-Einrichtungsleitung
Laufzeit des Panels	engmaschig: 2012-2025
Aktueller SUF	am 05.02.2025 veröffentlicht `doi:10.5157/NEPS:SC1:12.0.0` Daten aus 12 Panelwellen N=3.481 Elterninterviews in Welle 1 N=1.288 Zielpersonen in Welle 12

Startkohorte 2 – Kindergarten

Zielpersonen	Kita-Kinder zwei Jahre vor Einschulung bei Panelstart
Kontextpersonen	Elternteil, Kita-Gruppenbetreuung, Kita-Einrichtungsleitung, Lehrkräfte, Schulleitung
Laufzeit des Panels	engmaschig: 2011-2022
Aktueller SUF	am 09.07.2024 veröffentlicht `doi:10.5157/NEPS:SC2:11.0.0` Daten aus 11 Panelwellen N=2.948 Zielpersonen in Welle 1 Aufstockung um N=6.176 Zielpersonen in Welle 3 N=3.344 Zielpersonen in Welle 11

Startkohorte 3 – Klasse 5 (2010)

Zielpersonen	Schüler:innen der 5. Jahrgangsstufe bei Panelstart
Kontextpersonen	Elternteil, Lehrkräfte, Schulleitung
Laufzeit des Panels	engmaschig: 2010-2022
Aktueller SUF	am 28.11.2025 veröffentlicht `doi:10.5157/NEPS:SC3:13.1.0` Daten aus 13 Panelwellen N=5.774 Zielpersonen in Welle 1 Aufstockung um N=2.146 Zielpersonen in Welle 3 N=2.706 Zielpersonen in Welle 13

Startkohorte 4 – Klasse 9

Zielpersonen	Schüler:innen der 9. Jahrgangsstufe bei Panelstart
Kontextpersonen	Elternteil, Lehrkräfte, Schulleitung
Laufzeit des Panels	engmaschig: 2010-2025
Aktueller SUF	am 22.12.2025 veröffentlicht `doi:10.5157/NEPS:SC4:16.0.0` Daten aus 16 Panelwellen N=16.106 Zielpersonen in Welle 1 N=2.763 Zielpersonen in Welle 16

Startkohorte 5 – Studierende

Zielpersonen	Studierende im 1. Semester bei Panelstart
Laufzeit des Panels	engmaschig: 2010-2022
Aktueller SUF	am 30.09.2024 veröffentlicht `doi:10.5157/NEPS:SC5:19.0.0` Daten aus 19 Panelwellen N=17.909 Zielpersonen in Welle 1 N=5.662 Zielpersonen in Welle 19

Startkohorte 6 – Erwachsene

Zielpersonen	23 bis 64 Jahre alte Erwachsene bei Panelstart
Laufzeit des Panels	engmaschig: 2007-...
Aktueller SUF	am 10.08.2026 veröffentlicht `doi:10.5157/NEPS:SC6:17.0.0` Daten aus 17 Panelwellen N=6.778 Zielpersonen in Welle 1 (ALWA) Aufstockung um N=5.154 Zielpersonen in Welle 2 (1. NEPS-HE) Aufstockung um N=5.208 Zielpersonen in Welle 4 (3. NEPS-HE) N=4.075 Zielpersonen in Welle 17 (16. NEPS-HE)

Startkohorte 8 – Klasse 5 (2022)

Zielpersonen	Schüler:innen der 5. Jahrgangsstufe bei Panelstart
Kontextpersonen	Elternteil, Lehrkräfte, Schulleitung
Laufzeit des Panels	engmaschig: 2022-...
Aktueller SUF	am 23.07.2026 veröffentlicht `doi:10.5157/NEPS:SC8:2.0.0` Daten aus 2 Panelwellen N=5.763 Zielpersonen in Welle 1 N=5.215 Zielpersonen in Welle 2

Das Erhebungsprogramm des NEPS basiert auf einem konzeptionellen Rahmen, der inhaltliche Themenbereiche (*Säulen*) mit über den gesamten Lebensverlauf verteilten Bildungsphasen (*Etappen*) kombiniert. Im Zentrum stehen die in Abbildung 2 dargestellten Kernthemen der Entwicklung von Kompetenzen und Persönlichkeitsmerkmalen, der Formation von Bildungsentscheidungen, der Rolle unterschiedlicher Lernumwelten und sozialer Ungleichheiten – insbesondere nach sozialer Herkunft, Migrationshintergrund und Geschlecht – auf Bildungsprozesse sowie die Auswirkungen auf Bildungsrenditen. Diese Themen sind über den gesamten Bildungsverlauf eines Individuums von Relevanz. Charakteristisch für die Lebensverlaufsperspektive des NEPS ist die Prämisse, dass spätere Lebensphasen von früheren Erfahrungen bzw. Zuständen, äußeren Ereignissen und getroffenen Entscheidungen beeinflusst sind. Deshalb beginnt die Betrachtung von Bildungsverläufen im NEPS bei der vorschulischen Betreuung und reicht über die gesamte Zeit des Schulbesuchs, der Ausbildung und des Studiums bis hin zur beruflichen Weiterbildung und zur (nachgeholten) Bildung im Erwachsenenalter.

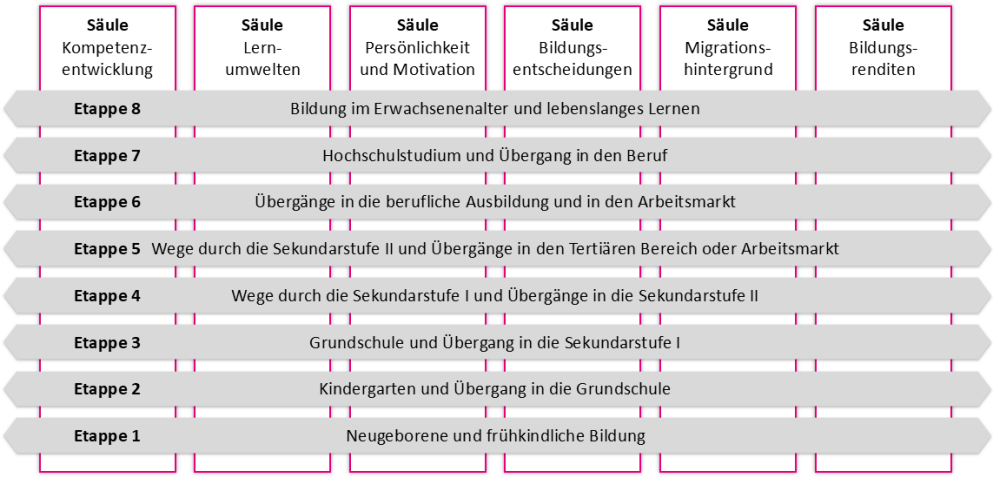

Abbildung 2: Rahmenkonzeption des NEPS

Ausführliche Informationen zum Aufbau des Nationalen Bildungspanels und zur initialen Rahmenkonzeption mit den sechs ursprünglichen Startkohorten enthält der Sammelband von Blossfeld und Roßbach (2019). Eine aktualisierte Fassung der Rahmenkonzeption des NEPS findet sich in Artelt und Sixt (2023) sowie in NEPS-Netzwerk (2023).

Beheimatet ist das NEPS am Leibniz-Institut für Bildungsverläufe (LIfBi) in Bamberg. Von 2009 bis 2013 war es zunächst als Projekt am Institut für bildungswissenschaftliche Längsschnittforschung (INBIL) der Otto-Friedrich-Universität Bamberg angesiedelt. Seit 2014 wird die Studie vom LIfBi in enger Kooperation mit einem deutschlandweiten Netzwerk interdisziplinärer Expert:innen fortgeführt. Details zur Geschichte des NEPS, den kooperierenden Forschungseinrichtungen, der Rahmenkonzeption und den verschiedenen Themenbereichen sind auf der entsprechenden LIfBi-Webseite zusammengestellt.

→ `www.lifbi.de > Forschung > Großprojekte > NEPS – Nationales Bildungspanel`

3 NEPScomp

Das LIfBi fördert die bildungswissenschaftliche Längsschnittforschung in Deutschland durch Bereitstellung einer grundlegenden, überregional und international bedeutsamen, forschungsbasierten Infrastruktur. Wesentlicher Bestandteil dieser Infrastruktur sind die aufbereiteten und dokumentierten Erhebungsdaten des NEPS, die vom Forschungsdatenzentrum (FDZ-LIfBi) in Form von Scientific-Use-Files kostenfrei angeboten werden.[1] Diese regulären Scientific-Use-Files bilden den Ausgangspunkt von NEPScomp.

Der Zugang zu den Daten basiert auf einem Datennutzungsvertrag, der das gesamte NEPS-Datenportfolio einschließt (siehe Kapitel 5). Seit der Veröffentlichung des ersten Scientific-Use-Files im Jahr 2011 bis zum Zeitpunkt der erstmaligen Erstellung dieses Manuals zum 1. September 2025 wurden 3.085 solcher Forschungsvorhaben und 4.070 Personen als Datennutzende registriert.

→ www.lifbi.de > Daten & Services > Datenzugang > NEPS – Forschungsprojekte

Die Arbeit mit den regulären Scientific-Use-Files erfordert ein vertieftes Wissen im Umgang mit komplexen Datenstrukturen. Gerade die erstmalige Nutzung der NEPS-Daten bzw. die Verwendung der Daten im Kontext von Seminaren, Forschungspraktika oder Bachelor- und Masterarbeiten geht mit zwei wesentlichen Herausforderungen einher: Zum einen steht in aller Regel nur ein begrenzter Zeitraum für die Bearbeitung einer Fragestellung zur Verfügung. Zum anderen ist die Erfahrung bezüglich der Handhabung derartiger Sekundärdaten üblicherweise noch limitiert. Beide Faktoren – Zeit zum Verstehen und Aufbereiten der Daten sowie Kenntnisse zum Treffen der richtigen Entscheidungen – sind jedoch zentral für die Nutzung der regulären Scientific-Use-Files des NEPS.

Um Einsteiger:innen dennoch die Möglichkeit für eigene Auswertungen auf Basis der NEPS-Daten zu geben, wurde **NEPScomp als zusätzliches Angebot** entwickelt. Im Vergleich zu den regulären Scientific-Use-Files zeichnet sich NEPScomp durch eine stark vereinfachte und reduzierte Datenstruktur aus. Gleichzeitig erlauben die Daten inhaltlich aussagefähige Analysen zu einem breiten Spektrum an Themen und einer Vielzahl an Fragestellungen. Dabei bleibt die Verbindung zu den umfassenderen Scientific-Use-Files der einzelnen Startkohorten gewahrt, so dass ein Wechsel jederzeit und ohne größeren Aufwand möglich ist. In den folgenden Abschnitten wird der Aufbau von NEPScomp näher erläutert. Besonderes Augenmerk liegt hierbei auf den Vereinfachungen gegenüber den regulären Scientific-Use-Files und deren Implikationen für das Analysepotenzial (siehe auch die Zusammenfassung in Kapitel 3.5).

1 Das NEPS-Datenportfolio besteht aus den sieben Panelstudien der einzelnen Startkohorten (siehe Kapitel 2). Mit jeder neuen Panelwelle einer Startkohorte wird der entsprechende Scientific-Use-File aktualisiert und um die neuen Erhebungsdaten ergänzt (*kumulierte Releases*). Hinzu kommen zwei Zusatzstudien, die sich mit der organisatorischen Reform der Oberstufe in Thüringen (zwei Erhebungen von 2010 bis 2011) und der G8-Reform in Baden-Württemberg (drei Erhebungen von 2011 bis 2013) befassen, aber nicht Teil von NEPScomp sind.

3.1 Erzeugung von NEPScomp

NEPScomp basiert auf den regulären Scientific-Use-Files der einzelnen Startkohorten in der zum Veröffentlichungszeitpunkt jeweils aktuellen Version (siehe auch die Auflistung in Kapitel 2). Im ersten Release war die Startkohorte 8 (Klasse 5 – 2022) zunächst nicht enthalten, da hier zum Veröffentlichungszeitpunkt noch keine Längsschnittinformationen im regulären Scientific-Use-File vorlagen. Seit dem Release 2.0.0 beinhaltet NEPScomp auch die aktuellen Daten dieser neuen Startkohorte (siehe auch Hinweis in den Release Notes, Kapitel A.5).

NEPScomp wird im Unterschied zu den regulären Scientific-Use-Files der einzelnen Startkohorten als integrierter Datenbestand mit sämtlichen Startkohorten in einem gemeinsamen Paket angeboten. Die Aufteilung nach Dateiformaten – **SPSS und Stata** – dient vor allem der Speichereffizienz; inhaltlich sind beide Datenbestände völlig identisch.[2] Entsprechend existiert auch nur ein einziger *Digital Object Identifier* (DOI), unter dem dieser Datenbestand registriert und der bei der Datenzitation zu berücksichtigen ist (siehe Kapitel 5.2).

NEPScomp stellt eine direkte Ableitung aus den **Download-Versionen** der regulären Scientific-Use-Files dar. Sensitive Merkmale, die sonst nur in den geschützten Datenzugangsumgebungen einer kontrollierten Datenfernverarbeitung (*RemoteNEPS*) oder eines Datensicherheitsraum im LIfBi (*On-site*) zugänglich sind, bleiben bei NEPScomp vollständig außen vor. Dies schließt in erster Linie geografische Zuordnungen wie Bundesland (außer in den Startkohorten 1 und 6), Regierungsbezirk und Kreis aus, aber auch Informationen zu den Bildungseinrichtungen der Zielpersonen (insbesondere aus den Befragungen der Schul- und Kindergartenleitungen), länderspezifische Angaben zu Herkunft, Staatsangehörigkeit und Sprachen sowie einzelne Merkmale mit einem erhöhten Reidentifikationsrisiko. Eine vollständige Liste der betreffenden Variablen findet sich in den jeweiligen Anonymisierungsreports der Startkohorten (siehe Kapitel 4.2).

NEPScomp nutzt weitestgehend die gleichen Metadaten (=Beschreibung der Daten, z. B. Variablennamen und Labels) wie die regulären Scientific-Use-Files. Die wesentlichen Abweichungen finden sich aufgrund der Anpassung des Datenformats im Kompetenzdatensatz und im Statusdatensatz (siehe Kapitel 3.4). Alle Daten sind mit **deutsch- und englischsprachigen Variablen- und Wertelabels** ausgestattet. Während die beiden Sprachversionen für SPSS als separate Datenpakete in NEPScomp enthalten sind, erfolgt der Wechsel zwischen den Sprachen in Stata direkt innerhalb eines Datensatzes über das Kommando:

```
label language [de/en]
```

2 Nutzende der Statistiksoftware R können beide Version über entsprechende R-Pakete laden (siehe auch Kapitel A.2).

NEPScomp übernimmt auch die beiden zentralen **Identifikator-Variablen** aus den regulären Scientific-Use-Files. Diese sind:

- **`ID_t`** = Kennung der Zielperson (*target*)
- **`wave`** = Kennung des Messzeitpunkts (*Erhebungswelle*)

Weitere Identifikatoren sind für den Umgang mit den NEPScomp-Daten nicht erforderlich, da die Datenstruktur ein Zusammenspielen von Informationen aus verschiedenen Datensätzen einer Startkohorte über einen einfachen 1:1-Merge erlaubt (siehe das Beispiel in Abbildung 3). Grundsätzlich erlaubt die Beibehaltung der Personen-ID auch das Anspielen zusätzlicher Informationen aus den Datensätzen der regulären NEPS-Daten an NEPScomp.

3.2 Datenstruktur von NEPScomp

NEPScomp beinhaltet für jede Startkohorte **acht Datensätze**, während ein regulärer Scientific-Use-File in der Download-Variante zwischen 19 (Startkohorte 2) und 41 (Startkohorte 4) Datensätze umfasst. Die Reduzierung ist sowohl auf eine Auswahl von Inhalten als auch auf eine Umstrukturierung der Daten zurückzuführen. Die NEPScomp-Datensätze sind nach Themen sortiert, d. h. alle Variablen in einem Datensatz gehören zum gleichen inhaltlichen Kernbereich. Dies vereinfacht das Auffinden geeigneter Variablen für die jeweilige Analyse und minimiert die Notwendigkeit des Zusammenspielens von Informationen aus verschiedenen Datensätzen. Zu beachten ist, dass im Unterschied zu den regulären Scientific-Use-Files bestimmte Variablen identisch in verschiedenen Datensätzen einer Startkohorte enthalten sein können. Dies ist immer dann der Fall, wenn die entsprechende Variable unterschiedlichen Oberkategorien im Themenverzeichnis der Online-Variablensuche (siehe Kapitel 4.4) zugeordnet ist. In Anlehnung an dieses Themenverzeichnis und die Rahmenkonzeption des NEPS (siehe Abbildung 2) sind folgende Datensätze je Startkohorte in NEPScomp verfügbar:

- Kompetenzen und Lernergebnisse = `Competency`
- Lernumwelten = `LearningEnv`
- Bildungsbezogene Ressourcen, Einstellungen und Verhalten = `Resources`
- Demografie und Biografie = `Demography`
- Individuum und Persönlichkeit = `Personality`
- Migration und kulturelle Diversität = `Migration`
- Gesellschaft, Familie und Teilhabe = `Society`
- Episodenstatus = `States`

Der Dateiname eines Datensatzes verweist auf den jeweiligen Themenbereich, aber auch auf die Startkohorte (`SC#`), den Datenbestand (`NEPScomp`) und die Versionsnummer der Datenveröffentlichung (`#-#-#`). So finden sich im Datensatz `SC2_NEPScomp_Competency_2-0-0` die Kompetenzvariablen der Startkohorte 2 des aktuellen NEPScomp-Releases.

Für die **Variablen** gilt festzuhalten, dass nur inhaltlich einem der genannten Themenbereiche zugeordnete Merkmale im NEPScomp-Datenangebot enthalten sind. Wenn eine Variable mehreren Themenbereichen zugeordnet ist, so wird diese Variable identisch in die jeweiligen Datensätze übernommen.[3] Die Variablen in NEPScomp enthalten in aller Regel Informationen, die von einer Zielperson selbst oder von einem Elternteil stammen.[4] Auf die in den regulären Scientific-Use-Files übliche Differenzierung der Datensätze nach Auskunftspersonen wird in NEPScomp zugunsten einer kompakten Darstellung von inhaltlich verwandten Informationen verzichtet. Dadurch wird die Unterscheidung zwischen `t`- und `p`-Variablen umso wichtiger:

- **`t`** als erster Buchstabe im Variablennamen = Auskunft von einer Zielperson (*target*)
- **`p`** als erster Buchstabe im Variablennamen = Auskunft von einem Elternteil (*parent*)

Daneben existieren weitere Konventionen für die Benennung von Variablen im NEPS-Kontext: Während die erste Ziffer dem Typus der Auskunftsperson vorbehalten ist, verweist die zweite Ziffer auf den jeweiligen Themenbereich (z. B. „3“ für Bildungsentscheidungen). Die weiteren Ziffern kennzeichnen die Variablennummer. Manche Variablen weisen zudem ein Suffix auf, welches durch einen Unterstrich abgetrennt ist. Das häufigste Suffix ist „_g#“. Es zeichnet jene Variablen aus, die nicht unmittelbar erhoben, sondern im Zuge der Datenedition generiert wurden. Typische Anwendungsfälle sind abgeleitete Berufs- und Bildungsklassifikationen sowie aus mehreren Einzelitems erzeugte Indizes wie die Persönlichkeitsdimensionen der „Big Five“. Eine detaillierte Darstellung aller Konventionen zur Bezeichnung der NEPS-Variablen kann in den separaten Data Manuals der NEPS-Startkohorten nachgeschlagen werden (siehe Kapitel 4.2).

