

LEO-Grundbildungsstudie (LEO 2018) Methodenbericht

Forschungsprojekt im Auftrag der

Universität Hamburg
Fakultät für Erziehungswissenschaft
Berufliche Bildung und Lebenslanges Lernen (EW 3)

von

Kantar – Public Division Deutschland
Landsberger Str. 284
80687 München

Ihre Ansprechpartner:

Frauke Bilger
Associate Director Arbeit und berufliche Bildung
Tel: 089 5600 - 1696
E-Mail: frauke.bilger@kantar.com

Alexandra Strauß
Senior Consultant Arbeit und berufliche Bildung
Tel: 089 5600 - 1812
E-Mail: alexandra.strauss@kantar.com

München, den 30. Januar 2019
Pj.-Nr: 315107950

* Der Titel LEO-Grundbildungsstudie war der anfängliche Arbeitstitel der Studie und wurde erst nach der Verfassung des Methodenberichts in den Titel LEO – Leben mit geringer Literalität überführt. Es handelt sich bei diesem Methodenbericht um eine von der Universität Hamburg gekürzte Fassung, da einige Aspekte bereits durch die Dokumente „User-Manual“ und „CAPI-Fragebogen“ aufgenommen wurden. Zudem wurde auf Passagen zu Variablen, die die technische Auswertung der Testhefte betreffen, verzichtet. Der ausführliche Methodenbericht kann auf Nachfrage beim LEO-Team der Universität Hamburg bezogen werden.

Inhaltsverzeichnis

1	Ziele und Herausforderungen	5
2	Konzeption der LEO-Grundbildungsstudie	7
2.1	Studiendesign	8
2.2	Das Fragenprogramm der LEO-Grundbildungsstudie	10
2.2.1	Entwicklung des Hintergrundfragenprogramms	10
2.2.2	LEO 2018-Literalitätstests in der CAPI-Methodik	12
2.2.3	Befragungsunterlagen	16
2.3	Erhöhung der Teilnahmebereitschaft	17
2.3.1	Kontaktaufnahme zu Beginn des Interviews	18
2.3.2	Maßnahmen im Interviewverlauf	18
2.3.3	Incentivierung	19
2.4	Aufgaben der Interviewerinnen und Interviewer	20
2.4.1	Ergebnisse aus dem Pretest zur Interviewerrolle als Testleitung	21
2.4.2	Generelle, standardisierte Interviewerschulung bei Kantar	23
2.4.3	Projektspezifische Interviewerschulung	24
2.5	Anlage der Stichproben	25
2.5.1	Anlage der (Pretest-)Basiserhebung der LEO-Grundbildungsstudie	26
2.5.2	Anlage der LEO 2018 Zusatzerhebung	27
2.5.3	Anlage der Zusatzerhebung im Pretest	28
2.5.4	Integration beider Stichproben der Haupterhebung	28
3	Feldarbeit	29
3.1	Durchschnittliche Interviewdauer	29
3.2	In LEO 2018 realisierte Interviews und Ausfälle	30
3.2.1	Ausschöpfung und Ausfälle der LEO 2018-Basiserhebung	31
3.2.2	Ausschöpfung und Ausfälle der LEO 2018-Zusatzerhebung	32
3.3	Qualitätskontrolle Interviewer	33
3.3.1	Allgemeine Informationen über die Kantar-CAPI-Line sowie den Kantar-Interviewerstab	33

3.3.2	Standardisierte Interviewerkontrolle durch die CAPI-Line	34
3.3.3	Projektspezifische Interviewerkontrolle	35
Exkurs: Teilnahmebereitschaft der LEO-Befragungspersonen für eine Wiederholungsbefragung		38
4	Datenaufbereitung und Auswertung der Testaufgaben	39
4.1	Aufbereitung der im CAPI erhobenen Daten	39
4.1.1	Prüfung auf Filterfehler	39
4.1.2	Vollständigkeit der Interviews und Umgang mit Antwortverweigerungen	40
4.1.3	Datenaufbereitung und Codierung	41
4.2	Bewertung der Testaufgaben für den Datensatz (<i>Editing</i>)	42
4.2.1	Schulung der Rater	43
4.2.2	Bewertung der Testaufgaben	45
4.3	Erneute Bewertung ausgewählter Testhefte aus LEO 2010 durch das Raterteam von LEO 2018	51
5	Testteilnahme und Testverweigerungen	53
5.1	Nichtteilnahme an der Befragung (<i>Unit-Nonresponse</i>)	54
5.1.1	Die Bildungsstruktur der Stichprobe	54
5.1.2	Die Bearbeitung der Alphahefte im Vergleich von LEO 2010 und LEO 2018	56
5.2	Nichtteilnahme an Testaufgaben (<i>Item-Nonresponse</i>)	58
5.3	Unvollständige Testbearbeitung: Gründe und Typen	61
5.3.1	Gründe der unvollständigen Testbearbeitung	62
6	Gewichtung und Hochrechnung	68
6.1	Umwandlung der Basisstichprobe (<i>Designgewichtung I</i>)	69
6.2	Anpassung an Bevölkerungsstrukturen (Kalibrierung)	70
6.3	Gewichtung der intergiierten LEO 2018-Stichprobe	72
6.4	Erstellung der Gewichtungs- und Hochrechnungsfaktoren	74
7	Literatur	75
Anhang: 77		
A1	Befragungsunterlagen	77
A1-1	LEO 2018-Listenheft	77
A1-2	Empfehlungsschreiben der Auftraggebenden	118
A1-3	Ankündigungsschreiben für die Zusatzerhebung	119
A1-4	LEO 2018-Datenschutzblatt	121
A1-5	Extrablatt für die Interviewenden	123

A2	Standartexte zur Vercodung	124
A2-1	Vercodung der Beruf nach ISCO 08	124
A2-2	Vercodung der Wirtschaftszweige nach NACE 2008	127
A2-3	Vercodung der Themen von Weiterbildungsaktivitäten nach ISCED-Fields	130