Ein zentraler Vorteil der NEPScomp-Datenstruktur betrifft die Vereinheitlichung auf das **LONG-Datenformat**. In den regulären Scientific-Use-Files der NEPS-Startkohorten finden sich unterschiedliche Formate – neben dem für Paneldaten üblichen *LONG*-Format auch das zumeist für Querschnittdaten genutzte *WIDE*-Format und das speziell für Lebensverlaufs- bzw. Episodendaten typische *SPELL*-Format. Jedes Format folgt einer anderen Logik der Ablage der Erhebungsdaten, wodurch insbesondere das Zusammenspielen von Informationen aus unterschiedlich strukturierten Datensätzen zu einer erheblichen Herausforderung wird. Um diese Hürde zu vermeiden, wurden alle NEPScomp-Datensätze auf das LONG-Datenformat standardisiert.

- Die Daten der einzelnen Erhebungswellen sind nicht als separate Datensätze abgelegt. Stattdessen werden die Daten einer neu hinzukommenden Erhebungswelle an die bereits exis-

3 Ein Beispiel ist die Variable `th18083` („Lernen/Arbeiten zuhause, technische Kenntnisse Lehrkräfte“), die in der Startkohorte 6 mehreren Themenbereichen zugeordnet ist und identisch in den Datensätzen `Competency`, `Demography`, `LearningEnv` und `Society` vorkommt.

4 Vereinzelt gibt es auch Angaben von den Lehrkräften zu den Zielkindern in NEPScomp. Diese Informationen kommen aus sog. „schülerindividuellen Einschätzbögen“ und sind durch das vorangestellte „`e`“ im Variablennamen identifizierbar (z. B. `e66600a` zur voraussichtlichen Übertrittsempfehlung in der Startkohorte 2).

tierenden Datensätze „angehangen“. Das Ergebnis sind integrierte Paneldatensätze mit den Angaben von mehreren Personen aus mehreren Erhebungswellen.

- Eine Zeile in einem Datensatz enthält die Angaben (zu) genau einer Zielperson aus genau einer Erhebungswelle.
- Wird eine Information von einer Person über mehrere Erhebungswellen hinweg wiederholt erhoben (z. B. Lebenszufriedenheit), dann sind die entsprechenden Angaben dieser Person in einer Variable – verteilt über mehrere Zeilen – gespeichert (z. B. `t514001` im Datensatz `SC4_NEPScomp_Personality`).
- Die Datenstruktur enthält Zellen für jede Variable und jede Erhebungswelle. Da nicht alle Fragen in jeder Erhebungswelle gestellt wurden, sind auch nicht für alle Zellen gültige Werte vorhanden. In diesen Fällen weist NEPScomp jeweils System-Missings aus (siehe Kapitel 3.3).
- Um eine einzelne Datenzeile eindeutig bestimmen zu können, bedarf es in den NEPScomp-Datensätzen der Kombination aus Personen-ID (`ID_t`) und Wellenidentifikator (`wave`). Dies ist vor allem für das Zusammenspielen von Informationen aus verschiedenen Datensätzen von zentraler Bedeutung (=*Merging*).

NEPScomp bietet hier den großen Vorteil, dass Variablen aus verschiedenen Datensätzen einer Startkohorte **1:1 kombiniert** werden können – ohne weitere Datenaufbereitung. Ein aufwändiges Umformatieren von Daten im Vorfeld des Zusammenspielens entfällt.

Abbildung 3: Beispiel für Merging-Syntax in Stata

```
** Datenpfad für NEPScomp auf lokalem Rechner definieren => z. B. Z:/Data
global datapath Z:/Data/NEPScomp_D_Stata_2-0-0/SC6

** Masterdatensatz öffnen => Society
use "${datapath}/SC6_NEPScomp_Society_2-0-0.dta", clear

** Anspielen der Zielvariablen => Geschlecht
** aus dem Using-Datensatz => Demography
** über Schlüsselvariablen => ID_t, wave
merge 1:1 ID_t wave using "${datapath}/SC6_NEPScomp_Demography_2-0-0.dta", ///
        keepusing(t700001) nogen

** Messzeitpunkte von Religiosität => Wellen 3, 7, 11, 15
tab wave t435000

** Welle für Querschnittanalyse auswählen => Welle 3 (=2. NEPS-Haupterhebung)
keep if wave==3

** Geschlechtsunterschiede prüfen => T-Test
ttest t435000, by(t700001)
```

Das Stata-Beispiel in Abbildung 3 illustriert für die Startkohorte 6 das Anspielen der Variable Geschlecht (`t700001`) aus dem `Demography`-Datensatz an den `Society`-Datensatz, um Un-

terschiede zwischen Männern und Frauen hinsichtlich ihrer Religiösität (`t435000`) querschnittlich analysieren zu können. Der eigentliche Merging-Prozess beschränkt sich auf einen einzigen Befehl (analog für SPSS in Kapitel A.1 und für R in Kapitel A.2).

Die Vereinheitlichung des Datenformats hat allerdings **Beschränkungen** zur Folge. Zum einen berücksichtigt NEPScomp keine gruppenbezogene Informationen. Damit fehlen insbesondere die Angaben von Lehrkräften zu den Klassen der Zielkinder. Mehrebenenanalytische Auswertungen sind nicht möglich.[5] Zum anderen verzichtet NEPScomp weitgehend auf episodenspezifische Informationen, was weitere Auswertungsverfahren ausschließt (z. B. Sequenzdatenanalyse). Stattdessen enthält der NEPScomp-Datensatz `States` mehrere Variablen zum jeweils aktuellen Status Quo und zur Dauer von ausgewählten Merkmalen oder Charakteristika im Lebensverlauf, die aus den entsprechenden Episodendaten generiert wurden. Die Umformatierung der ursprünglichen Kompetenzdaten hat zur Folge, dass einzelne Variablengruppen in NEPScomp ausgeschlossen sowie andere Variablengruppen in neuen Variablen zusammengeführt wurden (siehe Kapitel 3.4). Um diese Beschränkungen zu umgehen, ist der Rückgriff auf die regulären Scientific-Use-Files der einzelnen NEPS-Startkohorten erforderlich.

3.3 Definition fehlender Werte in NEPScomp

Die regulären Scientific-Use-Files des NEPS verfügen über bis zu zwanzig verschiedene Codes für fehlende Werte, die allesamt durch negative Werte gekennzeichnet sind. Sie reichen von „weiß nicht“ (-98) und „verweigert“ (-97) über „trifft nicht zu“ (-93) und „überfiltert“ (-54/-99) bis hin zu „Befragung abgebrochen“ (-91) oder itemspezifischen Ausfallgründen (-20).

NEPScomp weist dagegen alle fehlenden Werte als **System-Missing (.)** aus. Durch diese Vereinheitlichung entfällt die Notwendigkeit einer Spezifikation von Werten, die vorab von der jeweiligen Analyse ausgeschlossen werden sollen. System-Missings werden von SPSS, Stata und R automatisch als fehlende Werte erkannt. Mit der Aufhebung der ursprünglichen Differenzierung geht jedoch ein Informationsverlust einher. Demnach besteht keine Möglichkeit, ein „leeres“ Feld in der Datenmatrix einem bestimmten Ausfallgrund zuzuordnen. Es lässt sich nicht mehr nachvollziehen, ob ein fehlender Wert auf der Reaktion einer Befragungsperson basiert (z. B. „weiß nicht“) oder ob die Frage der Person gar nicht gestellt wurde (z. B. „überfiltert“). Für bestimmte Analysen, insbesondere methodologischer Natur, ist diese Unterscheidung von wesentlicher Bedeutung. Hier muss auf die regulären Scientific-Use-Files der NEPS-Startkohorten zurückgegriffen werden. In den meisten Analysen werden jedoch sämtliche Missings pauschal ausgeschlossen – und da erleichtert NEPScomp den Umgang mit fehlenden Werten.

System-Missings wurden bei NEPScomp außerdem eingesetzt, um komplett fehlende Erhebungswellen von Studienteilnehmenden „aufzufüllen“. In Panelstudien ist es ein übliches Phänomen, dass Teilnehmende für eine bzw. mehrere Wellen aussetzen (*temporäre Ausfälle*) oder ihre Mitwirkung noch vor Abschluss der Studie ganz beenden (*endgültige Ausfälle*). In den regulären

5 In den Scientific-Use-Files der jeweiligen Startkohorten finden sich die entsprechenden Informationen in den separaten Datensätzen `pEducator` und `pCourse[Class/German/Math]`.

Scientific-Use-Files existieren zu diesen Personen keine Datenzeilen für die fehlenden Wellen. NEPScomp verfolgt an dieser Stelle einen abweichenden Ansatz. Hat eine Zielperson an einer Welle nicht teilgenommen, dann bleibt die entsprechende Datenzeile für die jeweilige Person (`ID_t`) in der betreffenden Welle (`wave`) im Datensatz erhalten. Alle Werte in dieser Zeile bestehen dann aus System-Missings. Dadurch weist jede Startkohorte die gleiche Fallzahl über die verschiedenen Wellen hinweg auf. Für das Zusammenspielen von Informationen aus unterschiedlichen Datensätzen ergibt sich somit immer ein perfekter 1:1-Match. Bezüglich der Auswertung der Daten hat diese Modifikation keine Folgen, da die als System-Missing definierten Werte per se von den Analysen ausgeschlossen werden.

Eine weitere Änderung hinsichtlich der Definition fehlender Werte wurde bei ausgewählten Merkmalen vorgenommen, für die eine zeitliche Kontinuität angenommen werden kann. Dies betrifft insbesondere soziodemografische Informationen im Datensatz `Demography`, wie z. B. das Geburtsland, das Geburtsjahr oder das Geschlecht – hier sind intraindividuelle Veränderungen über die Zeit nicht möglich oder zumindest sehr selten. In den regulären Scientific-Use-Files weisen die jeweiligen Variablen nur für diejenigen Wellen einen gültigen Wert in den Daten auf, in denen diese Information tatsächlich erhoben wurde (z. T. auch nur einmalig). In den übrigen Wellen findet sich dagegen ein Missing-Code. Gerade für Querschnittanalysen kann diese Struktur problematisch sein, da unter Umständen Informationen aus anderen Wellen herangezogen werden müssen. Wurde beispielsweise das Geburtsjahr nur einmalig in der ersten Welle erfasst und man möchte dieses Merkmal im Rahmen einer Querschnittanalyse mit Daten der dritten Welle nutzen, dann müsste diese Information zunächst für jede Person aus der Zeile der ersten Welle in die Zeile der dritten Welle übernommen werden, was einige zusätzliche Schritte an Datenaufbereitung erfordert. NEPScomp erspart diese Prozedur, da die jeweils letztgültigen Angaben der Befragungsteilnehmenden für die in Kapitel A.4 aufgeführten Merkmale in alle vorherigen und folgenden Wellen kopiert sind. Um in dem Beispiel zu bleiben: Auch wenn die Person nach der ersten Welle nicht noch einmal nach ihrem Geburtsjahr gefragt wurde, so findet sich dennoch in der entsprechenden Variable für jede Welle ein gültiger Wert – nämlich die Angabe aus der ersten Erhebungswelle.[6] Durch diese Umstellung ist in den NEPScomp-Daten zwar nicht mehr direkt ersichtlich, in welcher Welle das jeweilige Merkmal abgefragt wurde. Allerdings stehen für diese Information andere Materialien zur Verfügung, insbesondere die Online-Variablensuche (siehe Kapitel 4.4) und die Erhebungsinstrumente der jeweiligen Startkohorte (siehe Kapitel 4.2).

An dieser Stelle sei noch darauf hingewiesen, dass NEPScomp komplett ohne „Kürzungen“ bezüglich der Anzahl gültiger Fälle in den Daten auskommt. Der Datenbestand verfügt insofern über die gleiche statistische Power wie die regulären Scientific-Use-Files der einzelnen Startkohorten.

6 Dies bedeutet nicht, dass die entsprechenden Merkmale in den Daten komplett konstant über die Zeit sind. So finden sich beispielsweise Änderungen der Geschlechtszugehörigkeit einer Person in den Daten. Dies kommt dadurch zustande, dass ein Teil dieser Merkmale in größeren Abständen mehrfach abgefragt wurde. Identifiziert sich eine Person in Welle 1 als „männlich“ und bei einer erneuten Abfrage der Geschlechtszugehörigkeit in Welle 6 als „weiblich“, dann findet sich in den NEPScomp-Daten für diese Person die Ausprägung „männlich“ in den Wellen 1 bis 5 und „weiblich“ ab der sechsten Welle.

3.4 Umformatierte Datensätze in NEPScomp

Mit der Umstellung auf ein einheitliches LONG-Datenformat mussten insbesondere die sonst nur im WIDE-Format vorliegenden Kompetenzdaten und die im SPELL-Format vorliegenden Episodendaten für NEPScomp zusätzlich aufbereitet werden. In beiden Fällen wurden die Variablen neu generiert, d. h. sie existieren in dieser Form bzw. unter diesen Namen nur in NEPScomp.

Kompetenzdaten

Die Entwicklung von Kompetenzen im Lebensverlauf ist ein zentraler Bestandteil des Nationalen Bildungspanels. Die Testung von Kompetenzen erstreckt sich in allen Startkohorten über mehrere Wellen hinweg und umfasst verschiedene Inhaltsbereiche (*Domänen*) wie Lesen, Mathematik, Naturwissenschaften. Eine vollständige Übersicht der durchgeführten Kompetenzmessungen im Zeitverlauf einer jeden Startkohorte findet sich in Kapitel A.3.

In den regulären Scientific-Use-Files sind die aus den Testungen resultierenden Daten im sog. WIDE-Format abgelegt. Das heißt, dass pro Fall nur eine einzige Datenzeile existiert und der jeweilige Erhebungszeitpunkt einer Messung über den Variablennamen erkennbar ist. Hintergrund für dieses Datenformat ist die Tatsache, dass nur ein geringer Teil der verfügbaren Kompetenzvariablen unmittelbar für längsschnittliche Panelanalysen geeignet ist. Für den überwiegenden Teil der Variablen sind, entweder aufgrund einer einmaligen Erhebung oder aus messtheoretischen Gründen, nur querschnittliche Auswertungen zulässig.[7] Um dies den Datennutzenden deutlich zu machen und keine unzulässigen Analysen zu suggerieren, wurden sämtliche Kompetenzdaten in das „querschnittliche“ WIDE-Format geschrieben.

NEPScomp bietet die Kompetenzdaten dagegen im einheitlichen LONG-Format an. Variablen mit inhaltlich äquivalenten Informationen wurden hierfür zu einer gemeinsamen Variable zusammengeführt. Zudem wurde eine Auswahl an Variablen getroffen, bei der es sich um Indizes (*Scores*) handelt, die aus den Einzelitems einer Testung gebildet sind und die eine Aussage über das Abschneiden im Gesamttest erlauben. Je nach Domäne sind das WLE-Schätzer (*Weighted Likelihood Estimates*), Summenwerte, Mittelwerte, Differenzmaße oder Anteilswerte. Die Einzelitems selbst sind nicht Bestandteil von NEPScomp.

Um ein Beispiel zu geben: In der Startkohorte 4 wurde zu drei verschiedenen Messzeitpunkten (Wellen 1, 7, 10) die mathematische Kompetenz der Studienteilnehmenden getestet. Die Ergebnisse zu den einzelnen Aufgaben (korrekt gelöst vs. nicht korrekt gelöst) finden sich nur im regulären Scientific-Use-File der Startkohorte 4. Die aus den Einzelaufgaben geschätzten Gesamtscores (3x WLE korrigiert, 3x WLE unkorrigiert) wurden dagegen für NEPScomp in jeweils eine Variable im LONG-Format überführt (1x WLE korrigiert = `ma_sc1`, 1x WLE unkorrigiert =

7 Zu beachten ist, dass es in aller Regel nicht sinnvoll ist, einen Test in identischer Form über mehrere Jahre bei den gleichen Personen wiederholt einzusetzen. Durch die alters- bzw. entwicklungsspezifische Konzipierung von Tests – z. B. mit angepasstem Schwierigkeitslevel – ist eine unmittelbare Vergleichbarkeit der Einzelvariablen über die Zeit nicht gegeben.

ma_sc1u). Die Umstrukturierung erforderte zudem eine Anpassung der Variablennamen gegenüber den regulären Scientific-Use-Files, da sich der Zeitbezug im LONG-Format aus der Kombination mit der Wellenvariable wave ergibt. Die Zuordnung von Variablen zu Domänen und zur Art des Scorings ist jedoch weiterhin über den Variablennamen möglich (siehe Kapitel A.3).

WICHTIG: Bei der Arbeit mit NEPScomp ist zu beachten, dass bestimmte Kompetenzvariablen trotz des Vorliegens von wiederholten Messungen und trotz der Ablage im LONG-Format **nicht** für Längsschnittanalysen geeignet sind. Dies betrifft insbesondere die korrigierten WLE-Schätzer, die durch das Suffix _sc1 gekennzeichnet sind.[8] Die betreffenden Variablen lassen sich durch einen deutlichen Hinweis im Variablenlabel (*QUERSCHNITT*) identifizieren.

Für alle durchgeführten Tests existieren umfangreiche Dokumentationsmaterialien (siehe Kapitel 4.2). Die original eingesetzten Testinstrumente („Aufgaben“) stehen aus Gründen der Testsicherheit allerdings nicht frei zur Verfügung.

Statusdaten

Ein ebenfalls zentrales Merkmal des Nationalen Bildungspanels ist die präzise Erfassung von Ereignissen (*Spells*) in relevanten Bereichen der individuellen Lebensverläufe der Studienteilnehmenden. Die daraus resultierenden Episodendaten dokumentieren unterschiedliche biografische Phasen – von Bildungswegen über Erwerbstätigkeiten und Weiterbildungsaktivitäten bis hin zu Partnerschafts- und Familienereignissen. In den regulären Scientific-Use-Files liegen diese Informationen als sog. SPELL-Datensätze vor. Sie enthalten monatsgenaue Angaben zum Beginn und Ende eines Ereignisses sowie Zusatzinformationen (sog. „Deskriptoren“ wie z. B. die Arbeitszeit in einem angegebenen Job). Das SPELL-Format unterscheidet sich von den Datenformaten LONG und WIDE insofern, als dass jedes Ereignis pro Fall eine eigene Datenzeile beansprucht. Treten zwischen zwei Erhebungszeitpunkten mehrere Ereignisse auf, beispielsweise durch einen Wechsel des Arbeitsplatzes, wird jedes dieser Ereignisse bzw. Spells in einer separaten Zeile erfasst. Dadurch kann es für dieselbe Person und denselben Messzeitpunkt mehrere Datenzeilen geben. Der korrekte Umgang mit Spelldaten stellt eine besondere Herausforderung dar. Aus diesem Grund verfügen bereits die regulären Scientific-Use-Files über zusätzlich aus den Einzelinformationen generierte (Hilfs)Variablen und (Hilfs)Datensätze, die die Arbeit mit den biografischen Informationen erleichtern sollen.