1 Ziele und Herausforderungen

Mit der LEO-Grundbildungsstudie (im Folgenden: LEO 2018) stellt sich die Universität Hamburg¹ zum zweiten Mal der Herausforderung die Größenordnung des *funktionalen Analphabetismus*² in der 18- bis 64-jährigen Wohnbevölkerung (im Folgenden: Erwachsene) zu bestimmen. Die zentrale Herausforderung der zweiten LEO-Erhebung liegt nicht mehr darin, die Machbarkeit der Studie unter Beweis zu stellen, sondern darin, vergleichbare Ergebnisse zur vorangegangenen Erhebung vorzulegen.³ Basierend auf der standardisierten face-to-face-Befragung unter deutschsprachigen 18- bis 64-Jährigen konnten mithilfe der für die leo. – Level-One-Studie (im Folgenden LEO 2010) speziell entwickelten Lese- und Schreibtests (Grotlüschen, Riekmann & Buddeberg 2012a) gezeigt werden, dass 14 Prozent, hochgerechnet 7,5 Millionen, der Erwachsenen in Deutschland die Textebene unterschreiten. D.h., dass eine Person zwar einzelne Sätze lesen oder schreiben kann, nicht jedoch zusammenhängende – auch kürzere – Texte (Grotlüschen, Riekmann & Buddeberg 2012b, S. 19). Wird sich nun dieses bildungspolitisch zentrale Ergebnis bestätigen oder wird gar mit einem Anstieg oder einem Rückgang zu rechnen sein?

Aufgrund der hohen Bedeutsamkeit der Ergebnisse aus LEO 2010 – die federführende Projektleiterin, Prof. Dr. Anke Grotlüschen, wurde am Weltalphabetisierungstag zur "Botschafterin für Alphabetisierung 2015" benannt – wurde die LEO 2018-Erhebung als *Stand-Alone*-Studie konzipiert. Noch LEO 2010 wurde im Huckepackverfahren im Anschluss an die Erhebung des Adult Education Survey (AES 2010) durchgeführt (Bilger, Jäckle, Rosenblatt & Strauß, 2012). Hintergrundinformationen für LEO 2010 waren entsprechend die Informationen, die im Rahmen des AES 2010 erhoben wurden. Dies gilt für LEO 2018 nicht mehr. Für die LEO-Grundbildungsstudie wurde ein gut halbstündiges Hintergrundfragenprogramm konzipiert, entwickelt, zusammengestellt, in einem Pretest erprobt und unter 7.192 Erwachsenen erhoben. Das den wegen der Vergleichbarkeit weitgehend unveränderten Literalitätstests vorgeschaltete Hintergrundfragenprogramm beinhaltet neben Fragen nach soziodemografischen und berufsbezogenen Informationen vorrangig Fragen nach selbsteingeschätzten Grundkompetenzen und schriftsprachlichen Praktiken in verschiedenen Domänen (Kapitel 2.2).

Das Studiendesign wurde für LEO 2018 aufgrund der Vergleichbarkeit analog zu LEO 2010 konzipiert (Kapitel 2.1). Vorgesehen waren sowohl eine für die erwachsene Wohnbevölkerung repräsentative Basiserhebung als auch eine Zusatzerhebung. Mithilfe der Zusatzerhebung unter Erwachsenen, die höchstens einen Hauptschulabschluss aufweisen, wurde sichergestellt, dass auch in LEO 2018 die insgesamt erhobene Nettostichprobe

¹ Die Studie wird unter Federführung von Prof. Dr. Anke Grotlüschen und Dr. Klaus Buddeberg durchgeführt. Darüber hinaus gehören dem *LEO-Team* in Hamburg in alphabetischer Reihenfolge Gregor Dutz, Lisanne Heilmann, Christopher Stammer und Kirsten Vittali an.

² Das *LEO-Team* (s. Fußnote 1) spricht inzwischen nicht mehr von der *Größenordnung funktionaler Analphabeten*, sondern von der *Größenordnung gering Literalisierter*. Weil die Festlegung der begrifflichen Veränderung parallel zur Berichtslegung erfolgte, erlauben wir uns in diesem projektinternen Bericht die alte begriffliche Fassung beizubehalten.

³ Damit stellt sich die Universität Hamburg auch der im *Grundsatzpapier zur Nationalen Dekade* formulierten Herausforderung, mit LEO 2018 „an die Studie von 2010/11 anzuschließen und eine Fortschreibung der Ergebnisse zu sichern“ (ohne Autor 2017, S. 8).

ausreichend viele Fälle für die getrennte Betrachtung von Literalitätsgraden auf unterem Kompetenzniveau enthält.

Auch in LEO 2018 wurden Hintergrundfragen und Literalitätstests in die Methodik des persönlich-mündlichen, computergestützten Interviews (CAPI: *computer assisted personal interview*) integriert. Das Fragenprogramm für Basis- (n=6.681) und Zusatzerhebung (n=511) war dabei identisch. Die CAPI-Interviews wurden von Interviewerinnen und Interviewern des Kantar-Stabes am Wohnort der Befragungspersonen durchgeführt. Sie stellten sich dabei auch der besonderen Herausforderung, die Rolle der Testleitung zu übernehmen (Kapitel 2.4). Für die Erhebung wurden vor allem aus Gründen der Vergleichbarkeit dieselben Stichprobenanalgen für LEO 2018 geplant, wie zuvor für LEO 2010 (Kapitel 2.5).

Der hiermit von Kantar Public vorgelegte Bericht wurde von Frauke Bilger und Alexandra Strauß erstellt⁴ und beinhaltet methodische Informationen zur Konzeption (Kapitel 2), Feldarbeit (Kapitel 3), Datenaufbereitung (Kapitel 4 und 6) und Testteilnahme bzw. -verweigerung (Kapitel 5).

⁴ Wir danken unseren ehemaligen Praktikantinnen Lorena Göriz und Mascha Nolte für ihre Unterstützung.