Für NEPScomp wurde der Informationsgehalt der ursprünglichen Episodendaten im Zuge der Vereinheitlichung der Datenstruktur deutlich reduziert. Der Datensatz States beinhaltet neu generierte Variablen, die für ausgewählte Ereignisse jeweils den aktuellen Status zum Erhebungszeitpunkt (entspricht dem Datum des Interviews) und die Dauer des Ereignisses (entspricht der Anzahl an Monaten von Beginn der Episode bis zum Zeitpunkt des Interviews) angeben. Je nach Startkohorte und damit Alter der Zielpersonen sind zum Teil unterschiedliche Ereignisse in den Daten vertreten. So beinhaltet NEPScomp für die jüngeren Kohorten beispielsweise Angaben zu Kinderbetreuungsepisoden, während in den älteren Kohorten Informationen zu Bildung, Ausbildung, Erwerbstätigkeiten und Elternzeiten vorliegen. Die beiden neu generierten Merkmalstypen sind durch ein Suffix im Variablennamen gekennzeichnet:

8 Das _sc steht in den Competency-Datensätzen als Abkürzung für „score“, nicht für eine bestimmte Startkohorte (zur Bedeutung der verschiedenen Suffixes in den Kompetenzdaten siehe Kapitel A.3).

- **_g1** = aktueller Status eines biografischen Ereignisses
- **_g2** = Dauer dieses aktuellen Status seit Beginn (in Monaten)

Die Beschränkung auf einzelne Ereignistypen mit jeweils zwei Merkmalstypen ist gleichbedeutend mit dem Verzicht auf einen Großteil der in den ursprünglichen Episodendaten verfügbaren Informationen. Der bereitgestellte Ausschnitt spiegelt zudem bestimmte (Vor)Entscheidungen bei der Aufbereitung der Daten wider, zum Beispiel die Auswahl einer primären Erwerbstätigkeit im Falle paralleler Jobepisoden. Für vertiefende Analysen zu biografischen Verläufen ist daher die Nutzung der regulären Scientific-Use-Files inkl. Dokumentation dringend zu empfehlen.

3.5 Zusammenfassung

Tabelle 1: Zentrale Unterschiede zwischen regulären Scientific-Use-Files und NEPScomp

	Reguläre SUF	NEPScomp
Startkohorten	SC1 bis SC6 + SC8: N=7	SC1 bis SC6 + SC8: N=7
Datensätze	startkohortenspezifisch: N=19-41	einheitlich: N=8 pro Startkohorte
- Sortierung	nach verschiedene Kriterien	nach inhaltlichen Themen
- Variablen	keine Mehrfachzuordnung zu verschiedenen Datensätzen	mit Mehrfachzuordnung zu verschiedenen Datensätzen
Datenformat	datensatzspezifisch: LONG, WIDE, SPELL	einheitlich: LONG
Identifikatoren	datensatzspezifisch: ID_t, wave, ID_i, ID_e, ID_cc, splink, ID_int etc.	einheitlich: ID_t, wave
Merging	datensatzspezifisch: 1:1, 1:m etc.	einheitlich: 1:1
Fehlende Werte	bis zu 20 Kodes: negative Werte	einheitlich: System-Missing
Reduzierung	keine	Episodendaten Sensible Informationen Methodeninformationen Einzelitems Kompetenztestung Aggregatdaten zu Klassen, Schulen, Kindergärten
Datenzugang	Download, Remote, On-site	nur Download

4 Datendokumentation

4.1 Dokumentation zu NEPScomp

Die unmittelbare Dokumentation des NEPScomp-Datenbestands umfasst dieses Datenmanual, das Codebuch und die Release Notes:

Datenmanual Das vorliegende NEPScomp-Manual fungiert als allgemeine Beschreibung und zentraler Einstieg in die Arbeit mit NEPScomp. Es stattet die Datennutzenden mit wichtigen Hinweisen zum Umgang mit den Daten sowie mit Links zu ergänzenden Informationsquellen aus. Eine sorgfältige Lektüre vor Beginn der geplanten Analysen empfiehlt sich in jedem Fall.

Codebuch Das Codebuch stellt eine tabellarische Übersicht bereit, die sowohl die Struktur der NEPScomp-Daten mit allen Datensätzen und Variablen (inkl. Variablenname, Variablenlabel und Fragetext) als auch wellenspezifische Grundauszählungen bzw. Häufigkeitsverteilungen zu jeder Variable (inkl. Wertelabel) enthält.

Release Notes Jede Veröffentlichung einer neuen Datenversion wird von sog. „Release Notes“ begleitet. Diese Notizen dokumentieren sowohl Änderungen am Datenbestand gegenüber der jeweils vorhergehenden Datenversion als auch festgestellte Fehler in den Daten – einschließlich Korrekturhinweisen, sofern dies möglich ist. Zwischenzeitlich zur Kenntnis gelangte Probleme mit den Daten werden fortlaufend in den Release Notes auf der Webseite protokolliert. Die zum Zeitpunkt der Veröffentlichung von NEPScomp aktuelle Liste findet sich auch im Anhang dieses Datenmanuals (siehe Kapitel A.5).

Tutorial Im Anschluss an den ersten NEPScomp-Release im Jahr 2025 wurden vom FDZ-LIfBi zwei Video-Präsentationen zur Vorstellung dieses Datenangebots erstellt. Die Kurzfassung fokussiert ausschließlich auf NEPScomp; die Langfassung des Tutorials geht zusätzlich auf die regulären Scientific-Use-Files der einzelnen NEPS-Startkohorten ein.

Alle Dokumentationsmaterialien sind über die entsprechende Webseite abrufbar und werden dort auch regelmäßig aktualisiert:

→ `www.lifbi.de > Daten & Services > Daten und Dokumentation > NEPScomp`

4.2 Weiterführende Dokumentation

Um die Erhebung der Daten des Nationalen Bildungspanels nachvollziehen zu können, ist ein Rückgriff auf die umfangreiche Dokumentation zu den regulären Scientific-Use-Files der einzelnen NEPS-Startkohorten unumgänglich. Neben den eingesetzten Fragebögen geben insbesondere die Methodenberichte einen detaillierten Einblick in den Ablauf der Feldarbeit einer jeden

Panelwelle. Darüber hinaus informieren verschiedene Berichte über relevante Aspekte der Datenaufbereitung. Abbildung 4 und die nachfolgenden Erläuterungen vermitteln einen Eindruck vom **allgemeinen Angebot an NEPS-Datendokumentation**. Die aufgeführten Materialien beziehen sich auf die regulären Scientific-Use-Files und sind zumeist startkohortenspezifisch. Im Kontext von NEPScomp stellen sie eine Grundlage zur Vertiefung des Hintergrundwissens dar.

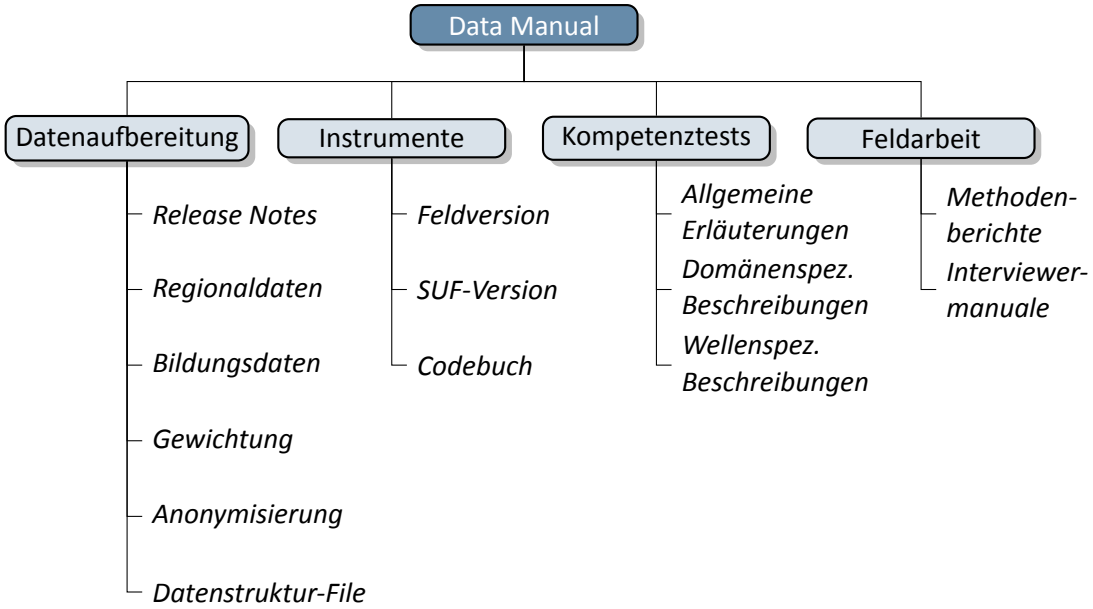

Abbildung 4: NEPS-Dokumentatonsmaterialien

- Das Data Manual, das Codebuch und die Release Notes erfüllen für die Scientific-Use-Files der einzelnen NEPS-Startkohorten die gleiche Funktionen wie für NEPScomp (siehe Kapitel 4.1).
- Bei den Erhebungsinstrumenten wird zwischen Feldversion und SUF-Version unterschieden. Für jede Welle in jeder Startkohorte können die original eingesetzten Feldinstrumente als Fragebögen bzw. Programmiervorlagen von der Webseite abgerufen werden. Diese Dokumente liegen nur in deutscher Sprache vor. Die SUF-Instrumente sind hingegen auch in englischer Sprache verfügbar. Zudem bieten sie ergänzende Informationen mit konkretem Bezug zum jeweiligen Scientific-Use-File (z. B. SUF-Variablennamen und -Labels, Zuordnung Datensatz).
- Die zur Testung der Kompetenzen administrierten Aufgaben sind nicht per Download zugänglich. Dafür bieten sowohl allgemeine Berichte (zu Erhebungszeitpunkten und Namenskonventionen, Skalierungsmodell, längsschnittlicher Verknüpfung) als auch wellenspezifische Berichte (zu getesten Domänen inkl. Testsituation, Ablauf, Dauer, Anzahl an Aufgaben, Erhebungsmodus) und domänenspezifische Berichte (zu Konstruktbeschreibungen, Testverfahren, Beispielitems, psychometrischen Eigenschaften und Skalierung) einen umfassenden Einblick in die Kompetenzmessungen des NEPS.
- Die Methodenberichte und Interviewerhandbücher dienen in erster Linie dem Verständnis des Ablaufs der Datenerhebungen (*Feldarbeit*). Diese Dokumente sind von den durchführen-

den Erhebungsinstituten erstellt. Sie dokumentieren die Vorbereitung einer Befragung (z. B. Schulung der Interviewer:innen, Rekrutierung der Befragungsteilnehmenden), deren Durchführung (z. B. Einsatz der Interviewer:innen, Kontaktierung der Befragungsteilnehmenden, Anreizstrategie) und die Ergebnisse (z. B. Realisierung der Stichprobe, Datenqualität).

- Eine Reihe von Einzeldokumenten adressiert spezifische Themen der Datenaufbereitung und Datenbereitstellung. Dazu zählen die Anonymisierung von sensiblen Informationen, die Ableitung von Bildungsklassifikationen, die Erstellung von Gewichtungsvariablen sowie die Nutzung von kleinräumigen Regionalindikatoren.
- Das sog. „Datenstruktur-File“ ist letztlich ein dem Scientific-Use-File entsprechendes Datenpaket – nur ohne Beobachtungen bzw. Fälle. Dieser geleerte Datenbestand ist ohne Vertrag frei zugänglich. Er fungiert einerseits als Anschauungsmaterial für die Datenstruktur eines Scientific-Use-Files. Andererseits kann Syntax auf Grundlage dieses Files vorbereitet werden, was insbesondere bei der Analyse von NEPS-Daten in der geschützten Umgebung des LIfBi-Datensicherheitsraums von Vorteil ist.

Alle genannten Materialien sind ohne Beschränkung über die Webseiten der einzelnen NEPS-Startkohorten abrufbar. Aktuell befindet sich ein übergreifendes Download-Portal in Vorbereitung, welches alle Dokumente versammelt und welches eine gezielte Suche bzw. Auswahl über entsprechende Filteroptionen ermöglicht.

→ www.lifbi.de > Daten & Services > Daten und Dokumentation > NEPS

4.3 Publikationen

Weitere Informationsquellen für das gesamte NEPS-Datenportfolio sind:

- Die „NEPS-Bibliografie“ listet alle bekannten Veröffentlichungen auf, in denen NEPS-Daten verwendet wurden. Die Artikel, Buchbeiträge und Monografien vermitteln nicht nur einen Eindruck von den Analysepotenzialen des NEPS, sondern können auch als Inspiration bzw. Referenz für eigene Auswertungen genutzt werden.
- Die „NEPS Survey Paper“ sind eine LIfBi-eigene Publikationsreihe mit Beiträgen, die sich vor allem methodologischen Aspekten und Fragen des Umgangs mit den NEPS-Daten widmen. Die Reihe richtet sich primär an Personen, die selbst mit den Daten arbeiten und an weiterführenden Informationen zu entsprechenden Themen interessiert sind.
- Die „LIfBi Working Papers“ werden ebenfalls vom LIfBi herausgegeben. Neben Beschreibungen, Analysen und Ergebnisberichten rund um das Nationale Bildungspanel umfassen sie auch Artikel zu anderen Studien bzw. Daten, für die das Institut (mit)verantwortlich ist.
- Die „NEPS-Transferberichte“ adressieren Wissenschaft, Journalismus, Bildungspolitik, Bildungsadministration und Bildungspraxis. Sie fokussieren auf Analysen mit hoher gesellschaftlicher Relevanz (z. B. Corona & Bildung) und stellen diese kompakt und allgemeinverständlich vor.

→ www.lifbi.de > Forschung > Großprojekte > NEPS > NEPS-Bibliografie

4.4 Serviceangebote

Um die Arbeit der Datennutzenden bestmöglich zu unterstützen, wurden diverse Serviceleistungen seitens des FDZ-LIfBi etabliert. Diese Angebote beziehen sich allgemein auf die Forschungsdateninfrastruktur des LIfBi, können aber auch für den Umgang mit NEPScomp wertvolle Informationen und Hilfestellungen liefern.

- Die „Online-Variablensuche“ ermöglicht die Recherche von Variablen über das gesamte Instrumentarium des Nationalen Bildungspanels (und weiterer Studien des LIfBi) hinweg. Das Tool verfügt über eine Volltextsuche nach Schlagworten sowie über verschiedene Filteroptionen und ein hierarchisches Themenverzeichnis. Mit der Variablensuche lässt sich schnell und intuitiv herausfinden, ob geeignete Variablen für die geplante Analyse in den Daten vorhanden sind. Darüber hinaus können ergänzende Informationen zu allen Variablen angezeigt werden – von der Frageformulierung und dem Antwortschema über Variablenname und Vorkommen in Studien bzw. Wellen bis hin zu Quellverweisen. Auch Variablen aus dem NEPScomp-Datenbestand sind in der Variablensuche nachschlagbar, sofern der ursprüngliche Variablenname aus den regulären Scientific-Use-Files übernommen wurde (siehe Kapitel 3.1).
- Die „NEPS-Datenschulungen“ sind mehrfach im Jahr vom FDZ-LIfBi durchgeführte Online-Kurse. Die kostenlosen Veranstaltungen bestehen aus verschiedenen Modulen, die separat besucht werden können. Während die Basismodule vor allem Kenntnisse über die allgemeine Rahmenkonzeption der NEPS-Studie vermitteln und dem Einstieg in die Arbeit mit den NEPS-Daten bzw. der Dokumentation dienen, geht es bei den Erweiterungsmodulen um ausgewählte Schwerpunktthemen wie den Umgang mit Kompetenzdaten, Episodendaten etc. Eines der Basismodule widmet sich explizit der Einführung in das NEPScomp-Datenangebot. Die Anmeldung zur Teilnahme an einer Schulung erfolgt formlos via E-Mail (siehe Kapitel 5.4).
- Mehrere „Video-Tutorials“ begleiten und ergänzen die Datenschulungen visuell. Die Videos thematisieren unter anderem das NEPS-Studiendesign mit den verschiedenen Startkohorten, die Datenstruktur der regulären Scientific-Use-Files und die Herausforderungen beim Zusammenspielen von Daten aus unterschiedlichen Datensätzen (*Merging*). NEPScomp steht als Tutorial sowohl in einer Kurzfassung als auch in einer Langfassung zur Verfügung.
- Das „Forum4MICA“ (*Making Information Commonly Available*) ist eine Online-Plattform für den offenen Austausch zwischen Datennutzenden bzw. informationssuchenden Personen und Mitarbeitenden von Forschungsdatenzentren bzw. Datenanbietern. Das Forum bietet eine niedrigschwellige Möglichkeit, direkt mit Expert:innen und anderen Datennutzenden in Kontakt zu treten. Es können Fragen gestellt, Antworten gegeben, Kommentare hinterlassen oder einfach die Inhalte gezielt nach Auskünften durchsucht werden. Für Themen rund um NEPScomp existiert ein separater Bereich – zu finden in der Kategorie „FDZ-LIfBi“ als Unterkategorie „NEPScomp“. Vorgesehen ist dieser Bereich zum einen für die Kommunikation zwischen Datennutzenden und Mitarbeitenden des FDZ-LIfBi bzw. der NEPS-Studie. Zum anderen bietet er Raum für einen Austausch zwischen den Datennutzenden selbst. Ziel ist es,

Probleme und mögliche Lösungen miteinander zu teilen und öffentlich zur Verfügung zu stellen. Dies kann Fragen zum allgemeinen Verständnis der Daten betreffen, aber auch konkrete Herausforderungen, die sich aus der Arbeit mit NEPScomp ergeben, etwa bei der Programmierung von Syntax. Je mehr Personen sich daran beteiligen und je mehr Beiträge gepostet werden, desto schneller entwickelt sich dieser Forumsbereich zu einem wertvollen Wissensarchiv für alle. Darüber hinaus soll das Forum eine wichtige Rolle bei der Weiterentwicklung von NEPScomp spielen (siehe Kapitel 4.5).

- Der FDZ-Newsletter „LIfBi data“ wird halbjährlich per E-Mail an registrierte Personen versandt und über die Webseite zugänglich gemacht. Er informiert über aktuelle Entwicklungen hinsichtlich der Forschungsdateninfrastruktur des Instituts mit Fokus auf Datenveröffentlichungen, Schulungsveranstaltungen, Konferenzen, Neuerungen im Dokumentations- und Serviceangebot sowie Hinweise auf kürzlich erschienene Publikationen unter Verwendung von Daten des LIfBi. Die Anmeldung für den Newsletter erfolgt unter Angabe einer gültigen E-Mail-Adresse und des Namens.

Die aufgelisteten Serviceangebote finden sich neben weiteren Dienstleistungen zur Auswahl auf dieser Webseite:

→ `www.lifbi.de > Daten und Services > Services`

4.5 Weiterentwicklung

Ein Teil der NEPS-Startkohorten befindet sich noch im regulären Erhebungsturnus von ca. einer Erhebung pro Jahr (siehe Abbildung 1). Andere Startkohorten wurden mittlerweile aus der engmaschigen Begleitung in eine weitmaschige Begleitung mit Kurzerhebungen im Abstand von fünf Jahren überführt (*NEPS next*). Zudem wird das Nationale Bildungspanel perspektivisch um neue Startkohorten aufgefrischt, so wie es zuletzt 2022 mit der neuen Fünftklässler:innen-Stichprobe der Startkohorte 8 geschehen ist. Mit anderen Worten: Der NEPS-Datenbestand wächst weiterhin. Entsprechend soll auch NEPScomp in regelmäßigen Abständen aktualisiert werden – um die neuen Daten zu integrieren und das Datenangebot gemäß seiner Nutzungszwecke weiter zu optimieren. Letzteres ist ein Anspruch, der sich am besten gemeinsam mit den Nutzenden, also primär Studierenden und ihren Seminarleitungen bzw. Betreuungspersonen sowie Einsteiger:innen, umsetzen lässt.

Rückmeldungen via Forum4MICA

Dem zuvor in Kapitel 4.4 als Serviceangebot erwähnten Forum4MICA kommt bei der kontinuierlichen Weiterentwicklung von NEPScomp eine besondere Bedeutung zu. Der speziell für NEPScomp eingerichtete Bereich kann und soll nicht zuletzt auch für Anregungen und Vorschläge zur Verbesserung des Datenangebots genutzt werden. Wird beispielsweise im Zuge der Datenaufbereitung für eigene Analysen eine neue (Hilfs)Variable generiert, die auch für andere Nutzende von Interesse sein könnte, dann kann diese Variable in eine zukünftige Version von NEPScomp

aufgenommen werden. Voraussetzung hierfür ist, dass zuvor ein Austausch über die neue Variable – inkl. Syntax zur Erzeugung – im Forum stattgefunden hat. Daneben ist aber auch allgemeines Feedback jederzeit herzlich willkommen.

Die Beiträge im Online-Forum sind grundsätzlich offen und können von allen interessierten Personen durchsucht und gelesen werden. Eine aktive Beteiligung mit eigenen Postings – Fragen, Antworten, Kommentaren etc. – erfordert nicht mehr als eine einmalige Registrierung unter Angabe einer gültigen E-Mail-Adresse und eines Nutzernamens. Das Erstellen von Beiträgen selbst ist unkompliziert und weitestgehend intuitiv. Man findet im Forum aber auch Hilfetexte sowie eine allgemeine Anleitung mit Verweis auf die vielfältigen Funktionen und Möglichkeiten.

→ `www.lifbi.de > Daten & Services > Services > Forum4MICA`

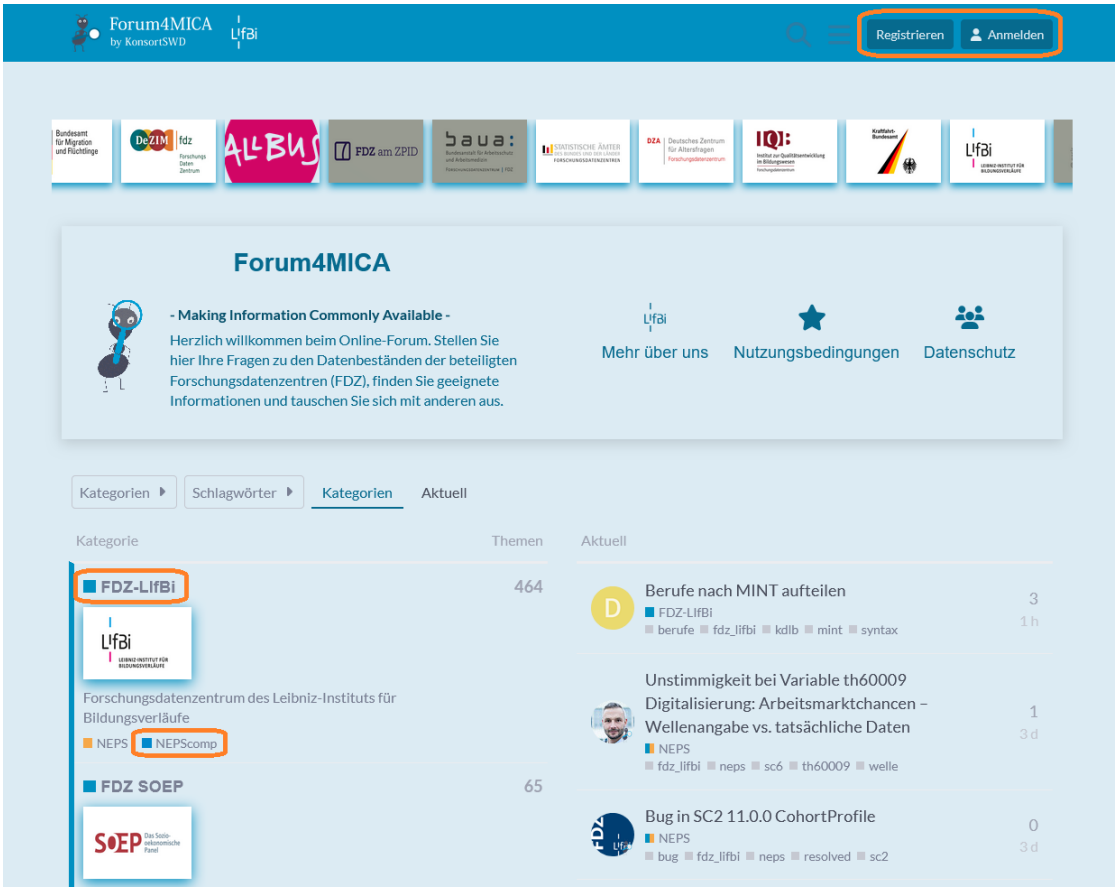

Abbildung 5: Startseite Forum4MICA mit Markierungen (Aufruf am 25.08.2025)

5 Datennutzung

5.1 Zugang zu den Daten

NEPScomp ist ein Bestandteil der Forschungsdateninfrastruktur des Nationalen Bildungspanels. Diese Infrastruktur umfasst neben NEPScomp alle regulären Scientific-Use-Files der aktuell sieben NEPS-Startkohorten sowie der beiden abgeschlossenen Zusatzstudien Thüringen und Baden-Württemberg. Alle Daten und Dokumentationsmaterialien sind kostenfrei nutzbar. Die Bereitstellung erfolgt durch das Forschungsdatenzentrum des Leibniz-Instituts für Bildungsverläufe (FDZ-LIfBi). Der Zugang zu den Daten ist an den Abschluss eines **NEPS-Datennutzungsvertrags** gebunden. Um NEPScomp verwenden zu können, sind folgende Schritte und Regelungen zu beachten:

Antrag ausfüllen Zunächst ist das Antragsformular „NEPS-Datennutzungsvertrag“ (in deutscher und englischer Sprache verfügbar) von der Webseite herunterzuladen und auszufüllen. Die Angaben zum Analysevorhaben umfassen im Wesentlichen einen Titel, eine Kurzbeschreibung und ein Enddatum für die Datennutzung. Ein Antrag kann auch von mehreren Personen gemeinsam gestellt werden.

- Antragsberechtigt sind ausschließlich Personen mit einer Anbindung an eine wissenschaftliche Einrichtung (z. B. Universität, Fachhochschule, Leibniz-Institut).
- Alle an einem Antrag beteiligten Personen müssen der gleichen wissenschaftlichen Einrichtung angehören bzw. über eine entsprechende Anbindung verfügen.
- Studierende können nicht selbst einen Antrag stellen, sondern erhalten Zugang zu den NEPS-Daten als „zusätzlich Beteiligte“ an einem Datennutzungsvertrag. Das heißt, der Antrag muss formal von einer angestellten Person der wissenschaftlichen Einrichtung gestellt werden – etwa von der Seminarleitung oder der betreuenden Person einer Qualifizierungsarbeit. Die persönlichen Angaben der antragstellenden Person sind auf der ersten Seite des Formulars einzutragen. Die zusätzlich an der Datennutzung beteiligten Studierenden sind auf der dritten Seite anzugeben. Bei Bedarf kann diese Seite auch mehrfach einem Antrag beigefügt werden.
- Alle in einem Antrag namentlich erwähnten Personen müssen die Vertragsbedingungen zur Kenntnis nehmen und das Dokument auf der letzten Seite eigenhändig unterschreiben.

  → `www.lifbi.de > Daten & Services > Datenzugang > NEPS -> Datenzugang`

Antrag einreichen Der vollständig ausgefüllte und unterschriebene Antrag wird als Scan per E-Mail oder per Post an das FDZ-LIfBi gesendet. Die erforderlichen Kontaktinformationen sind in der Kopfzeile auf der ersten Seite des Antragsformulars angegeben.

- Ein elektronischer Versand beschleunigt in aller Regel die Bearbeitung.

Antrag bearbeiten Das FDZ-LIfBi überprüft die Angaben im Antragsformular und richtet für alle im Formular genannten Personen ein Nutzerkonto ein – sofern ein solches nicht bereits vorhanden ist. Darüber hinaus werden die Angaben zum Forschungsvorhaben in eine Datenbank eingepflegt, um die Informationen auf der entsprechenden Webseite darstellen zu können.

→ www.lifbi.de > Daten & Services > Datenzugang > NEPS-Forschungsprojekte

Zugangsdaten übermitteln Jede am Datennutzungsvertrag beteiligte Person erhält in zwei separaten E-Mails eine individuelle Nutzerkennung (nu-Nummer) und ein Erstpasswort. Die antragstellende Person erhält außerdem das gegengezeichnete Vertragsdokument zurückgesandt.

- Die E-Mail mit dem Erstpasswort enthält auch Instruktionen zur Änderung dieses Erstpasworts in ein persönliches Passwort. Nur mit einem geeigneten persönlichen Passwort – in Kombination mit der jeweiligen Nutzerkennung – ist der Download von NEPS-Daten nach Anmeldung an der Webseite möglich.
- Zu beachten ist, dass die Synchronisierung des neuen, persönlichen Passwortes etwas Zeit in Anspruch nimmt. Entsprechend darf das Einloggen auf der Webseite nicht unmittelbar nach der Änderung des Erstpasworts erfolgen. Nach ein paar Minuten sollte die Anmeldung funktionieren.
- Die Kombination aus individueller Nutzerkennung und persönlichem Passwort bleibt konstant. Wenn ein Passwort vergessen wird, kann jederzeit über die Webseite ein neues Erstpasswort angefordert werden. Ist die Nutzerkennung nicht mehr bekannt, dann bitte Kontakt zum FDZ-LIfBi (siehe Kapitel 5.4) aufnehmen.

Daten herunterladen Jede Person mit gültigen Zugangsdaten kann sich auf der Webseite anmelden und die gewünschten NEPS-Datenpakete beliebig oft herunterladen. Im Unterschied zu den regulären Scientific-Use-Files steht NEPScomp ausschließlich als Download-Variante zur Verfügung.[9]

- Nach dem Download gilt die vertragliche Festlegung, dass eine Weitergabe der NEPS-Daten an unberechtigte Personen (sog. „Dritte“, d. h. Personen ohne gültigen Nutzungsvertrag) nicht gestattet ist.

  → www.lifbi.de > Daten & Services > Daten und Dokumentation > NEPS

Vertrag nachträglich anpassen Sofern sich zu einem späteren Zeitpunkt der Bedarf ergibt, die Nutzungsdauer der NEPS-Daten im Rahmen des vereinbarten Forschungsvorhabens zu verlängern oder weitere Personen in die Datennutzung aufzunehmen, so kann dies über das ebenfalls auf der Webseite angebotene „Änderungsformular“ beantragt werden.

9 Die Zusatzformulare für eine RemoteNEPS- bzw. eine On-site-Ergänzung sind im Hinblick auf NEPScomp ohne Bedeutung. Sie werden erst dann relevant, wenn Zugang zu bestimmten Informationen in den regulären Scientific-Use-Files benötigt wird, die aus Gründen des Datenschutzes in den Download-Varianten nicht enthalten sind.

- Die antragstellende Person des Basisvertrags muss auch im Änderungsformular auf der ersten Seite eingetragen sein und das Dokument unterschreiben. Neu hinzukommende Personen sind auf der zweiten Seite einzutragen. Auch von diesen Personen ist eine Unterschrift erforderlich.

  → www.lifbi.de > Daten & Services > Datenzugang > NEPS –> Datenzugang

5.2 Zitation der Daten

Jede Datenveröffentlichung des NEPS wird bei DataCite registriert und mit einem *Digital Object Identifier* (DOI) versehen. Dieser Identifier hat drei Funktionen: Erstens kennzeichnet er das jeweilige Datenprodukt (siehe Tabelle 2). Hierbei steht 10.5157 für die Registrierungsnummer des LIfBi, NEPScomp für die Bezeichnung des Datenbestands, SUF für Scientific-Use-File und 2.0.0 für die Datenversion.[10] Zweitens ist jeder DOI mit einer „Landing Page“ verknüpft, die weiterführende Informationen zum Datenprodukt enthält. Drittens ermöglicht dieser Identifikator eine einfache und präzise Zitation der verwendeten Daten. Eine korrekte Datenzitation gehört ebenso wie die Angabe von Literaturquellen zur **guten wissenschaftlichen Praxis**, denn nur auf dieser Grundlage sind Replikationsanalysen zur Überprüfung der berichteten Befunde überhaupt möglich.

Tabelle 2: Versionshistorie NEPScomp

Version	Digital Object Identifier	Veröffentlicht am
2.0.0 (current)	doi:10.5157/NEPScomp:SUF:2.0.0	**23.10.2026**
1.0.0	doi:10.5157/NEPScomp:SUF:1.0.0	01.09.2025

Personen, die Analysen auf Basis von NEPScomp oder anderen NEPS-Daten durchführen, sind verpflichtet, in ihren Arbeiten die Verwendung der Daten inklusive DOI auszuweisen. NEPScomp ist durch folgende Zitation im Literaturverzeichnis anzuzeigen:

> NEPS-Netzwerk. (2026). *NEPScomp, National Educational Panel Study*. Leibniz-Institut für Bildungsverläufe (LIfBi), Bamberg. https://doi.org/10.5157/NEPScomp:SUF:2.0.0

10 Die erste Ziffer der Versionskennung bezieht sich auf die Anzahl der sog. Hauptreleases. Diese Ziffer wird hochgezählt, sobald eine neue Version von NEPScomp mit neuen Daten veröffentlicht wird, z. B. durch Ergänzung der bisherigen Daten um neue Erhebungswellen und/oder eine neue Startkohorte. Die zweite Ziffer gibt Auskunft über „Major Updates“, bei denen es nicht primär um die Erweiterung des Datenbestands, sondern um die Korrektur von Fehlern in den Daten geht. Die dritte Ziffer informiert über „Minor Updates“, bei denen zum Beispiel Variablenlabel oder Übersetzungen geändert wurden – diese kleineren Korrekturen erfordern keine Anpassung von Syntax. Diese dreistellige Versionskennung findet sich auch in der Bezeichnung des Datenbestands selbst sowie der darin enthaltenen Datensätze. Jede neue Version von NEPScomp – egal ob Hauptrelease oder Update – ersetzt die vorherige Version vollständig.

Zusätzlich sollte an geeigneter Stelle auf die NEPS-Studie hingewiesen werden:

> Diese Arbeit nutzt Daten des Nationalen Bildungspanels (NEPS; vgl. Blossfeld & Roßbach, 2019). Das NEPS wird vom Leibniz-Institut für Bildungsverläufe (LIfBi, Bamberg) in Kooperation mit einem deutschlandweiten Netzwerk durchgeführt.

Die bibliografischen Angaben der dazugehörigen Quelle lauten:

> Blossfeld, H.-P., & Roßbach, H.-G. (Hrsg.). (2019). *Education as a lifelong process: The German National Educational Panel Study (NEPS). Edition ZfE* (2nd ed.). Springer VS. https://doi.org/10.1007/978-3-658-23162-0

Für die Zitation des vorliegenden Datenmanuals empfiehlt sich folgende Vorlage:

> FDZ-LIfBi. (2026). *Datenmanual NEPScomp, Version 2.0.0*. Bamberg, Leibniz-Institut für Bildungsverläufe, National Educational Panel Study.

Weitere Zitationsvorgaben, zum Beispiel zu den Erhebungsinstrumenten oder den Feld- und Methodenberichten, können im Data Manual der jeweiligen NEPS-Startkohorte nachgeschlagen werden (siehe Kapitel 4.2).

5.3 Empfehlungen

Vor Beginn der Analyse informieren Insbesondere dieses Datenmanual liefert eine Reihe von Hinweisen bezüglich der Entscheidung, ob NEPScomp eine geeignete Datenbasis für das eigene Analysevorhaben darstellt oder nicht. Darüber hinaus stehen weitere Materialien für das notwendige Hintergrundwissen zur Verfügung (siehe Kapitel 4.2).

Regelungen des Datenzugangs beachten Voraussetzung für die Nutzung von NEPScomp ist die Beteiligung an einem entsprechenden Vertrag. Die darin festgelegten Regeln sind zu beachten. Dies gilt insbesondere für das Verbot einer Weitergabe der Daten an Dritte sowie jeglichen Versuch einer Re-Identifikation von Personen, Haushalten oder Einrichtungen. Alle NEPS-Daten sind vertraulich zu behandeln und nur für den vereinbarten Zweck zu nutzen (siehe Kapitel 5.1).

Verständnis für die Daten schaffen Das Nationale Bildungspanel ist eine überaus umfangreiche Panelstudie mit komplexem Erhebungsdesign. NEPScomp nimmt zwar zahlreiche Vereinfachungen vor; dennoch sollte bei der Arbeit mit den Daten auch die eigentliche Dokumentation der jeweils interessierenden NEPS-Startkohorte berücksichtigt werden. Vor allem die Erhebungsinstrumente und die Methodenberichte vermitteln eine detaillierte Vorstellung davon, welche Informationen wie gesammelt wurden (siehe Kapitel 4.2).

Prüfung der Daten sicherstellen Für jede Analyse ist es von zentraler Bedeutung, dass die richtigen Variablen ausgewählt und angemessen eingesetzt werden. NEPScomp erfordert hier besondere Aufmerksamkeit, weil teilweise Informationen von verschiedenen Auskunftspersonen (Zielpersonen, Eltern, pädagogisches Fachpersonal) in gemeinsamen Datensätzen enthalten sind (unterscheidbar über den ersten Buchstaben im Variablennamen: `t`,

p, e) und weil das „vollständige“ LONG-Datenformat in aller Regel mit fehlenden Werten in den Häufigkeitsverteilungen der einzelnen Variablen einhergeht (siehe Kapitel 3.2).