2 Konzeption der LEO-Grundbildungsstudie

Die LEO-Grundbildungsstudie (LEO 2018) soll zur Vorgängerstudie (LEO 2010) vergleichbare Ergebnisse vorlegen. Dafür ist ein vergleichbares Studiendesign umzusetzen (Kapitel 2.1). Grundlage von LEO 2018 sind demnach – analog zu LEO 2010 – eine Basiserhebung unter Erwachsenen, also den 18- bis 64-Jährigen, (n=6.681) angelegt als Random Route nach ADM⁵-Standard mit CAPI⁶-Methode (Kapitel 2.5.1), ergänzt um eine Zusatzstichprobe von 511 Erwachsenen im unteren Bildungsbereich (höchster schulischer Abschluss: Hauptschulabschluss). Die Zusatzerhebung erfolgte analog zu LEO 2010 auf Basis von Altadressen (genauer s. Kapitel 2.5.2). Die Gesamtstichprobe umfasst insgesamt 7.192 Interviews (genauer s. Kapitel 3.2).

Aufgrund des erstmaligen *Stand-Alone*-Charakters der Studie war neben der Vergleichbarkeit der Ergebnisse aus den Literalitätstests zudem ein Hintergrundfragenprogramm neu zu konzipieren, zusammenzustellen, zu entwickeln (Kapitel 2.2), für die CAPI-Erhebung anzupassen und zu erheben. Der der Haupterhebung vorgelegte Pretest (Kapitel 2.2) diente demnach der Überprüfung folgender Punkte:

- **Instrumententest:**
Vor allem die Instrumente der neu eingesetzten Fragen im neu konzipierten Hintergrundfragenprogramm waren zu überprüfen (Kapitel 2.2.1). Hierbei wurden nicht nur Auffälligkeiten in den Pretestdaten selbst berücksichtigt, sondern auch die direkt von den Interviewerinnen und Interviewern benannten Kritikpunkte im Rahmen eines Kurzfragenprogramms. Darüber hinaus wurde die durchschnittliche Interviewdauer mithilfe der Pretestdaten geprüft (Kapitel 2.2.2). Eine konkrete Abschätzung der Qualität der Befragungsunterlagen mit Blick auf eine potenzielle Erhöhung der Teilnahmebereitschaft folgte (Kapitel 2.3).
- **Verfahrenstest:**
Der Verfahrenstest diente nicht nur dazu, eine Größenordnung der rücklaufenden Testhefte abzuschätzen (Kapitel 2.2.2), sondern der Erprobung des Interviewverlaufs selbst. Die besondere Herausforderung im Einzelinterview liegt für die Interviewerinnen und Interviewer nicht nur darin, eine potenzielle Befragungsperson von einer Teilnahme am Interview zu überzeugen und danach in standardisiertem CAPI-Verfahren Fragen beantworten zu lassen, sondern vor allem in ihrer Rolle als Testleitung (Kapitel 2.4) während der Bearbeitung der Literalitätstests durch die Befragungspersonen.

⁵ ADM: *Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V.*

⁶ CAPI: *computer assisted personal interviews*

2.1 Studiendesign

Wenngleich die LEO-Grundbildungsstudie (LEO 2018) – anders als zuvor die leo. – Level-One-Studie (LEO 2010) – als *Stand-Alone-Studie* konzipiert und durchgeführt wurde, entspricht sie in ihrer methodischen Konzeption, Anlage und Durchführung der Vorgängerstudie (Bilger, Jäckle, Rosenblatt & Strauß, 2012). Eine schematische Darstellung des Designs von LEO 2018 findet sich in Abbildung 1.

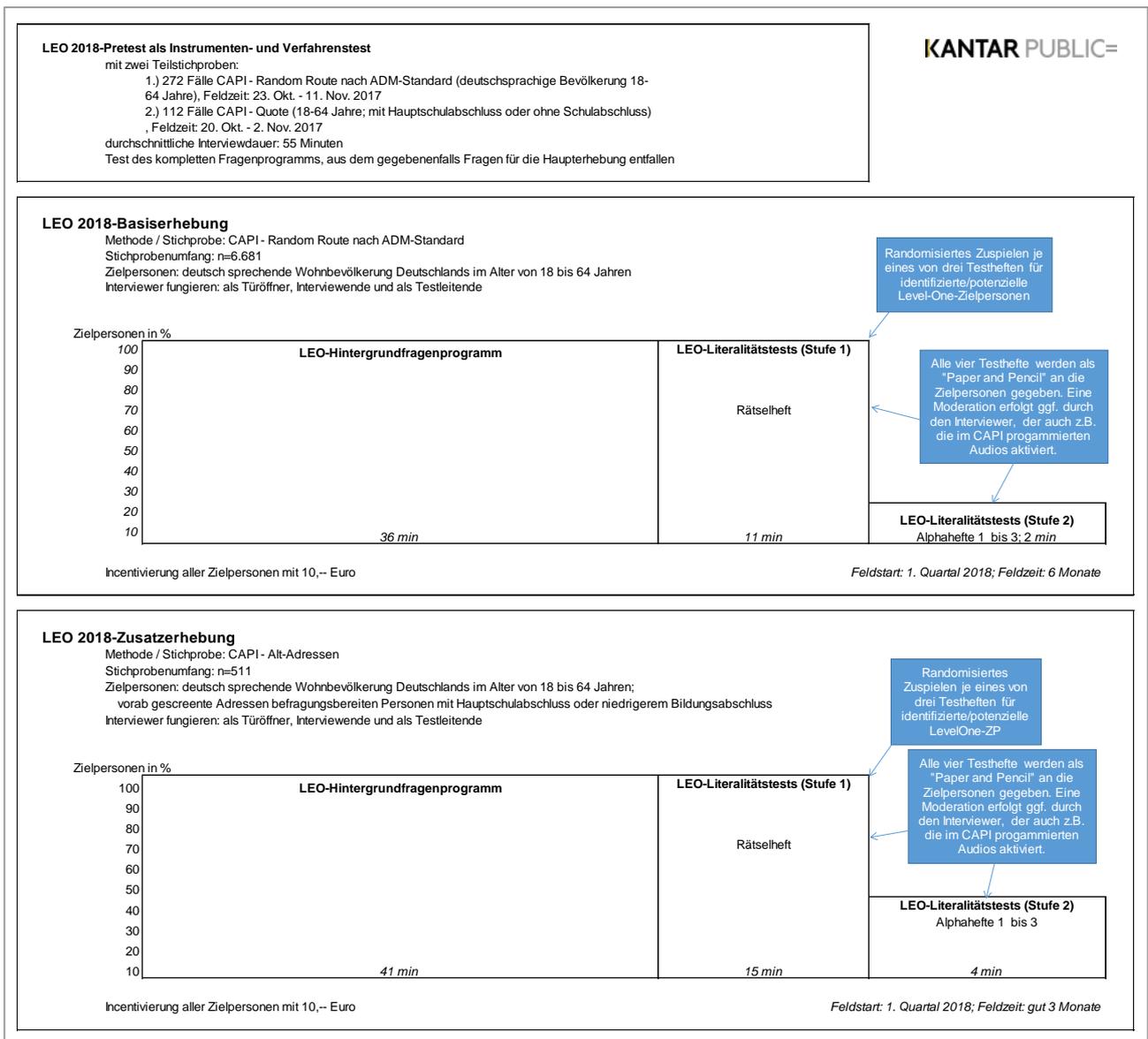


Abbildung 1: Design der LEO-Grundbildungsstudie (LEO 2018)

Die *Basiserhebung* erfolgte als CAPI-Befragung unter 18- bis 64-Jährigen, angelegt als Random Route nach ADM-Standard (Kapitel 2.5.1). Geplant wurden 6.500 bis 7.000 Interviews mit einer durchschnittlichen Interviewdauer von einer knappen Stunde. Im Rahmen der Erhebung sollten zunächst die Hintergrundfragen und

anschließend die Literalitätstests erhoben werden. Dabei wurde der erste Teil der Testaufgaben im sogenannten *Rätselheft* allen Zielpersonen vorgegeben. Erst nach Eingabe der Testergebnisse ins CAPI wurde dort bestimmt, ob die Zielperson ein weiteres Testheft erhält, und wenn ja, zufallsbasiert eines von insgesamt drei möglichen sogenannten *Alphaheften* vorgegeben. Das Fragenprogramm einschließlich der Literalitätstests der Basiserhebung wurde von vornherein so konzipiert, dass es eins zu eins in der Zusatzerhebung eingesetzt werden konnte.

Um auch in LEO 2018 sicherzustellen, dass in der Gesamtstichprobe ausreichend Testergebnisse der Personen mit niedrigem Bildungsniveau vorliegen, wurde analog zu LEO 2010 auch hier eine *Zusatzerhebung* geplant. Die Zusatzerhebung beruht auf einer Bruttostichprobe von Erwachsenen, also 18- bis 64-jährigen Personen, die über höchstens einen Hauptschulabschluss verfügen. Das für die Identifikation dieser Personen notwendige Screening erfolgte, analog zu LEO 2010, aus Kostengründen nicht im Feld, sondern auf Basis bereits vorliegender Informationen (Kapitel 2.5.2). Das Brutto-Sample der Zusatzerhebung basiert auf Adressen von Personen, die bereits an einer anderen Studie teilnahmen und sich dabei mit einer neuerlichen Befragung durch Kantar einverstanden erklärten. Um sicherzustellen, dass nur Personen, die höchstens über einen Hauptschulabschluss verfügen, in die Stichprobe gelangen (Stichwort: Vermeidung von *overcoverage*) wurde in der Zusatzerhebung das Interview dann beendet, wenn die Zielperson einen höheren schulischen Abschluss als „A Hauptschule, Volksschule“ in der entsprechenden Frage (**F030**) benannte. Um im Rahmen der Zusatzerhebung im Interviewverlauf möglichst zeitnah die Befragungsperson als Zielperson der Zusatzerhebung zu validieren, wurde die Frage nach dem Bildungshintergrund vergleichsweise weit vorne im Interviewverlauf verortet.⁷

Die zentralen Herausforderungen in der Studienkonzeption lagen darin, ...

- insbesondere in der für die Literalitätstestung zentralen Gruppe mit Schriftsprachschwierigkeiten das Abbruchrisiko zu minimieren. Hier musste insbesondere auf die Übergangsphasen zwischen dem Hintergrundfragenprogramm und dem Testteil und zwischen den Testteilen 1 (Rätselheft) und 2 (Alphahefte) geachtet werden. Die u.a. dafür durchgeführte Schulung der Kontaktinterviewerinnen und -interviewer (Konzept: *Train-the-trainer*) des Kantar-Stabes wird weiter unten in Kapitel 2.3 beschrieben.
- insbesondere Personen mit geringer Schriftsprachkompetenz, die häufig ein niedriges Bildungsniveau aufweisen, für die Befragung zu gewinnen. Hierfür wurde besonders auf die Ausdrucksform im Interview geachtet sowie auf eine geeignete Präsentation sonstiger Befragungsunterlagen (Kapitel 2.2). Darüber hinaus wurde allen Befragungspersonen ein monetäres Incentive in Höhe von 10,- € angeboten (s. Kapitel 2.3).
- die zeitliche Anlage des Gesamtinterviews möglichst so zu gestalten, dass ein Abbruchrisiko aufgrund zeitlicher und gegebenenfalls kognitiver Überlastung nicht verstärkt wird (Kapitel 2.2).

Auf die Maßnahmen zur Erhöhung der Teilnahmebereitschaft bzw. zur Minimierung des Abbruchrisikos wird gesondert in Kapitel 2.3 Bezug genommen, nach Vorstellung der Befragungsinstrumente in Kapitel 2.2. Besondere Herausforderungen für die Interviewenden werden in Kapitel 2.4 beschrieben und die Stichprobenanlagen in Kapitel 2.5.

⁷ Der Bildungshintergrund der Zielperson wird im Interview als zehnte Frage gestellt, chronologisch nach den Fragen zu Geschlecht, Alter und Haushaltszusammensetzung. Viele Informationen, die der Frage vorgelagert sind, sind Auskünfte (z.B. regionaler Art), die die Interviewerinnen und Interviewer ins CAPI eingeben können, ohne die Zielpersonen zu fragen.