Datenstruktur korrekt nutzen NEPScomp kombiniert längsschnittliche und querschnittliche Informationen in der einheitlichen LONG-Struktur. Über den Wellenidentifikator `wave` lassen sich die einzelnen Erhebungszeitpunkte identifizieren und bei Bedarf für Analysen auswählen (Stata: `keep if wave==3`). Der Personenidentifikator `ID_t` ist über alle Panelwellen hinweg persistent, d. h. eine Zielperson ist während des gesamten Panelverlaufs immer mit der gleichen `ID_t`-Nummer verknüpft. Eine Umwandlung des bereitgestellten LONG-Formats in das WIDE-Format ist auf Basis beider Identifikatoren über entsprechende Befehle – Stata: `reshape wide`, SPSS: `casestovars` – möglich.

Variablen aus verschiedenen Datensätzen miteinander verknüpfen Oftmals sind die für eine Analyse relevanten Informationen in unterschiedlichen Datensätzen abgelegt. NEPScomp umfasst pro Startkohorte acht Datensätze. Das Zusammenspielen ist hier aufgrund der vereinheitlichten Datenstruktur allerdings keine große Herausforderung. Über die Kombination aus Personen- und Wellenzugehörigkeit (`ID_t`, `wave`) lassen sich Variablen aus allen Datensätzen 1:1 miteinander verbinden. Bei Querschnittanalysen können entweder zunächst die fraglichen Variablen im Längsschnitt kombiniert und anschließend die Wellenauswahl über `wave` getroffen werden. Oder man beschränkt erst die Variablen in den jeweiligen Datensätzen auf die fragliche Welle und verknüpft diese Variablen dann über den Personenidentifikator `ID_t` (siehe Abbildung 3).

Beschränkungen bei der Analyse von Kompetenzdaten berücksichtigen Manche Kompetenzdomänen sind nur einmalig pro Startkohorte getestet. Die entsprechenden Variablen weisen dann nur für die jeweilige Welle gültige Werten auf und lassen sich auch nur im Querschnitt analysieren. Es gibt aber auch Variablen zu wiederholt eingesetzten Kompetenztests, die ebenfalls nur querschnittlich auswertbar sind. Dabei handelt es sich primär um korrigierte WLE-Schätzer (Suffix im Variablennamen: `_sc1`, vgl. Pohl und Carstensen, 2012). Die betreffenden Variablen sind allesamt durch einen deutlichen Hinweis im Label – QUERSCHNITT – gekennzeichnet. Diese Variablen sollten nicht für längsschnittliche Analysen verwendet werden, auch wenn für sie eine längsschnittliche Datenstruktur vorliegt (siehe auch die Übersichten in Kapitel A.3).

Auf reguläre Scientific-Use-Files wechseln NEPScomp ist aufgrund der reduzierten Inhalte nicht für alle Arten von Auswertungsverfahren geeignet, die grundsätzlich mit den NEPS-Daten möglich wären (z. B. Ereignisdatenanalysen, Mehrebenenanalysen, Regionalanalysen, Methodenanalysen, Gewichtung). Sobald tiefergehende Informationen für die Beantwortung einer Fragestellung mit den NEPS-Daten benötigt werden, empfiehlt sich der Umstieg auf die regulären Scientific-Use-Files der einzelnen NEPS-Startkohorten. NEPScomp dient primär dem Einstieg in die Arbeit mit den Daten des Nationalen Bildungspanels – keinesfalls als Ersatz für die umfassendere Dateninfrastruktur.

Daten angemessen zitieren In allen Arbeiten sollte die Verwendung der NEPS-Daten gemäß den Regeln guter wissenschaftlicher Praxis angezeigt werden. Auch für NEPScomp existieren eine DOI und eine entsprechende Vorgabe für die Datenzitation (siehe Kapitel 5.2).

Publikation melden Gemäß Datennutzungsvertrag ist das FDZ-LIfBi über jede Art von Publikation zu informieren, die aus der Arbeit mit den NEPS-Daten hervorgeht. Dazu zählen auch studentische Abschlussarbeiten etc. auf der Basis von NEPScomp. Für die Benachrichtigung genügt eine E-Mail mit Angabe der bibliografischen Informationen oder einem Link zur Veröffentlichung (siehe Kapitel 5.4).

Sich über das Forum4MICA austauschen Die öffentliche Online-Plattform verfügt über einen eigenen Bereich „NEPScomp“, der sowohl für die Klärung von Fragen und Problemen als auch für Feedback und Anregungen zur weiteren Entwicklung des Datenprodukts vorgesehen ist. Das FDZ-LIfBi hofft auf eine rege Beteiligung (siehe Kapitel 4.5).

5.4 Kontakt zum Forschungsdatenzentrum des LIfBi

Das FDZ-LIfBi verantwortet die Aufbereitung, Dokumentation und Bereitstellung von NEPScomp. Es fungiert zugleich als erste Anlaufstelle für Fragen und Anregungen rund um dieses Datenprodukt. Allgemeine Anliegen sollten dabei nach Möglichkeit zunächst über das Forum4MICA adressiert werden (siehe Kapitel 4.5). Selbstverständlich steht das FDZ-LIfBi auch für individuelle Auskünfte via E-Mail oder Telefon zur Verfügung.

E-Mail: `fdz@lifbi.de`
Web: → `www.lifbi.de` > `Institution` > `Abteilung 3` > `Forschungsdatenzentrum`
Telefon: +49 951 700 60-037

6 Literatur

Blossfeld, H.-P., & Roßbach, H.-G. (Hrsg.). (2019). *Education as a lifelong process: The German National Educational Panel Study (NEPS). Edition ZfE* (2nd ed.). Springer VS. https://doi.org/10.1007/978-3-658-23162-0

FDZ-LIfBi. (2026). *Datenmanual NEPScomp, Version 2.0.0*. Bamberg, Leibniz-Institut für Bildungsverläufe, National Educational Panel Study.

Fuß, D., Gnambs, T., Lockl, K., Attig, M., & Nusser, L. (2025). *Competence Data in NEPS: Overview of Measures and Variable Naming Conventions (Starting Cohorts 1 to 8), Revised Version 2025* (LIfBi Research Data Documentation Series). Bamberg, Germany: Leibniz Institute for Educational Trajectories.

NEPS-Netzwerk. (2026). *NEPScomp, National Educational Panel Study*. Leibniz-Institut für Bildungsverläufe (LIfBi), Bamberg. https://doi.org/10.5157/NEPScomp:SUF:2.0.0.

Pohl, S., & Carstensen, C. H. (2012). *NEPS Technical Report – Scaling the Data of the Competence Tests* (NEPS Working Paper Nr. 14). German National Educational Panel Study (NEPS). Bamberg.

A Appendix

A.1 Beispiel für Merging-Syntax in SPSS

```
** Using-Datensatz öffnen => SC6, Demography
** Variablen auswählen => ID_t wave Geschlecht
** nach Schlüsselvariablen sortieren => ID_t wave
GET FILE='${datapath}\NEPScomp_D_SPSS_2-0-0\SC6\SC6_NEPScomp_Demography_2-0-0.sav'
        /KEEP=ID_t wave t700001.
SORT CASES BY ID_t wave.
DATASET NAME using.

** Master-Datensatz öffnen => SC6, Society
** nach Schlüsselvariablen sortieren => ID_t wave
GET FILE='${datapath}\NEPScomp_D_SPSS_2-0-0\SC6\SC6_NEPScomp_Society_2-0-0.sav'
SORT CASES BY ID_t wave.
DATASET NAME master.

** Zusammenspielen der Datensätze => using master
** über Schlüsselvariablen => ID_t, wave
MATCH FILES
        /FILE='using'
        /FILE=*
        /BY ID_t wave.

** Messzeitpunkte von Religiosität => Wellen 3, 7, 11, 15
CROSSTABS
        /TABLES=wave BY t435000
        /FORMAT=AVALUE TABLES
        /CELLS=COUNT
        /COUNT ROUND CELL.

** Welle für Querschnittanalyse auswählen => Welle 3 (=2. NEPS-Haupterhebung)
SELECT IF (wave = 3).
EXECUTE.

** Geschlechtsunterschiede prüfen => T-Test
T-TEST GROUPS=t700001(1 2)
        /MISSING=ANALYSIS
        /VARIABLES=t435000
        /ES DISPLAY(TRUE)
        /CRITERIA=CI(.95).
```

A.2 Beispiel für Merging-Syntax in R

```
# Erforderliches Paket laden (alternativ "readstata13")
require("haven")

# Datenpfad für NEPScomp auf lokalem Rechner definieren => z.B. Z:/Data
setwd("Z:/Data/NEPScomp_D_Stata_2-0-0/SC6")

# Datensätze importieren => Society, Demography
nepscomp_society <- read_dta("SC6_NEPScomp_Society_2-0-0.dta")
nepscomp_demography <- read_dta("SC6_NEPScomp_Demography_2-0-0.dta")

# Anspielen der Zielvariablen => Geschlecht
# über Schlüsselvariablen => ID_t, wave
nepscomp_merged <- merge(nepscomp_society, nepscomp_demography, by=c("ID_t",
  "wave"))

# Messzeitpunkte von Religiosität => Wellen 3, 7, 11, 15
table(as_factor(nepscomp_merged$wave), as_factor(nepscomp_merged$t435000))

# Geschlechtsunterschiede prüfen => T-Test
# und Welle auswählen => Welle 3 (=2. NEPS-Haupterhebung)
t.test(t435000 ~ t700001, nepscomp_merged, wave == 3)
```

A.3 Kompetenzdaten

In der folgenden Tabelle sind – nach Startkohorten getrennt – alle in NEPScomp enthaltenen Kompetenzmaße aus den Testungen der Zielpersonen aufgelistet. Neben der Bezeichnung der Kompetenzdomäne und dem jeweiligen Kürzel im Variablennamen sind dort auch die Erhebungszeitpunkte (Wellen), die Art der Indexbildung eines Gesamt-Testscores aus den Einzelscores, das dazugehörige Suffix im Variablennamen und die Art der zulässigen Datenanalyse angegeben. Zu beachten ist, dass nicht alle Kompetenzmaße, die in einer Startkohorte wiederholt erhoben wurden, für Längsschnittanalysen geeignet sind (siehe Kapitel 3.4). Für weiterführende Informationen zu den Kompetenzmessungen sei auf die startkohortenspezifischen Dokumentationsmaterialien (siehe Kapitel 4.2) und dort speziell auf die Übersichten in Fuß et al., 2025 verwiesen.

Tabelle 3: Übersicht über die verfügbaren Kompetenzmaße

SC	Domäne	Kürzel	Erhebungswellen	Scoring-Typ	Suffix	Analyse
1	Sensomotorische Entwicklung	`cd`	1	WLE (korrigiert)	`sc1`	Querschnitt
1	Eltern-Kind-Interaktion (mehrdimens.)	`ih`	1, 2, 3			Längsschnitt
1	SON-R Subtest: Kategorisierung	`ca`	4	WLE (korrigiert)	`sc1`	Querschnitt
1	Belohnungsaufschub	`de`	4, 6, 8			Längsschnitt
1	Zahlenspanne vorwärts	`ds`	4, 7	Summe max. Ziffernpanne	`sc3a` `sc3b`	Längsschnitt Längsschnitt
1	Wortschatz	`vo`	4, 6, 8, 10	Summe	`sc3`	Längsschnitt
1	Prozedurale Metakognition Wortschatz	`mpvo`	8, 10	Differenzmaß Anteil korrekt	`sc5` `sc6`	Längsschnitt Längsschnitt
1	Mathematische Kompetenz	`ma`	5, 7, 9, 11	WLE (korrigiert) WLE (unkorrigiert)	`sc1` `sc1u`	Querschnitt Längsschnitt

(...)

Tabelle 3: (continued)

SC	Domäne	Kürzel	Erhebungswellen	Scoring-Typ	Suffix	Analyse
1	Prozedurale Metakognition Mathematik	mpma	9, 11	Differenzmaß Anteil korrekt	sc5 sc6	Längsschnitt Längsschnitt
1	Naturwissenschaftliche Kompetenz	sc	6, 8, 10	WLE (korrigiert) WLE (unkorrigiert)	sc1 sc1u	Querschnitt Längsschnitt
1	Prozedurale Metakognition Naturwissenschaften	mpsc	8, 10	Differenzmaß Anteil korrekt	sc5 sc6	Längsschnitt Längsschnitt
1	DGCF Schlussfolgern	dg	7, 10	WLE (korrigiert) WLE (unkorrigiert)	sc1 sc1u	Querschnitt Längsschnitt
1	DGCF Wahrnehmungsgeschwindigkeit	dg	7, 11	Summe	sc3a	Längsschnitt
1	Zahlenspanne rückwärts	bd	8	Summe max. Ziffernpanne	sc3a sc3b	Querschnitt Querschnitt
1	Lesegeschwindigkeit	rs	9, 11	Summe	sc3	Längsschnitt
1	Frühe Lesekompetenz	rx	9, 11	Summe	sc3	Längsschnitt
1	Prozedurale Metakognition Frühes Lesen	mprx	9, 11	Differenzmaß Anteil korrekt	sc5 sc6	Längsschnitt Längsschnitt
2	Grammatik	gr	1, 3	WLE (korrigiert) WLE (unkorrigiert)	sc1 sc1u	Querschnitt Längsschnitt
2	Prozedurale Metakognition Grammatik	mpgr	1, 3	Differenzmaß Anteil korrekt	sc5 sc6	Längsschnitt Längsschnitt
2	Wortschatz	vo	1, 3, 5	WLE (korrigiert) WLE (unkorrigiert)	sc1 sc1u	Querschnitt Längsschnitt

(...)

Tabelle 3: (continued)

SC	Domäne	Kürzel	Erhebungswellen	Scoring-Typ	Suffix	Analyse
2	Prozedurale Metakognition Wortschatz	mpvo	1, 3, 5	Differenzmaß Anteil korrekt	sc5 sc6	Längsschnitt Längsschnitt
2	Naturwissenschaftliche Kompetenz	sc	1, 3, 5, 9 1, 3, 5	WLE (korrigiert) WLE (unkorrigiert)	sc1 sc1u	Querschnitt Längsschnitt
2	Prozedurale Metakognition Naturwissenschaften	mpsc	1, 3, 5, 9	Differenzmaß Anteil korrekt	sc5 sc6	Längsschnitt Längsschnitt
2	Mathematische Kompetenz	ma	2, 3, 4, 6, 9, 11 2, 3, 4, 6, 9	WLE (korrigiert) WLE (unkorrigiert)	sc1 sc1u	Querschnitt Längsschnitt
2	Prozedurale Metakognition Mathematik	mpma	2, 3, 4, 6, 9, 11	Differenzmaß Anteil korrekt	sc5 sc6	Längsschnitt Längsschnitt
2	DGCF Schlussfolgern	dg	2, 4, 11	Summe	sc3b	Längsschnitt
2	DGCF Wahrnehmungsgeschwindigkeit	dg	2, 4	Summe	sc3a	Längsschnitt
2	Belohnungsaufschub	de	2, 6			Längsschnitt
2	Zahlenspanne vorwärts	ds	2	Summe	sc3a	Querschnitt
2	Zahlenspanne rückwärts	bd	2	Summe	sc3a	Querschnitt
2	Identifikation von Phonemen	ip	2	Summe	sc3	Querschnitt
2	Reime	ri	2	Summe	sc3	Querschnitt
2	Onset-Reim-Synthetisieren	on	2	Summe	sc3	Querschnitt
2	Frühe Buchstabenkenntnis	lk	2	Summe	sc3	Querschnitt

(...)

Tabelle 3: (continued)

SC	Domäne	Kürzel	Erhebungswellen	Scoring-Typ	Suffix	Analyse
2	Deklarative Metakognition	md	3, 5	Mittelwert	sc4	Querschnitt
2	Lesegeschwindigkeit	rs	4	Summe	sc3	Querschnitt
2	Frühe Lesekompetenz	rx	4	Summe	sc3	Querschnitt
2	Prozedurale Metakognition Frühes Lesen	mprx	4	Differenzmaß Anteil korrekt	sc5 sc6	Querschnitt Querschnitt
2	Hörverstehen Russisch	nr	4	WLE (korrigiert)	sc1	Querschnitt
2	Prozedurale Metakognition Russisch	mpnr	4	Differenzmaß Anteil korrekt	sc5 sc6	Querschnitt Querschnitt
2	Hörverstehen Türkisch	nt	4	WLE (korrigiert)	sc1	Querschnitt
2	Prozedurale Metakognition Türkisch	mpnt	4	Differenzmaß Anteil korrekt	sc5 sc6	Querschnitt Querschnitt
2	ICT Kompetenz	ic	5	WLE (korrigiert)	sc1	Querschnitt
2	Prozedurale Metakognition ICT	mpic	5	Differenzmaß Anteil korrekt	sc5 sc6	Querschnitt Querschnitt
2	Lesekompetenz	re	6, 9	WLE (korrigiert) WLE (unkorrigiert)	sc1 sc1u	Querschnitt Längsschnitt
2	Prozedurale Metakognition Lesen (gesamt)	mpre	6, 9	Differenzmaß Anteil korrekt	sc5 sc6	Längsschnitt Längsschnitt
2	Orthografie (2 Sets)	or	6	WLE (korrigiert) Set1 WLE (korrigiert) Set2	sc1a sc1b	Querschnitt Querschnitt

(...)

Tabelle 3: (continued)

SC	Domäne	Kürzel	Erhebungswellen	Scoring-Typ	Suffix	Analyse
2	Prozedurale Metakognition Orthografie (2 Sets)	mpor	6	Differenzmaß Anteil korrekt	sc5 sc6	Querschnitt Querschnitt
3	Mathematische Kompetenz	ma	1, 3, 5, 9	WLE (korrigiert) WLE (unkorrigiert)	sc1 sc1u	Querschnitt Längsschnitt
3	Prozedurale Metakognition Mathematik	mpma	1, 3, 5, 9	Differenzmaß Anteil korrekt	sc5 sc6	Längsschnitt Längsschnitt
3	Lesekompetenz	re	1, 3, 6, 9	WLE (korrigiert) WLE (unkorrigiert)	sc1 sc1u	Querschnitt Längsschnitt
3	Prozedurale Metakognition Lesen (gesamt)	mpre	1, 3, 6, 9	Differenzmaß Anteil korrekt	sc5 sc6	Längsschnitt Längsschnitt
3	Orthografie (2 Sets)	or	1, 3, 5	WLE (korrigiert) Set1 WLE (korrigiert) Set2	sc1a sc1b	Querschnitt Querschnitt
3	Prozedurale Metakognition Orthografie (2 Sets)	mpor	1, 3 1, 3, 5	Differenzmaß Anteil korrekt	sc5 sc6	Querschnitt Querschnitt
3	Lesegeschwindigkeit	rs	1, 5	Summe	sc3	Längsschnitt
3	DGCF Wahrnehmungsgeschwindigkeit	dg	1, 6	Summe	sc3a	Längsschnitt
3	DGCF Schlussfolgern	dg	1, 6	Summe	sc3b	Längsschnitt
3	Naturwissenschaftliche Kompetenz	sc	2, 5, 8	WLE (korrigiert) WLE (unkorrigiert)	sc1 sc1u	Querschnitt Längsschnitt
3	Prozedurale Metakognition Naturwissenschaften	mpsc	2, 5, 8	Differenzmaß	sc5	Längsschnitt

(...)