2.2 Das Fragenprogramm der LEO-Grundbildungsstudie

Zur Bestimmung der Größenordnung *funktionaler Analphabeten* liegt der Fokus der LEO-Grundbildungsstudie auf der Erhebung von Schriftsprachkompetenzen. Dazu wurden in LEO 2018 die bereits in LEO 2010 erprobten, standardisierten Tests der Lese- und Schreibfähigkeiten der Zielpersonen eingesetzt. Die LEO-Grundbildungsstudie ist als *Stand-Alone-Studie* angelegt. Wie Abbildung 1 im letzten Abschnitt zeigt, war für die Erhebung ein Hintergrundfragenprogramm neu zusammenzustellen, das durchschnittlich gut eine halbe Stunde Zeit in der CAPI-Befragung umfassen sollte. Für den Pretest war eine rund 10 Minuten längere Version des Hintergrundfragenprogramms vorgesehen. Im Anschluss an die Hintergrundfragen erfolgte – analog zu LEO 2010 – die durch die Interviewerin bzw. den Interviewer entsprechend der CAPI-Vorgaben moderierte Testerhebung. Auf der ersten Stufe erhielten alle Befragungspersonen elf Aufgaben, präsentiert in Form eines selbst auszufüllenden Rätselhefts (*Paper & Pencil*). Die dort von den Befragungspersonen eingetragenen Ergebnisse wurden nach Fertigstellung des Rätselhefts von den Interviewerinnen und Interviewern ins CAPI eingegeben. Dort wurde ein Score berechnet. Sofern dieser eine festgelegte Fehlerschwelle unterschritt, wurde mithilfe des CAPI-Programms zufällig eines von drei Alphaheften („Joschi“, „Helga“ oder „Leschek“) wiederum mit jeweils vier Aufgaben ausgewählt. Das ausgewählte Alphaheft wurde der Befragungsperson mit Bitte um Bearbeitung übergeben.

2.2.1 Entwicklung des Hintergrundfragenprogramms

Die Literalitätstests erfolgten auch in der zweiten Erhebung im Jahr 2018 im Anschluss an eine standardisierte Befragung, in der erstmals der Grundbildungsbegriff erweitert wird. Im Hintergrundfragenprogramm wird eine subjektive Einschätzung von Grundbildungskompetenzen in verschiedenen Segmenten oder Domänen erfragt. Damit erweitert sich die Informationsbreite von LEO 2018 maßgeblich gemäß den Anforderungen des *Grundsatzpapiers der Nationalen Dekade 2016*. Erfragt wurden:

- Gesundheitsbezogene Praktiken und Grundkompetenzen
- Finanzielle Praktiken und Grundkompetenzen
- Familiäre Grundkompetenzen (Mobilität, Familie/Kinder, Alltag)
- Digitale Grundkompetenzen und Praktiken
- Politische Praktiken und Grundkompetenzen
- Arbeitsorientierte Grundbildung

Ergänzt wurden die Fragenblöcke durch Fragen nach personen- und berufs- bzw. arbeitsbezogenen Informationen, Informationen über die Haushaltszusammensetzung, Informationen zum Migrationshintergrund, Informationen zur selbsteingeschätzten Mutter- bzw. Fremdsprachbefähigung sowie zum Weiterbildungsverhalten bzw. zur Bildungsberatung.

Die Instrumente für das Hintergrundfragenprogramm wurden bis Ende August 2017 von der Universität Hamburg eigens entwickelt oder aus anderen Studien zusammengestellt und für LEO 2018 adaptiert. Eine erste Beschreibung der auszuwählenden Instrumente erfolgte bereits auf der ersten LEO-Beiratssitzung⁸ im Mai

⁸ In alphabetischer Reihenfolge gehören dem Beirat von LEO 2018 folgende Personen an:
Prof. Dr. Helmut Bremer – Universität Duisburg-Essen, Prof. Dr. Ingrid Gogolin – Universität Hamburg, Prof. Dr. Johannes Hartig – DIPF Frankfurt, Prof. Dr. Klaus Hurrelmann – Hertie School of Governance Berlin, Prof. Dr. Carola Iller – Universität Hildesheim, Prof. Dr. Corinna Kleinert – IIfBI, Universität Bamberg, Prof. Dr. Nele McElvany – Technische Universität Dortmund, Prof. Dr. Beatrice Rammstedt – GESIS Mannheim, Prof. Dr. Josef Schrader – DIE Bonn, Prof. Dr. Heike Solga – WZB Berlin.

2017. Im September 2017 brachte Kantar Public die von der Universität Hamburg ausgewählten Instrumente in eine für eine Breitenbefragung geeignete Reihenfolge und beriet bei der Modifikation der Instrumente für die LEO-Grundbildungsstudie im CAPI-Erhebungsmodus. Eine Zusammenschau der Fragenblöcke der LEO 2018-Erhebung ist in Abbildung 2 gezeigt.

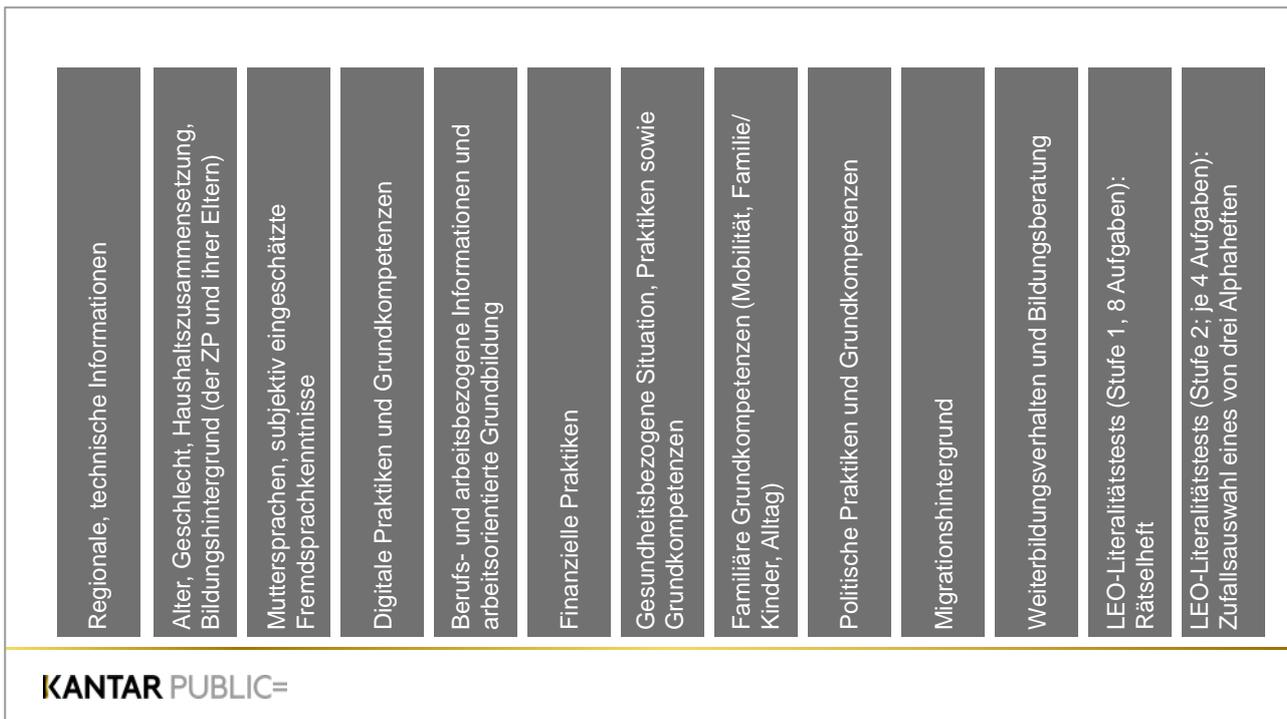


Abbildung 2: Zusammengefasste Fragenblöcke der LEO-Grundbildungsstudie

Mithilfe des Pretests (Oktober/November 2017) konnten die Instrumente des Hintergrundfragenprogramms überprüft werden. Dafür erstellte Kantar Public einen Chart-Bericht, der am 6. Dezember 2017 auf einem gemeinsamen Workshop mit dem LEO-Team an der Universität Hamburg vorgestellt und diskutiert wurde. Die Diskussionsergebnisse und Entscheidungen für Änderungen verschiedener Punkte des Pretestdesigns und des Pretestfragenprogramms für die Haupterhebung wurden nachträglich im projektinternen Chartbericht als Protokollergebnis dokumentiert (Bilger, Strauß, Gensicke, Baier, & Backes, 2017).

Der Pretest zeigte, als gemeinsames Fazit der Universität Hamburg und Kantar Public, dass die gewählten Abläufe und Erhebungsinstrumente – auch des Hintergrundfragenprogramms – insgesamt sehr weit entwickelt waren. Auf Detailebene lagen Auffälligkeiten vor, für die Anpassungen für die Haupterhebung vorgenommen werden sollten. Der Kürzungsbedarf wurde auf Basis von 272 Fällen unter 18- bis 64-Jährigen in der CAPI-Befragungsmethode angelegt als Random Route nach ADM-Standard sowie 112 Fällen unter 18- bis 64-Jährigen in einer CAPI-Befragung mit Quotenstichprobe⁹, insgesamt also 384 Fällen auch rechnerisch ermittelt. Berechnet wurde eine durchschnittliche Interviewdauer des Hintergrundfragenprogramms im Pretest von 40 Minuten, getrennt nach den beiden Stichproben dauerten die Pretestinterviews der Zusatzerhebung etwas

⁹ Um die *gescreenten* Altadressen für die Haupterhebung nicht nutzen und verschlanken zu müssen, wurde für den Pretest ein einfacheres Stichprobenverfahren gewählt. Inhaltlich wurde im Pretest ermittelt, ob sich eine Testerhebung unter Personen mit niedrigerem Bildungsniveau entsprechend des Verfahrens von LEO 2010 nach wie vor realisieren ließe. Dafür reichte die Stichprobenanlage als *Quota Sampling* aus (Kapitel 2.5.3).

länger (s.u. Tabelle 1, Kapitel 2.2.2). Ein weiteres Ergebnis (s. nächster Abschnitt) zeigt zwar, dass die durchschnittliche Bearbeitungszeit für je Rätsel- und Alphahefte etwas kürzer als angenommen war, allerdings in der für die LEO-Grundbildungsstudie zentralen Zielgruppe dennoch vergleichsweise hoch liegt. Vor diesem Hintergrund und vor dem Hintergrund der Hinweise von knapp einem Drittel der im Pretest eingesetzten Interviewerinnen und Interviewer, dass das Interview (erheblich) zu lang sei, wurde einvernehmlich unsere Empfehlung, zur Minimierung von Interviewbelastung und somit des Abbruchrisikos, das Interview dennoch zu kürzen angenommen.

Neben der Entnahme einiger Instrumente wurden in einigen Fällen Antwortvorgaben verschlankt, um die durchschnittliche Interviewdauer für die Haupterhebung zu reduzieren. Wie bereits erwähnt, zeigten sowohl die Hinweise der Interviewerinnen und Interviewer als auch die Pretestergebnisse keinen einschlägigen Änderungsbedarf des Fragenprogramms. Gleichwohl wurde eine Vielzahl von Verbesserungsmöglichkeiten im Detail ermittelt und für die Haupterhebung angepasst:

- Einzelne sprachliche Vereinfachungen bzw. sprachliche Ausschärfung des Sachverhalts
- Vereinzelte Änderung der Fragenabfolgen im Detail
- Vereinheitlichung von Antwortvorgaben z.B. in Fragen mit Skalen an verschiedenen Stellen im Fragenprogramm
- Leichte Anpassung von Filterführungen

Detaillierte Informationen zum Hintergrund von Verbesserungsvorschlägen sind, wie bereits erwähnt, im projektinternen Chartbericht zum Pretest dokumentiert (Bilger, Strauß, Gensicke, Baier, & Backes, 2017).

2.2.2 LEO 2018-Literalitätstests in der CAPI-Methodik

Ziel der Studie ist es, die Größenordnung *funktionaler Analphabeten* in Form einer Quote unter den Erwachsenen zu bestimmen. Zu diesem Zweck wurden analog zu LEO 2010 standardisierte Literalitätstests der Zielpersonen durchgeführt. Dabei wurden wenige bereits in LEO 2010 eingesetzte Testaufgaben leicht geändert und eine neue Aufgabe ergänzt. Weil der überwiegende Teil der Testaufgaben in LEO 2018 übernommen wurde, ist davon auszugehen, dass ein direkter Vergleich mit den Ergebnissen aus LEO 2010 vorgenommen werden kann.