Tabelle 3: (continued)

SC	Domäne	Kürzel	Erhebungswellen	Scoring-Typ	Suffix	Analyse
				Anteil korrekt	sc6	Längsschnitt
3	ICT Kompetenz	ic	2, 5, 9	WLE (korrigiert) WLE (unkorrigiert)	sc1 sc1u	Querschnitt Längsschnitt
3	Prozedurale Metakognition ICT	mpic	2, 5, 9	Differenzmaß Anteil korrekt	sc5 sc6	Längsschnitt Längsschnitt
3	Deklarative Metakognition	md	2, 6	Mittelwert	sc4	Querschnitt
3	Wortschatz	vo	2	Summe	sc3	Querschnitt
3	Prozedurale Metakognition Wortschatz	mpvo	2	Differenzmaß Anteil korrekt	sc5 sc6	Querschnitt Querschnitt
3	Hörverstehen Russisch	nr	3, 6	WLE (korrigiert)	sc1	Querschnitt
3	Prozedurale Metakognition Russisch	mpnr	3, 6	Differenzmaß Anteil korrekt	sc5 sc6	Längsschnitt Längsschnitt
3	Hörverstehen Türkisch	nt	3, 6	WLE (korrigiert)	sc1	Querschnitt
3	Prozedurale Metakognition Türkisch	mpnt	3, 6	Differenzmaß Anteil korrekt	sc5 sc6	Längsschnitt Längsschnitt
3	Hörverstehen	li	6	WLE (unkorrigiert)	sc1u	Querschnitt
3	Prozedurale Metakognition Hörverstehen	mpli	6	Differenzmaß Anteil korrekt	sc5 sc6	Querschnitt Querschnitt
3	Englisch Lesekompetenz	ef	7, 9	WLE (korrigiert) WLE (unkorrigiert)	sc1 sc1u	Querschnitt Längsschnitt

(...)

Tabelle 3: (continued)

SC	Domäne	Kürzel	Erhebungswellen	Scoring-Typ	Suffix	Analyse
3	Prozedurale Metakognition Englisch	mpef	7, 9	Differenzmaß Anteil korrekt	sc5 sc6	Längsschnitt Längsschnitt
3	Wissenschaftspropädeutik	st	9	WLE (korrigiert)	sc1	Querschnitt
3	Prozedurale Metakognition Wissenschaftspropädeutik	mpst	9	Differenzmaß Anteil korrekt	sc5 sc6	Querschnitt Querschnitt
4	Naturwissenschaftliche Kompetenz	sc	1, 5, 14	WLE (korrigiert) WLE (unkorrigiert)	sc1 sc1u	Querschnitt Längsschnitt
4	Prozedurale Metakognition Naturwissenschaften	mpsc	1, 5, 14	Differenzmaß Anteil korrekt	sc5 sc6	Längsschnitt Längsschnitt
4	Mathematische Kompetenz	ma	1, 7, 10	WLE (korrigiert) WLE (unkorrigiert)	sc1 sc1u	Querschnitt Längsschnitt
4	Prozedurale Metakognition Mathematik	mpma	1, 7, 10	Differenzmaß Anteil korrekt	sc5 sc6	Längsschnitt Längsschnitt
4	ICT Kompetenz	ic	1, 7, 14	WLE (korrigiert) WLE (unkorrigiert)	sc1 sc1u	Querschnitt Längsschnitt
4	Prozedurale Metakognition ICT	mpic	1, 7	Differenzmaß Anteil korrekt	sc5 sc6	Längsschnitt Längsschnitt
4	Lesegeschwindigkeit	rs	1	Summe	sc3	Querschnitt
4	Wortschatz	vo	1	Summe	sc3	Querschnitt
4	Prozedurale Metakognition Wortschatz	mpvo	1	Differenzmaß Anteil korrekt	sc5 sc6	Querschnitt Querschnitt

(...)

Tabelle 3: (continued)

SC	Domäne	Kürzel	Erhebungswellen	Scoring-Typ	Suffix	Analyse
4	Lesekompetenz	re	2, 7, 10	WLE (korrigiert) WLE (unkorrigiert)	sc1 sc1u	Querschnitt Längsschnitt
4	Prozedurale Metakognition Lesen (gesamt)	mpre	2, 7, 10	Differenzmaß Anteil korrekt	sc5 sc6	Längsschnitt Längsschnitt
4	DGCF Wahrnehmungsgeschwindigkeit	dg	2, 10	Summe	sc3a	Längsschnitt
4	DGCF Schlussfolgern	dg	2, 10	Summe	sc3b	Längsschnitt
4	Deklarative Metakognition	md	2	Mittelwert	sc4	Querschnitt
4	Hörverstehen Russisch	nr	2	WLE (korrigiert)	sc1	Querschnitt
4	Prozedurale Metakognition Russisch	mpnr	2	Differenzmaß Anteil korrekt	sc5 sc6	Querschnitt Querschnitt
4	Hörverstehen Türkisch	nt	2	WLE (korrigiert)	sc1	Querschnitt
4	Prozedurale Metakognition Türkisch	mpnt	2	Differenzmaß Anteil korrekt	sc5 sc6	Querschnitt Querschnitt
4	Englisch Lesekompetenz	ef	3, 7	WLE (korrigiert) WLE (unkorrigiert)	sc1 sc1u	Querschnitt Längsschnitt
4	Prozedurale Metakognition Englisch	mpef	3, 7	Differenzmaß Anteil korrekt	sc5 sc6	Längsschnitt Längsschnitt
4	Wissenschaftspropädeutik	st	7	WLE (korrigiert)	sc1	Querschnitt
4	Prozedurale Metakognition Wissenschaftspropädeutik	mpst	7	Differenzmaß Anteil korrekt	sc5 sc6	Querschnitt Querschnitt

(...)

Tabelle 3: (continued)

SC	Domäne	Kürzel	Erhebungswellen	Scoring-Typ	Suffix	Analyse
5	Mathematische Kompetenz	ma	1, 12	WLE (korrigiert) WLE (unkorrigiert)	sc1 sc1u	Querschnitt Längsschnitt
5	Prozedurale Metakognition Mathematik	mpma	1, 12	Differenzmaß Anteil korrekt	sc5 sc6	Längsschnitt Längsschnitt
5	Lesekompetenz	re	1, 12	WLE (korrigiert) WLE (unkorrigiert)	sc1 sc1u	Querschnitt Längsschnitt
5	Prozedurale Metakognition Lesen (gesamt)	mpre	1, 12	Differenzmaß Anteil korrekt	sc5 sc6	Längsschnitt Längsschnitt
5	Lesegeschwindigkeit	rs	1	Summe	sc3	Querschnitt
5	ICT Kompetenz	ic	5	WLE (korrigiert)	sc1	Querschnitt
5	Prozedurale Metakognition ICT	mpic	5	Differenzmaß Anteil korrekt	sc5 sc6	Querschnitt Querschnitt
5	Ökonomische Kompetenz	ba	7	WLE (korrigiert)	sc1	Querschnitt
5	Englisch Lesekompetenz	ef	12	WLE (korrigiert)	sc1	Querschnitt
5	Prozedurale Metakognition Englisch	mpef	12	Differenzmaß Anteil korrekt	sc5 sc6	Querschnitt Querschnitt
6	Lesekompetenz	re	3, 5, 9	WLE (korrigiert) WLE (unkorrigiert)	sc1 sc1u	Querschnitt Längsschnitt
6	Prozedurale Metakognition Lesen (gesamt)	mpre	3, 5, 9	Differenzmaß Anteil korrekt	sc5 sc6	Längsschnitt Längsschnitt

(...)

Tabelle 3: (continued)

SC	Domäne	Kürzel	Erhebungswellen	Scoring-Typ	Suffix	Analyse
6	Lesegeschwindigkeit	rs	3, 5	Summe	sc3c	Querschnitt
6	Mathematische Kompetenz	ma	3, 9	WLE (korrigiert) WLE (unkorrigiert)	sc1 sc1u	Querschnitt Längsschnitt
6	Prozedurale Metakognition Mathematik	mpma	3, 9	Differenzmaß Anteil korrekt	sc5 sc6	Längsschnitt Längsschnitt
6	Naturwissenschaftliche Kompetenz	sc	5, 14	WLE (korrigiert) WLE (unkorrigiert)	sc1 sc1u	Querschnitt Längsschnitt
6	Prozedurale Metakognition Naturwissenschaften	mpsc	5, 14	Differenzmaß Anteil korrekt	sc5 sc6	Längsschnitt Längsschnitt
6	ICT Kompetenz	ic	5, 14	WLE (korrigiert) WLE (unkorrigiert)	sc1 sc1u	Querschnitt Längsschnitt
6	Prozedurale Metakognition ICT	mpic	5	Differenzmaß Anteil korrekt	sc5 sc6	Querschnitt Querschnitt
6	DGCF Wahrnehmungsgeschwindigkeit	dg	7	Summe	sc3a	Querschnitt
6	DGCF Schlussfolgern	dg	7	Summe	sc3b	Querschnitt
6	Wortschatz	vo	7	Summe	sc3	Querschnitt
6	Prozedurale Metakognition Wortschatz	mpvo	7	Differenzmaß Anteil korrekt	sc5 sc6	Querschnitt Querschnitt
8	Mathematische Kompetenz	ma	1	WLE (korrigiert)	sc1	Querschnitt

(...)

Tabelle 3: (continued)

SC	Domäne	Kürzel	Erhebungswellen	Scoring-Typ	Suffix	Analyse
8	Mathematische Kompetenz (Bildungstrend)	mb	1	WLE (Bildungstrend Metrik)	sc1b	Querschnitt
8	Lesekompetenz	re	1	WLE (korrigiert)	sc1	Querschnitt
8	Lesekompetenz (Bildungstrend)	rb	1	WLE (Bildungstrend Metrik)	sc1b	Querschnitt
8	Allgemeinwissen Allgemeinwissen Geisteswissenschaften Allgemeinwissen Naturwissenschaften Allgemeinwissen Sozialwissenschaften	gk	2	Summe	sc3 sc3a sc3b sc3c	Querschnitt Querschnitt Querschnitt Querschnitt
8	Prozedurale Metakognition Allgemeinwissen (gesamt)	mpgk	2	Differenzmaß Anteil korrekt	sc5 sc6	Querschnitt Querschnitt
8	Wortschatz	vo	2	Summe	sc3	Querschnitt
8	Hörverstehen Russisch	nr	2	WLE (korrigiert)	sc1	Querschnitt
8	Hörverstehen Türkisch	nt	2	WLE (korrigiert)	sc1	Querschnitt
8	ICT Kompetenz	ic	2	WLE (korrigiert)	sc1	Querschnitt
8	Digitale Kompetenz	dc	2	WLE (korrigiert)	sc1	Querschnitt

A.4 Zeitkonstante Variablen

Die nachfolgende Auflistung enthält alle Variablen – sortiert nach Startkohorten – mit Merkmalen, die sich über die Zeit nicht oder nur sehr selten verändern. Diese Variablen wurden für NEPScomp insofern editiert, als dass gültige Werte aus einer Erhebung in einer bestimmten Welle auch in die Wellen zuvor und danach übertragen wurden (siehe Kapitel 3.3).

Tabelle 4: Überblick über alle „aufgefüllten“ Variablen

Kohorte	Variable	Label
SC1	p201451	Beginn TV Konsum
SC1	p400000_g1	Geburtsland Befragter (Deutschland/Ausland; bereinigt)
SC1	p400010_g2	Geburtsland im Ausland Befragte/r (kategorisiert)
SC1	p400015	Befragter in Ost- oder Westdeutschland geboren?
SC1	p40003m	Zuzugsdatum nach Deutschland Befragter: Monat
SC1	p40003y	Zuzugsdatum nach Deutschland Befragter: Jahr
SC1	p400070_g2	Geburtsland Mutter Befragter (kategorisiert)
SC1	p400090_g2	Geburtsland Vater Befragter (kategorisiert)
SC1	p400500_g3	Herkunftsgruppe
SC1	p40113m	Erwerb deutscher Staatsangehörigkeit Befragte/r: Monat
SC1	p40113y	Erwerb deutscher Staatsangehörigkeit Befragte/r: Jahr
SC1	p403000_g1	Geburtsland Partner (Deutschland/Ausland; bereinigt)
SC1	p403010_g2	Geburtsland im Ausland Partner (kategorisiert)
SC1	p403015	Partner in Ost- oder Westdeutschland geboren?
SC1	p403030	Alter Zuzug nach Deutschland Partner
SC1	p403070_g2	Geburtsland Mutter des Partners (kategorisiert)
SC1	p403090_g2	Geburtsland Vater des Partners (kategorisiert)
SC1	p40403m	Datum Erwerb deutsche Staatsangehörigkeit Partner: Monat
SC1	p40403y	Datum Erwerb deutsche Staatsangehörigkeit Partner: Jahr
SC1	p406000_g1	Geburtsland des Zielkindes (Deutschland/Ausland; bereinigt)
SC1	p406010_g2	Geburtsland des Zielkindes (kategorisiert)
SC1	p40603m	Zuzugsdatum des Zielkindes nach Deutschland (Monat)
SC1	p40603y	Zuzugsdatum des Zielkindes nach Deutschland (Jahr)
SC1	p410000_g1D	Erstsprache/Muttersprache Kind (deutsch/nicht deutsch)
SC1	p410002_g1D	Weitere Erst-/Muttersprache Kind (deutsch/nicht deutsch)
SC1	p41002m	Beginn Deutsch lernen (Monat)
SC1	p41002y	Beginn Deutsch lernen (Jahr)
SC1	p410030	Erstsprache/Muttersprache ermitteln – bilinguales Kind, über Kind
SC1	p410040	Deutschsprachigkeit Kind (Autovariable)
SC1	p410050	Bilingualität Kind (Autovariable)
SC1	p410070	Muttersprache Zielkind
SC1	p413000_g1D	Erstsprache/Muttersprache Befragter(deutsch/nicht deutsch)
SC1	p413000_g2	Erstsprache/Muttersprache Befragter(kategorisiert)
SC1	p413002_g1D	Weitere Erst-/Muttersprache Befragter(deutsch/nicht deutsch)
SC1	p413002_g2	Weitere Erst-/Muttersprache Befragter(kategorisiert)
SC1	p413030	Herkunftssprache ermitteln - bilingualer befragter Elternteil
SC1	p413040	Deutschsprachigkeit befragter Elternteil (Autovariable)
SC1	p413050	Bilingualität befragter Elternteil (Autovariable)
SC1	p414000_g1D	Erstsprache/Muttersprache Partner (deutsch/nicht deutsch)
SC1	p414000_g2	Erstsprache/Muttersprache Partner (kategorisiert)

(...)

Tabelle 4: (continued)

Kohorte	Variable	Label
SC1	p414002_g1D	Weitere Erst-/Muttersprache Partner (deutsch/nicht deutsch)
SC1	p414002_g2	Weitere Erst-/Muttersprache Partner (kategorisiert)
SC1	p414030	Erstsprache/Muttersprache ermitteln - bilingualer Partner
SC1	p414040	Deutschsprachigkeit Partner/ Partnerin (Autovariable)
SC1	p414050	Partner bilingual (Autovariable)
SC1	p529000_D	Gewicht bei Geburt (aggregiert)
SC1	p529001_D	Größe bei Geburt (aggregiert)
SC1	p529007	Gewicht in Gramm bei der U4
SC1	p529008	Größe in Zentimetern bei der U4
SC1	p529010	Datum der U4: Monat
SC1	p529011	Datum der U4: Jahr
SC1	p529012	Gewicht des Kindes in Gramm bei der U6
SC1	p529013	Größe des Kindes in Zentimetern bei der U6
SC1	p529014	Datum der U6: Monat
SC1	p529015	Datum der U6: Jahr
SC1	p529020	Gewicht des Kindes in Gramm bei der U7
SC1	p529021	Größe des Kindes in Zentimetern bei der U7
SC1	p529022	Datum der U7: Monat
SC1	p529023	Datum der U7: Jahr
SC1	p529024	Gewicht des Kindes in Gramm bei der U7a
SC1	p529025	Größe des Kindes in Zentimetern bei der U7a
SC1	p529026	Datum der U7a: Monat
SC1	p529027	Datum der U7a: Jahr
SC1	p529028	Gewicht des Kindes in Gramm bei der U8
SC1	p529029	Größe des Kindes in Zentimetern bei der U8
SC1	p529030	Datum der U8: Monat
SC1	p529031	Datum der U8: Jahr
SC1	p529032	Gewicht des Kindes in Gramm bei der U9
SC1	p529033	Größe des Kindes in Zentimetern bei der U9
SC1	p529034	Datum der U9: Monat
SC1	p529035	Datum der U9: Jahr
SC1	p529100	Schwangerschaftswoche bei Geburt
SC1	p529108	Verlegung in Kinderklinik
SC1	p529109	Aufenthaltszeit in Klinik
SC1	p529110	Geburtskomplikationen: Saugglocke oder Zangengeburt
SC1	p529111	Geburtskomplikationen: Notkaiserschnitt
SC1	p529112	Geburtskomplikationen: Frühgeburt
SC1	p529113	Geburtskomplikationen: Sauerstoffmangel
SC1	p529114	Geburtskomplikationen: Beckenendlage
SC1	p529115	Geburtskomplikationen: keine Komplikationen
SC1	p700010	Geschlecht Zielkind
SC1	p70012m	Geburtsdatum Zielkind (Monat)
SC1	p70012y	Geburtsdatum Zielkind (Jahr)
SC1	p71203m	Datum Einschulung Zielkind: Monat
SC1	p71203y	Datum Einschulung Zielkind: Jahr
SC1	p72802m	Datum Feststellung sonderpädagogischer Förderbedarf: Monat
SC1	p72802y	Datum Feststellung sonderpädagogischer Förderbedarf: Jahr
SC1	p731116	Geschlecht Partner
SC1	p73111m	Beginn Wohnen mit Partner: Monat

(...)