Die Lese- und Schreibfähigkeit der Befragungspersonen wurde analog zur Erhebung im Jahr 2010 am Ende der Befragung, also im Anschluss an das Hintergrundfragenprogramm getestet. Ebenfalls analog zu LEO 2010 wurden im Rahmen des LEO 2018-CAPI-Fragenprogramms standardisierte Fragen und Erläuterungen eingesetzt, mit denen die Interviewerinnen und Interviewer die Tests in der Interviewsituation präsentieren und deren Bearbeitung dokumentieren. Die Tests wurden den Befragungspersonen wiederum in Form farbig gedruckter Rätsel- und Alphahefte im Rahmen der Erhebung vorgelegt (Abbildung 3).¹⁰ Der CAPI-Teil der Erhebung mit Hinweisen für die Interviewerinnen und Interviewer zur Übergabe der Testhefte und Aufgaben in der Testsituation mussten so aufeinander abgestimmt sein, dass die Interaktion von Interviewerin bzw. Interviewer und Befragungsperson den Testanforderungen entsprechend gelang. Zudem sollten die Interviewerinnen und Interviewer nach jeder (Teil-)Aufgabe angeben, ob sie vollständig bearbeitet wurde. Sofern dies nicht der Fall war, wurden sie gebeten, die Gründe für die unvollständige Bearbeitung offen anzugeben und im CAPI einzutragen, ob die Zielperson den Test weiterbearbeitet oder an der gegebenen Stelle abbricht (s. Kapitel 5).

¹⁰ Aufgrund der Geheimhaltungsvereinbarung mit der Universität Hamburg werden hier beispielhaft ausschließlich die Deckblätter des Rätselhefts sowie der drei Alphahefte gezeigt, nicht aber die vollständigen Testhefte dokumentiert.

Insgesamt wurden den Befragungspersonen elf Aufgaben (inkl. der einen neu eingesetzten Aufgabe) im Rahmen des sogenannten Rätselheftes zur Bearbeitung vorgelegt. Ziel der in den Rätselheften angeführten sogenannten Filteraufgaben war es, Personen mit vergleichsweise großen Schwierigkeiten beim Lesen und Schreiben zu identifizieren. Im Anschluss an die Bearbeitung des Rätselheftes waren die Interviewerinnen und Interviewer aufgefordert, die schriftlich in den Heften niedergelegten Antworten der Zielperson auszuwerten und die Ergebnisse in das CAPI-Instrument zu übertragen (s. auch Kapitel 2.4.3). Auf Basis der so erhobenen Testergebnisse konnten maximal elf Punkte erreicht werden.¹¹ Der genaue Punktwert oder Score wurde im CAPI auf Basis der eingegebenen Ergebnisse errechnet und mit einer vorab durch die Universität Hamburg definierten Fehlerschwelle verglichen. Sofern eine Person neun oder mehr Punkte erreichte, endete das Interview. Bei acht oder weniger Punkten – also einer Unterschreitung der Fehlerschwelle – wurde im CAPI-Programm eine randomisierte Auswahl von eines von drei Alphaheften für eine weitere Testung der Zielperson vorgenommen. Knapp ein Viertel aller in LEO 2018 befragten Zielpersonen (24 Prozent) erhielt auf diese Weise ein zufällig ausgewähltes Alphaheft zur Bearbeitung.

Vor Testbeginn waren die Interviewerinnen und Interviewer angehalten sicherzustellen, dass sich nach Möglichkeit keine weitere Person im Raum aufhält bzw. wenn doch eine etwaige Unterstützung durch diese Person bei der Bearbeitung der Testaufgaben ausgeschlossen wird. Zudem wurde darauf hingewiesen, dass eine entsprechende Unterstützung durch die Interviewerinnen und Interviewer selbst ebenfalls unzulässig ist¹² (s. auch Kapitel 2.4.1). Hierdurch wurde so weit wie möglich sichergestellt, dass die Tests ausschließlich von der Befragungsperson selbst bearbeitet werden. Das Vorgehen ist in Abbildung 3 dargestellt.

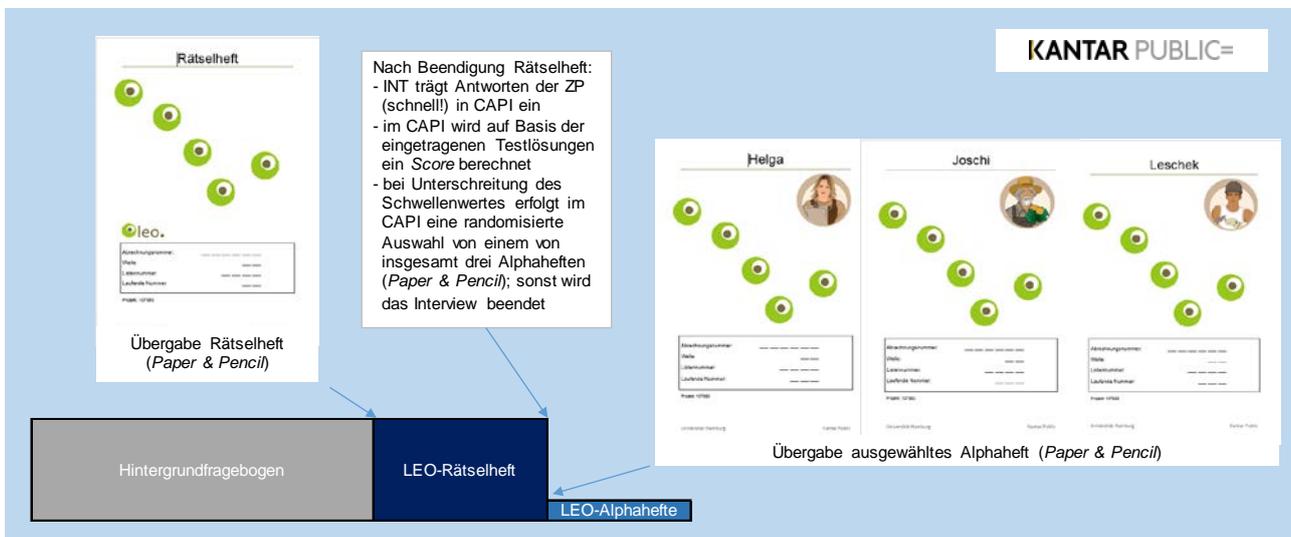


Abbildung 3: Zweistufiger Literalitätstest

¹¹ Dieser Wert liegt aufgrund der neu eingefügten Aufgabe um einen Punkt höher als in LEO 2010. Dort konnten maximal 10 Punkte im Rätselheft erreicht werden.