Tabelle 4: (continued)

Kohorte	Variable	Label
SC1	p73111y	Beginn Wohnen mit Partner: Jahr
SC1	p73157m	Beginn berufliche Ausbildung: Monat
SC1	p73157y	Beginn berufliche Ausbildung: Jahr
SC1	p73167m	Beginn berufliche Ausbildung Partner: Monat
SC1	p73167y	Beginn berufliche Ausbildung Partner: Jahr
SC1	p731702	Geschlecht Auskunftsperson
SC1	p73170m	Geburtsdatum Befragter: Monat
SC1	p73170y	Geburtsdatum Befragter: Jahr
SC1	p73175y	Geburtsjahr Partner
–	–	–
SC2	p400000_g1	Geburtsland Befragter (Deutschland/Ausland; bereinigt)
SC2	p40003m	Zuzugsdatum nach Deutschland Befragter: Monat
SC2	p40003y	Zuzugsdatum nach Deutschland Befragter: Jahr
SC2	p400070_g1D	Geburtsland Mutter Befragter (Deutschland/Ausland; bereinigt)
SC2	p400090_g1D	Geburtsland Vater Befragter (Deutschland/Ausland; bereinigt)
SC2	p40113m	Datum Erwerb der deutschen Staatsangehörigkeit Befragter: Monat
SC2	p40113y	Datum Erwerb der deutschen Staatsangehörigkeit Befragter: Jahr
SC2	p403000_g1	Geburtsland Partner (Deutschland/Ausland; bereinigt)
SC2	p403030	Alter Zuzug nach Deutschland Partner
SC2	p403070_g1D	Geburtsland Mutter des Partners (Deutschland/Ausland; bereinigt)
SC2	p403090_g1D	Geburtsland Vater des Partners (Deutschland/Ausland; bereinigt)
SC2	p40403m	Datum Erwerb deutsche Staatsangehörigkeit Partner: Monat
SC2	p40403y	Datum Erwerb deutsche Staatsangehörigkeit Partner: Jahr
SC2	p406000_g1	Geburtsland des Zielkindes (Deutschland/Ausland; bereinigt)
SC2	p40603m	Zuzugsdatum des Zielkindes nach Deutschland (Monat)
SC2	p40603y	Zuzugsdatum des Zielkindes nach Deutschland (Jahr)
SC2	p410000_g1D	Erstsprache/Muttersprache Kind (deutsch/nicht deutsch)
SC2	p410002_g1D	Weitere Erst-/Muttersprache Kind (deutsch/nicht deutsch)
SC2	p41002m	Beginn Deutsch lernen (Monat)
SC2	p41002y	Beginn Deutsch lernen (Jahr)
SC2	p410030	Erstsprache/Muttersprache ermitteln – bilinguales Kind, über Kind
SC2	p413000_g1D	Erstsprache/Muttersprache Befragter(deutsch/nicht deutsch)
SC2	p413002_g1D	Weitere Erst-/Muttersprache Befragter(deutsch/nicht deutsch)
SC2	p413030	Herkunftssprache ermitteln - bilingualer befragter Elternteil
SC2	p414000_g1D	Erstsprache/Muttersprache Partner (deutsch/nicht deutsch)
SC2	p414002_g1D	Weitere Erst-/Muttersprache Partner (deutsch/nicht deutsch)
SC2	p414030	Erstsprache/Muttersprache ermitteln - bilingualer Partner
SC2	p414100_g1D	Erstsprache/Muttersprache neuer Partner (deutsch/nicht deutsch)
SC2	p414102_g1D	Weitere Erst-/Muttersprache neuer Partner (deutsch/nicht deutsch)
SC2	p414130	Erst-/Muttersprache ermitteln - bilingualer neuer Partner
SC2	p529000	Gewicht bei Geburt
SC2	p529001	Größe bei Geburt
SC2	p529101	Stadium der Entwicklung bei Geburt (Frühgeborenes)
SC2	p529102	Probleme nach der Geburt
SC2	p529108	Verlegung in Kinderklinik
SC2	p700010	Geschlecht Zielkind
SC2	p70012m	Geburtsdatum Zielkind (Monat)
SC2	p70012y	Geburtsdatum Zielkind (Jahr)
SC2	p71202m	Datum erster Kindergartenbesuch (Monat)

(...)

Tabelle 4: (continued)

Kohorte	Variable	Label
SC2	p71202y	Datum erster Kindergartenbesuch (Jahr)
SC2	p71203m	Datum Einschulung Zielkind (Monat)
SC2	p71203y	Datum Einschulung Zielkind (Jahr)
SC2	p72802m	Datum Feststellung sonderpädagogischer Förderbedarf (Monat)
SC2	p72802y	Datum Feststellung sonderpädagogischer Förderbedarf (Jahr)
SC2	p731116	Geschlecht Partner
SC2	p73111m	Beginn Wohnen mit Partner: Monat
SC2	p73111y	Beginn Wohnen mit Partner: Jahr
SC2	p731702	Geschlecht Auskunftsperson
SC2	p73170y	Geburtsdatum Befragter: Jahr
SC2	p73175y	Geburtsjahr Partner
SC2	pb1002a	Geburtsjahr Geschwister im HH 1
SC2	pb1002b	Geburtsjahr Geschwister im HH 2
SC2	pb1002c	Geburtsjahr Geschwister im HH 3
SC2	pb1002d	Geburtsjahr Geschwister im HH 4
SC2	pb1002e	Geburtsjahr Geschwister im HH 5
SC2	pb1002f	Geburtsjahr Geschwister im HH 6
SC2	pb1002g	Geburtsjahr Geschwister im HH 7
SC2	pb1002h	Geburtsjahr Geschwister im HH 8
SC2	pb1002i	Geburtsjahr Geschwister im HH 9
SC2	pb1002j	Geburtsjahr Geschwister im HH 10
SC2	t400000_g1D	Geburtsland (Deutschland/Ausland)
SC2	t41000a_g1	Muttersprache (Anzahl Nennungen)
SC2	t41000a_g2D	Muttersprache (Nennung 1, vergröbert)
SC2	t41000a_g3D	Muttersprache (Nennung 2, vergröbert)
SC2	t41000a_g4D	Muttersprache (Nennung 3, vergröbert)
SC2	t41000a_g5D	Muttersprache (Nennung 4, vergröbert)
SC2	t428500	Migrationsbio
–	–	–
SC3	p400000_g1	Geburtsland Befragter (Deutschland/Ausland; bereinigt)
SC3	p40003m	Zuzugsdatum nach Deutschland Befragter (Monat)
SC3	p40003y	Zuzugsdatum nach Deutschland Befragter (Jahr)
SC3	p400070_g1D	Geburtsland Mutter Befragter (Deutschland/Ausland; bereinigt)
SC3	p400090_g1D	Geburtsland Vater Befragter (Deutschland/Ausland; bereinigt)
SC3	p40113m	Datum Erwerb der deutschen Staatsangehörigkeit Befragter (Monat)
SC3	p40113y	Datum Erwerb der deutschen Staatsangehörigkeit Befragter (Jahr)
SC3	p403000_g1	Geburtsland Partner (Deutschland/Ausland; bereinigt)
SC3	p403030	Alter Partner bei Zuzug nach Deutschland
SC3	p403070_g1D	Geburtsland Mutter des Partners (Deutschland/Ausland; bereinigt)
SC3	p403090_g1D	Geburtsland Vater des Partners (Deutschland/Ausland; bereinigt)
SC3	p40403m	Erwerb deutsche Staatsangehörigkeit Partner, Datum (Monat)
SC3	p40403y	Erwerb deutsche Staatsangehörigkeit Partner, Datum (Jahr)
SC3	p406000_g1	Geburtsland des Zielkindes (Deutschland/Ausland; bereinigt)
SC3	p40603m	Zuzugsdatum des Zielkindes nach Deutschland (Monat)
SC3	p40603y	Zuzugsdatum (Zuzugsjahr) des Zielkindes nach Deutschland
SC3	p410000_g1D	Erstsprache/Muttersprache Kind (deutsch/nicht deutsch)
SC3	p410002_g1D	Weitere Erst-/Muttersprache Kind (deutsch/nicht deutsch)
SC3	p41002m	Beginn Deutsch lernen (Monat)
SC3	p41002y	Beginn Deutsch lernen (Jahr)

(...)

Tabelle 4: (continued)

Kohorte	Variable	Label
SC3	p410030	Erstsprache/Muttersprache ermitteln – bilinguales Kind, über Kind
SC3	p410040	Deutschsprachigkeit Kind (Autovariable)
SC3	p410050	Bilingualität Kind (Autovariable)
SC3	p413000_g1D	Erstsprache/Muttersprache Befragter(deutsch/nicht deutsch)
SC3	p413002_g1D	Weitere Erst-/Muttersprache Befragter(deutsch/nicht deutsch)
SC3	p413030	Herkunftssprache ermitteln - bilingualer befragter Elternteil
SC3	p413040	Deutschsprachigkeit befragter Elternteil (Autovariable)
SC3	p413050	Bilingualität befragter Elternteil (Autovariable)
SC3	p414000_g1D	Erstsprache/Muttersprache Partner (deutsch/nicht deutsch)
SC3	p414002_g1D	Weitere Erst-/Muttersprache Partner (deutsch/nicht deutsch)
SC3	p414030	Erstsprache/Muttersprache ermitteln - bilingualer Partner
SC3	p414040	Deutschsprachigkeit Partner/ Partnerin (Autovariable)
SC3	p414050	Bilingualität Partner/ Partnerin (Autovariable)
SC3	p414100_g1D	Erstsprache/Muttersprache neuer Partner (deutsch/nicht deutsch)
SC3	p414102_g1D	Weitere Erst-/Muttersprache neuer Partner (deutsch/nicht deutsch)
SC3	p414130	Erst-/Muttersprache ermitteln - bilingualer neuer Partner
SC3	p529000	Gewicht bei Geburt
SC3	p529001	Größe bei Geburt
SC3	p529101	Stadium der Entwicklung bei Geburt (Frühgeborenes)
SC3	p529102	Probleme nach der Geburt
SC3	p529108	Verlegung in Kinderklinik
SC3	p700010	Geschlecht Zielkind
SC3	p70012m	Geburtsdatum Zielkind (Monat)
SC3	p70012y	Geburtsdatum Zielkind (Jahr)
SC3	p71202m	Datum erster Kindergartenbesuch (Monat)
SC3	p71202y	Datum erster Kindergartenbesuch (Jahr)
SC3	p71203m	Datum Einschulung Zielkind (Monat)
SC3	p71203y	Datum Einschulung Zielkind (Jahr)
SC3	p72802m	Datum Feststellung sonderpädagogischer Förderbedarf (Monat)
SC3	p72802y	Datum Feststellung sonderpädagogischer Förderbedarf (Jahr)
SC3	p731116	Geschlecht Partner
SC3	p73111m	Beginn Gemeinsames Wohnen mit Partner (Monat)
SC3	p73111y	Beginn Gemeinsames Wohnen mit Partner (Jahr)
SC3	p731702	Geschlecht Auskunftsperson
SC3	p73170y	Geburtsdatum Befragter: Jahr
SC3	p73175y	Geburtsjahr Partner
SC3	t400000_g1D	Geburtsland (Deutschland/Ausland)
SC3	t400030	Alter bei Ankunft in Deutschland
SC3	t400070_g1D	Mutter: Geburtsland (Deutschland/Ausland)
SC3	t400090_g1D	Vater: Geburtsland (Deutschland/Ausland)
SC3	t400220_g1D	Mutter der Mutter: Geburtsland (Deutschland/Ausland)
SC3	t400240_g1D	Vater der Mutter: Geburtsland (Deutschland/Ausland)
SC3	t400260_g1D	Mutter des Vaters: Geburtsland (Deutschland/Ausland)
SC3	t400280_g1D	Vater des Vaters: Geburtsland (Deutschland/Ausland)
SC3	t405000	Geburt in Deutschland oder im Ausland
SC3	t405500_g1	Befragter in Deutschland geboren? (bereinigt)
SC3	t40552m	Zuzugsdatum nach Deutschland Befragter: Monat
SC3	t40552y	Zuzugsdatum nach Deutschland Befragter: Jahr
SC3	t405560_g1D	Geburtsland Mutter Befragter (Deutschland/Ausland)

(...)

Tabelle 4: (continued)

Kohorte	Variable	Label
SC3	t405570_g1D	Geburtsland Großmutter mütterlicherseits (Deutschland/Ausland)
SC3	t405580_g1D	Geburtsland Großvater mütterlicherseits (Deutschland/Ausland)
SC3	t405600_g1D	Geburtsland Vater Befragter (Deutschland/Ausland)
SC3	t405610_g1D	Geburtsland Großmutter väterlicherseits (Deutschland/Ausland)
SC3	t405620_g1D	Geburtsland Großvater väterlicherseits (Deutschland/Ausland)
SC3	t40607m	Einbürgerungsmonat
SC3	t40607y	Einbürgerungsjahr
SC3	t41000a_g1	Muttersprache (Anzahl Nennungen)
SC3	t41000a_g2D	Muttersprache (Nennung 1, vergröbert)
SC3	t41000a_g3D	Muttersprache (Nennung 2, vergröbert)
SC3	t41000a_g4D	Muttersprache (Nennung 3, vergröbert)
SC3	t41000a_g5D	Muttersprache (Nennung 4, vergröbert)
SC3	t410010_g1	Zweitsprache (Anzahl Nennungen)
SC3	t410010_g2D	Zweitsprache (Nennung 1, vergröbert)
SC3	t410010_g3D	Zweitsprache (Nennung 2, vergröbert)
SC3	t410010_g4D	Zweitsprache (Nennung 3, vergröbert)
SC3	t410010_g5D	Zweitsprache (Nennung 4, vergröbert)
SC3	t410020	Alter Spracherwerb
SC3	t41010a_g1	Mutter: Muttersprache (Anzahl Nennungen)
SC3	t41010a_g2D	Mutter: Muttersprache (Nennung 1, vergröbert)
SC3	t41010a_g3D	Mutter: Muttersprache (Nennung 2, vergröbert)
SC3	t41010a_g4D	Mutter: Muttersprache (Nennung 3, vergröbert)
SC3	t41010a_g5D	Mutter: Muttersprache (Nennung 4, vergröbert)
SC3	t41012a_g1	Vater: Muttersprache (Anzahl Nennungen)
SC3	t41012a_g2D	Vater: Muttersprache (Nennung 1, vergröbert)
SC3	t41012a_g3D	Vater: Muttersprache (Nennung 2, vergröbert)
SC3	t41012a_g4D	Vater: Muttersprache (Nennung 3, vergröbert)
SC3	t41012a_g5D	Vater: Muttersprache (Nennung 4, vergröbert)
SC3	t414000_g1D	erste Muttersprache (deutsch/nicht deutsch)
SC3	t414002_g1D	zweite Muttersprache (deutsch/nicht deutsch)
SC3	t414030	Herkunftssprache ermitteln - bilinguales Target
SC3	t428500	Migrationshintergrund (nicht in Deutschland geboren)
SC3	t70000d	Geburtsdatum - Tag
SC3	t700031	Geschlecht
SC3	t70004m	Geburtsdatum - Monat
SC3	t70004y	Geburtsdatum - Jahr
SC3	t700101_g1	Geburtsgemeinde (West/Ost)
SC3	t731101	Familienform bis zum 15. Lebensjahr
SC3	t731204	Geburtsdatum Mutter
SC3	t731254	Geburtsdatum Vater
SC3	tf32302	erster Auszug aus Elternhaus
SC3	tf3231m	Haushaltsgründungsmonat
SC3	tf3231y	Haushaltsgründungsjahr
SC3	tf3233m	Auszugsmonat aus Elternhaus
SC3	tf3233y	Auszugsjahr aus Elternhaus
–	–	–
SC4	p400000_g1	Geburtsland Befragte/r (Deutschland/Ausland; bereinigt)
SC4	p40003m	Zuzugsdatum nach Deutschland Befragter (Monat)
SC4	p40003y	Zuzugsdatum nach Deutschland Befragter (Jahr)

(...)

Tabelle 4: (continued)

Kohorte	Variable	Label
SC4	p400070_g1D	Geburtsland Mutter Befragter (Deutschland/Ausland; bereinigt)
SC4	p400090_g1D	Geburtsland Vater Befragter (Deutschland/Ausland; bereinigt)
SC4	p40113m	Datum Erwerb der deutschen Staatsangehörigkeit Befragter (Monat)
SC4	p40113y	Datum Erwerb der deutschen Staatsangehörigkeit Befragter (Jahr)
SC4	p403000_g1	Geburtsland Partner (Deutschland/Ausland; bereinigt)
SC4	p403030	Alter Partner bei Zuzug nach Deutschland
SC4	p403070_g1D	Geburtsland Mutter des Partners (Deutschland/Ausland; bereinigt)
SC4	p403090_g1D	Geburtsland Vater des Partners (Deutschland/Ausland; bereinigt)
SC4	p40403m	Erwerb deutsche Staatsangehörigkeit Partner, Datum (Monat)
SC4	p40403y	Erwerb deutsche Staatsangehörigkeit Partner, Datum (Jahr)
SC4	p406000_g1	Geburtsland des Zielkindes (Deutschland/Ausland; bereinigt)
SC4	p40603m	Zuzugsdatum (Zuzugsmonat) des Zielkindes nach Deutschland
SC4	p40603y	Zuzugsdatum (Zuzugsjahr) des Zielkindes nach Deutschland
SC4	p410000_g1D	Erstsprache/Muttersprache Kind (deutsch/nicht deutsch)
SC4	p410002_g1D	Weitere Erst-/Muttersprache Kind (deutsch/nicht deutsch)
SC4	p41002m	Beginn Deutsch lernen (Monat)
SC4	p41002y	Beginn Deutsch lernen (Jahr)
SC4	p410030	Herkunftssprache ermitteln – bilinguales Kind, über Kind
SC4	p410040	Deutschsprachigkeit Kind (Autovariable)
SC4	p410050	Bilingualität Kind (Autovariable)
SC4	p413000_g1D	Erstsprache/Muttersprache Befragter(deutsch/nicht deutsch)
SC4	p413002_g1D	Weitere Erst-/Muttersprache Befragter(deutsch/nicht deutsch)
SC4	p413030	Herkunftssprache ermitteln - bilingualer befragter Elternteil
SC4	p414000_g1D	Erstsprache/Muttersprache Partner (deutsch/nicht deutsch)
SC4	p414002_g1D	Weitere Erst-/Muttersprache Partner (deutsch/nicht deutsch)
SC4	p414030	Herkunftssprache ermitteln - bilingualer Partner
SC4	p414100_g1D	Erstsprache/Muttersprache neuer Partner (deutsch/nicht deutsch)
SC4	p414102_g1D	Weitere Erst-/Muttersprache neuer Partner (deutsch/nicht deutsch)
SC4	p414130	Erst-/Muttersprache ermitteln - bilingualer neuer Partner
SC4	p529000	Gewicht bei Geburt
SC4	p529001	Größe bei Geburt
SC4	p529101	Stadium der Entwicklung bei Geburt (Frühgeborenes)
SC4	p529102	Probleme nach der Geburt
SC4	p529108	Verlegung in Kinderklinik
SC4	p700010	Geschlecht Zielkind
SC4	p70012m	Geburtsdatum Zielkind (Monat)
SC4	p70012y	Geburtsdatum Zielkind (Jahr)
SC4	p71202m	Datum erster Kindergartenbesuch (Monat)
SC4	p71202y	Datum erster Kindergartenbesuch (Jahr)
SC4	p71203m	Datum Einschulung Zielkind (Monat)
SC4	p71203y	Datum Einschulung Zielkind (Jahr)
SC4	p72802m	Datum Feststellung sonderpädagogischer Förderbedarf (Monat)
SC4	p72802y	Datum Feststellung sonderpädagogischer Förderbedarf (Jahr)
SC4	p731116	Geschlecht Partner
SC4	p73111m	Beginn Gemeinsames Wohnen mit Partner (Monat)
SC4	p73111y	Beginn Gemeinsames Wohnen mit Partner (Jahr)
SC4	p731702	Geschlecht Auskunftsperson
SC4	p73170y	Geburtsjahr Befragter
SC4	p73175y	Geburtsjahr Partner

(...)

Tabelle 4: (continued)

Kohorte	Variable	Label
SC4	t400000_g1D	Geburtsland (Deutschland/Ausland)
SC4	t400030	Alter bei Ankunft in Deutschland
SC4	t400070_g1D	Mutter: Geburtsland (Deutschland/Ausland)
SC4	t400090_g1D	Vater: Geburtsland (Deutschland/Ausland)
SC4	t400220_g1D	Mutter der Mutter: Geburtsland (Deutschland/Ausland)
SC4	t400240_g1D	Vater der Mutter: Geburtsland (Deutschland/Ausland)
SC4	t400260_g1D	Mutter des Vaters: Geburtsland (Deutschland/Ausland)
SC4	t400280_g1D	Vater des Vaters: Geburtsland (Deutschland/Ausland)
SC4	t405000_g1	Geburt in Deutschland oder im Ausland (bereinigt)
SC4	t405500_g1	Befragter in Deutschland geboren? (bereinigt)
SC4	t40552m	Zuzugsdatum nach Deutschland Befragter (Monat)
SC4	t40552y	Zuzugsdatum nach Deutschland Befragter (Jahr)
SC4	t405560_g1D	Geburtsland Mutter Befragter (Deutschland/Ausland)
SC4	t405570_g1D	Geburtsland Großmutter mütterlicherseits (Deutschland/Ausland)
SC4	t405580_g1D	Geburtsland Großvater mütterlicherseits (Deutschland/Ausland)
SC4	t405600_g1D	Geburtsland Vater Befragter (Deutschland/Ausland)
SC4	t405610_g1D	Geburtsland Großmutter väterlicherseits (Deutschland/Ausland)
SC4	t405620_g1D	Geburtsland Großvater väterlicherseits (Deutschland/Ausland)
SC4	t40607m	Einbürgerungsmonat
SC4	t40607y	Einbürgerungsjahr
SC4	t407000_g1	Partner geboren In-/Ausland (bereinigt)
SC4	t41000a_g1	Muttersprache (Anzahl Nennungen)
SC4	t41000a_g2D	Muttersprache (Nennung 1, vergröbert)
SC4	t41000a_g3D	Muttersprache (Nennung 2, vergröbert)
SC4	t41000a_g4D	Muttersprache (Nennung 3, vergröbert)
SC4	t41000a_g5D	Muttersprache (Nennung 4, vergröbert)
SC4	t410010_g1	Zweitsprache (Anzahl Nennungen)
SC4	t410010_g2D	Zweitsprache (Nennung 1, vergröbert)
SC4	t410010_g3D	Zweitsprache (Nennung 2, vergröbert)
SC4	t410010_g4D	Zweitsprache (Nennung 3, vergröbert)
SC4	t410010_g5D	Zweitsprache (Nennung 4, vergröbert)
SC4	t410020	Alter Spracherwerb
SC4	t41010a_g1	Mutter: Muttersprache (Anzahl Nennungen)
SC4	t41010a_g2D	Mutter: Muttersprache (Nennung 1, vergröbert)
SC4	t41010a_g3D	Mutter: Muttersprache (Nennung 2, vergröbert)
SC4	t41010a_g4D	Mutter: Muttersprache (Nennung 3, vergröbert)
SC4	t41010a_g5D	Mutter: Muttersprache (Nennung 4, vergröbert)
SC4	t41012a_g1	Vater: Muttersprache (Anzahl Nennungen)
SC4	t41012a_g2D	Vater: Muttersprache (Nennung 1, vergröbert)
SC4	t41012a_g3D	Vater: Muttersprache (Nennung 2, vergröbert)
SC4	t41012a_g4D	Vater: Muttersprache (Nennung 3, vergröbert)
SC4	t41012a_g5D	Vater: Muttersprache (Nennung 4, vergröbert)
SC4	t414000_g1D	erste Muttersprache (deutsch/nicht deutsch)
SC4	t414002_g1D	zweite Muttersprache (deutsch/nicht deutsch)
SC4	t414030	Herkunftssprache ermitteln - bilinguales Target
SC4	t700001	Geschlecht
SC4	t70000m	Geburtsdatum: Monat
SC4	t70000y	Geburtsdatum: Jahr
SC4	t700031	Geschlecht

(...)

Tabelle 4: (continued)

Kohorte	Variable	Label
SC4	t70004m	Geburtsdatum - Monat
SC4	t70004y	Geburtsdatum - Jahr
SC4	t700101_g1	Geburtsgemeinde (West/Ost)
SC4	t731101	Familienform bis zum 15. Lebensjahr
SC4	t731205	Geburtsdatum Mutter (1)
SC4	t731206	Geburtsdatum Mutter (2)
SC4	t731255	Geburtsdatum Vater (1)
SC4	t731256	Geburtsdatum Vater (2)
SC4	tf32302	erster Auszug aus Elternhaus
SC4	tf32311	Geschlecht des/der Partner/in
SC4	tf3231m	Haushaltsgründungsmonat
SC4	tf3231y	Haushaltsgründungsjahr
SC4	tf3232y	Geburtsjahr Partner/in
SC4	tf3233m	Auszugsmonat aus Elternhaus
SC4	tf3233y	Auszugsjahr aus Elternhaus
–	–	–
SC5	t405000_g1	Geburt in Deutschland oder im Ausland (bereinigt)
SC5	t405010_g2	Land des Geburtsorts (kategorisiert)
SC5	t40503m	Zuzugsmonat
SC5	t40503y	Zuzugsjahr
SC5	t405060_g1	Geburtsort Mutter (bereinigt)
SC5	t405070_g2	Mutter: Geburtsland (kategorisiert)
SC5	t405080_g1	Mutter lebt in Deutschland (bereinigt)
SC5	t405090_g1	Geburtsort Vater (bereinigt)
SC5	t405100_g2	Vater: Geburtsland (kategorisiert)
SC5	t405200	Großelternteil im Ausland geboren und nach 1950 nach Deutschland zugezogen
SC5	t40521a_g1	Mutter der Mutter im Ausland geboren und nach 1950 zugezogen (bereinigt)
SC5	t40521b_g1	Mutter des Vaters im Ausland geboren und nach 1950 zugezogen (bereinigt)
SC5	t40521c_g1	Vater der Mutter im Ausland geboren und nach 1950 zugezogen (bereinigt)
SC5	t40521d_g1	Vater des Vaters im Ausland geboren und nach 1950 zugezogen (bereinigt)
SC5	t405220_g2	Mutter der Mutter: Geburtsland (kategorisiert)
SC5	t405230_g2	Vater der Mutter: Geburtsland (kategorisiert)
SC5	t405240_g2	Mutter des Vaters: Geburtsland (kategorisiert)
SC5	t405250_g2	Vater des Vaters: Geburtsland (kategorisiert)
SC5	t40607m	Einbürgerungsdatum (Monat)
SC5	t40607y	Einbürgerungsdatum (Jahr)
SC5	t414000_g2	erste Muttersprache (kategorisiert)
SC5	t414002_g2	zweite Muttersprache (kategorisiert)
SC5	t414020	Beginn Zweitspracherwerb Deutsch
SC5	t414030	Hauptsprache bilinguale ZP
SC5	t414060_g2	Hilfsvariable: Herkunftssprache Target (kategorisiert)
SC5	t414100_g2	Mutter: Erstsprache/Muttersprache (kategorisiert)
SC5	t414120_g2	Vater: Erstsprache/Muttersprache (kategorisiert)
SC5	t42800x_g2	Herkunftsland (kategorisiert)
SC5	t42803x	Hilfsvariable: Geburtsland Mutter SU, TRK
SC5	t42804x	Hilfsvariable: Geburtsland Vater SU, TRK
SC5	t700001	Geschlecht
SC5	t70000m	Geburtsdatum: Monat
SC5	t70000y	Geburtsdatum: Jahr

(...)

Tabelle 4: (continued)

Kohorte	Variable	Label
SC5	t700101_g1	Geburtsgemeinde (West/Ost)
SC5	t731101	Familienform bis zum 15. Lebensjahr
SC5	t731204	Geburtsjahr Mutter
SC5	t731254	Geburtsjahr Vater
SC5	ts31203	Geschlecht des/der Partners/in
–	–	–
SC6	t405010_g2	Land des Geburtsorts (kategorisiert)
SC6	t40503m	Zuzugsdatum Monat
SC6	t40503y	Zuzugsdatum Jahr
SC6	t405060_ha	Mutter: Geburtsort (harmonisiert)
SC6	t405070_g2	Mutter: Geburtsland (kategorisiert)
SC6	t405090_ha	Vater: Geburtsort (harmonisiert)
SC6	t405100_g2	Vater: Geburtsland (kategorisiert)
SC6	t405200	Großelternteil im Ausland geboren und nach 1950 nach Deutschland zugezogen
SC6	t40521a_g1	Mutter der Mutter im Ausland geboren und nach 1950 zugezogen (bereinigt)
SC6	t40521b_g1	Mutter des Vaters im Ausland geboren und nach 1950 zugezogen (bereinigt)
SC6	t40521c_g1	Vater der Mutter im Ausland geboren und nach 1950 zugezogen (bereinigt)
SC6	t40521d_g1	Vater des Vaters im Ausland geboren und nach 1950 zugezogen (bereinigt)
SC6	t405220_g2	Mutter der Mutter: Geburtsland (kategorisiert)
SC6	t405230_g2	Vater der Mutter: Geburtsland (kategorisiert)
SC6	t405240_g2	Mutter des Vaters: Geburtsland (kategorisiert)
SC6	t405250_g2	Vater des Vaters: Geburtsland (kategorisiert)
SC6	t40607m	Einbürgerungsmonat
SC6	t40607y	Einbürgerungsjahr
SC6	t407000_g1	Partner geboren In-/Ausland (bereinigt)
SC6	t413020	Alter Beginn Deutschlernen
SC6	t413100_g2	Erstsprache/Muttersprache Mutter (kategorisiert)
SC6	t413120_g2	Erstsprache/Muttersprache Vater (kategorisiert)
SC6	t700001	Geschlecht
SC6	t70000m	Geburtsmonat
SC6	t70000y	Geburtsjahr
SC6	t700101_g1	Geburtsgemeinde (West/Ost)
SC6	t731101	Familienform bis zum 15. Lebensjahr
SC6	t73120y	Geburtsdatum Mutter
SC6	t73125y	Geburtsdatum Vater
SC6	tf32311	Geschlecht des/der Partners/in (LAT)
SC6	tf3232y	Geburtsdatum (-monat/-jahr) Partner/in (LAT) (Jahr)
–	–	–
SC8	p40003m	Zuzugsdatum nach Deutschland Befragte/r: Monat
SC8	p40003y	Zuzugsdatum nach Deutschland Befragte/r: Jahr
SC8	p400070_g1D	Geburtsland Mutter Befragte/r (Deutschland/Ausland; bereinigt)
SC8	p400090_g1D	Geburtsland Vater Befragte/r (Deutschland/Ausland; bereinigt)
SC8	p400500_g2	Fehlende/widersprüchliche Geburtslandangaben beim Generationenstatus
SC8	p401100	Deutsche Staatsangehörigkeit Befragte/r
SC8	p401110	Deutsche Staatsangehörigkeit seit Geburt Befragte/r
SC8	p40113m	Erwerb deutscher Staatsangehörigkeit Befragte/r: Monat
SC8	p40113y	Erwerb deutscher Staatsangehörigkeit Befragte/r: Jahr
SC8	p401160	Beantragung deutscher Staatsangehörigkeit Befragte/r
SC8	p403000_g1	Geburtsland Partner/in (Deutschland/Ausland; bereinigt)

(...)

Tabelle 4: (continued)

Kohorte	Variable	Label
SC8	p403070_g1D	Geburtsland Mutter Partner/in (Deutschland/Ausland; bereinigt)
SC8	p403090_g1D	Geburtsland Vater Partner/in (Deutschland/Ausland; bereinigt)
SC8	p404000	Deutsche Staatsangehörigkeit Partner/in
SC8	p404010	Deutsche Staatsangehörigkeit seit Geburt Partner/in
SC8	p40403m	Erwerb deutscher Staatsangehörigkeit Partner/in: Monat
SC8	p40403y	Erwerb deutscher Staatsangehörigkeit Partner/in: Jahr
SC8	p406000_g1	Geburtsland in Deutschland Zielkind (bereinigt)
SC8	p40603m	Zuzugsdatum nach Deutschland Zielkind: Monat
SC8	p40603y	Zuzugsdatum nach Deutschland Zielkind: Jahr
SC8	p407050_g1D	Staatsangehörigkeit Zielkind (deutsch/nicht deutsch)
SC8	p407055	Zweite Staatsangehörigkeit Zielkind: ja/nein
SC8	p407060_g1D	Zweite Staatsangehörigkeit Zielkind (deutsch/nicht deutsch)
SC8	p410000_g1D	Erstsprache/Muttersprache Kind (deutsch/nicht deutsch)
SC8	p410002_g1D	weitere Erstsprache/Muttersprache Kind (deutsch/nicht deutsch)
SC8	p410030	Herkunftssprache ermitteln - bilinguales Kind
SC8	p410040	Deutschsprachigkeit Kind (Autovariable)
SC8	p410050	Bilingualität Kind (Autovariable)
SC8	p428927	[HILF] Geburtsland im Ausland befragte Person
SC8	p428928	[HILF] Geburtsland im Ausland Mutter Partner/in
SC8	p428929	[HILF] Geburtsland im Ausland Vater Partner/in
SC8	p428931	[HILF] Geburtsland Kind im Ausland
SC8	p529000	Gewicht bei Geburt
SC8	p700011	Geschlecht nach Geburtenregister Zielkind
SC8	p70012m	Geburtsdatum Zielkind: Monat
SC8	p70012y	Geburtsdatum Zielkind: Jahr
SC8	p731122	Geschlecht nach Geburtenregister Partner/in
SC8	p731706	Geschlecht nach Geburtenregister Befragte/r
SC8	p731707	Geschlechtsidentität Befragte/r
SC8	p73170y	Geburtsjahr Befragte/r
SC8	p73175y	Geburtsjahr Partner/in
SC8	t400000_g1D	Geburtsland (Deutschland/Ausland)
SC8	t400500_g2	Fehlende/widersprüchliche Geburtslandangaben beim Generationenstatus
SC8	t401152_g1D	Erste Staatsangehörigkeit (deutsch/nicht deutsch)
SC8	t401153_g1D	Zweite Staatsangehörigkeit (deutsch/nicht deutsch)
SC8	t414000_g1D	erste Muttersprache (deutsch/nicht deutsch)
SC8	t414002_g1D	zweite Muttersprache (deutsch/nicht deutsch)
SC8	t414030	Herkunftssprache ermitteln - bilinguales Target
SC8	t414040	Hilfsvariable: Deutschsprachigkeit Target
SC8	t414050	Hilfsvariable: Bilingualität Target
SC8	t428923	Auswahl Geburtsland Elternteil A oder B
SC8	t428924	Auswahl Geburtsland Großeltern
SC8	t700032	Geschlecht bei Geburt
SC8	t700033	Geschlechtsidentität
SC8	t70004m	Geburtsdatum: Monat
SC8	t70004y	Geburtsdatum: Jahr
SC8	t731152_g1D	Geburtsland Elternteil A (Deutschland/Ausland; bereinigt)
SC8	t731153_g1D	Geburtsland Mutter Elternteil A (Deutschland/Ausland; bereinigt)
SC8	t731154_g1D	Geburtsland Vater Elternteil A (Deutschland/Ausland; bereinigt)
SC8	t731162_g1D	Geburtsland Elternteil B (Deutschland/Ausland; bereinigt)

(...)

Tabelle 4: (continued)

Kohorte	Variable	Label
SC8	t731163_g1D	Geburtsland Mutter Elternteil B (Deutschland/Ausland; bereinigt)
SC8	t731164_g1D	Geburtsland Vater Elternteil B (Deutschland/Ausland; bereinigt)
SC8	tx80501	Kinder-/Schülerliste: Geschlecht Zielkind
SC8	tx80506	Kinder-/Schülerliste: Staatsangehörigkeit Zielkind
SC8	tx8050m	Kinder-/Schülerliste: Geburtsdatum (Monat) Zielkind
SC8	tx8050y	Kinder-/Schülerliste: Geburtsdatum (Jahr) Zielkind

A.5 Release Notes

Die Release Notes dokumentieren Veränderungen der vorliegenden Datenversion gegenüber der jeweils vorgängigen Datenversion. Sie enthalten Hinweise auf relevante Anpassungen bzgl. der Datenedition sowie auf bekannten (und ggf. behobene) Probleme bzw. Fehler in den Daten. Die (tages)aktuell gepflegten Release Notes finden sich auf der Webseite als Bestandteil der Datendokumentation (siehe Kapitel 4.1).

```
==================================================
**
** NEPScomp - RELEASE NOTES
** Hinweise und Aktualisierungen zu Version 2.0.0
** (doi:10.5157/NEPScomp:SUF:2.0.0)
**
==================================================

Allgemeine Hinweise:

        - detaillierte Informationen zur Vorbereitung und Handhabung der Daten sind im
          *Datenmanual* auf der Website von NEPScomp zu finden

        - Bitte melden Sie Probleme mit den Daten oder (vermeintliche) Fehler jederzeit
          an das Forschungsdatenzentrum (FDZ-LIfBi: fdz@lifbi.de) -- vielen Dank!

==================================================
** Änderungen in Release Version 2.0.0
** im Vergleich zu vorherigem Release
==================================================

        - Integration der Erhebungsdaten der NEPS-Startkohorte 8 (SUF 2.0.0)
          => neu: Wellen 1 und 2
          => keine Übernahme der Variablen zur prozeduralen Metakognition für
             die Domänen "Digitale Kompetenz", "ICT", "Russisch" und "Türkisch"
             (Anteil korrekt + Differenzmaß) aus dem regulären Scientific-Use-
             File aufgrund notwendiger Korrekturen in den Originaldaten

        - Aktualisierung der Erhebungsdaten der NEPS-Startkohorte 3 (SUF 13.1.0)

        - Ergänzung der Erhebungsdaten der NEPS-Startkohorte 4 (SUF 16.0.0)
          => neu: Welle 16

        - Ergänzung der Erhebungsdaten der NEPS-Startkohorte 6 (SUF 17.0.0)
          => neu: Wellen 16 und 17

        - Mehrfachzuordnung von Variablen pro Startkohorte implementiert
          => neu: bei Zuordnung eines Items zu verschiedenen Oberthemen wird
             die entsprechende Variable identisch in die jeweiligen Datensätze
             übernommen
```