¹² S. z.B. Interviewerhinweis im CAPI-Programm (T0a): „Für den folgenden Teil des Interviews soll nach Möglichkeit keine dritte Person im Raum sein. Falls dies nicht zu vermeiden ist, stellen Sie sicher, dass diese dritte Person der ZP bei der Beantwortung der Fragen NICHT hilft. ...Es ist wichtig, dass Sie keine weiteren Erläuterungen geben und der ZP bei der Lösung der Aufgaben NICHT helfen!“

Jedes Alphaheft enthielt vier Aufgaben. Zufällig erfolgte im Pretest durch die randomisierte Auswahl eine Gleichverteilung der Alphahefte. Gerundet auf ganze Prozentpunkte erhielten jeweils acht Prozent der Befragungspersonen eines der drei Alphahefte „Joschi“, „Helga“ oder „Leschek“.

In LEO 2010 lag eine besondere Herausforderung darin, die Kompetenztests an die Umfeldbedingungen und Abläufe einer CAPI-Erhebung anzupassen. Weil dieser Schritt bereits in LEO 2010 erfolgreich umgesetzt wurde, galt es in LEO 2018 geringe Anpassungen aufgrund der wenigen leicht veränderten Aufgaben bzw. der einen neuen Aufgabe vorzunehmen und ansonsten das für LEO 2010 erarbeitete Prozedere in den Pretest zu übernehmen. Dort wurde geprüft, ob die (damals) getroffenen Entscheidungen nach wie vor zu vergleichsweise geringen Abbruchquoten im Rahmen der Literalitätstests unter den Befragungspersonen führen und ob es weitere Möglichkeiten gibt, die Qualität insbesondere des Testteils der Erhebung zu halten bzw. verbessern.

Sehr detailliert wurde im Pretest der Frage nach der Testteilnahmebereitschaft bzw. einem -abbruch der Befragungspersonen (Abbrecherquote 4 Prozent im Pretest; zum Vergleich 3 Prozent in LEO 2010) nachgegangen. Signifikante Unterschiede, die eine Änderung der Testaufgaben sowohl in Rätsel- als auch Alphaheften mit sich gebracht hätten, waren nicht zu erkennen (genauer s. Bilger, Strauß, Gensicke, Baier, & Backes, 2017).

Im Nachgang des Pretests wurde die durchschnittliche Interviewdauer berechnet. Tabelle 1 zeigt die Ergebnisse für den Pretest insgesamt und getrennt nach Stichprobe und dem Bildungsniveau der Zielpersonen. Insgesamt dauerte ein Pretestinterview durchschnittlich 55 Minuten. Dieser Wert lag so hoch, wie erwartet. Er setzt sich aus den erwarteten Durchschnittswerten für den Hintergrundfragebogen (durchschnittlich 40 Minuten)¹³, die Bearbeitung des Rätselhefts (durchschnittlich 14 Minuten) und die Bearbeitung eines Alphahefts (durchschnittlich 2 Minuten) zusammen. Im Angebot zur Durchführung wurde aufgrund der Annahme für eine größere Erweiterung des Rätselhefts von einer durchschnittlichen Bearbeitungsdauer von 20 Minuten ausgegangen. Aufgrund der nachgelagerten Diskussionen wurde dieser Punkt inhaltlich durch die sogenannte Verlinkungsstudie, deren Feldarbeiten Anfang November 2018 abgeschlossen wurden und für die es einen eigenen Methodenbericht geben wird abgelöst. Vor diesem Hintergrund konnten die Annahmen auf einen Erwartungswert etwa analog zu LEO 2010 bezüglich der Interviewdauer korrigiert werden. Damals wurde eine durchschnittliche Bearbeitungszeit für das Rätselheft von einer Viertelstunde ermittelt. Somit fallen die durchschnittlichen Bearbeitungszeiten für das Rätsel- und Alphaheft im Pretest erwartungsgemäß aus.

¹³ Bereits im Angebot wurde die durchschnittliche Interviewdauer für das LEO-Hintergrundfragenprogramm mit durchschnittlich gut 30 Minuten und für den Pretest mit durchschnittlich 40 Minuten vorgesehen.

Tabelle 1: Durchschnittliche Pretest-Interviewdauer gerundet auf Minuten insgesamt und getrennt nach Stichproben und Bildungshintergrund

Durchschnittliche Interviewdauer in Minuten	alle Pretest-interviews (n=384)	Pretest-interviews der Basisstichprobe (n=272)	Pretest-interviews der Zusatzstichprobe (n=112)	ZP ¹⁾ : niedrige Bildung ²⁾ (n=183)	ZP ¹⁾ : mittlere bis hohe Bildung ²⁾ (n=197)
Hintergrundfragen	40	39	43	40	39
Rätselheft	14	13	17	15	13
Alphaheft	2	2	4	4	2
Gesamt	55	53	63	55	51
Neue Basis: Personen mit Rätsel- und Alphaheft:	(n=92)	(n=57)	(n=35)	(n=63)	(n=29)
Joschi	11	8	14	13	9
Helga	7	6	9	8	7
Leschek	9	8	11	10	8
Dauer Alphahefte	9	8	11	10	8
Gesamt	62	57	70	61	57

¹⁾ ZP: Zielperson

²⁾ Niedrige Bildung: Die Zielperson verfügt höchstens über einen Hauptschulabschluss.
Mittlere bis hohe Bildung: Die Zielperson verfügt über einen höheren Abschluss als einen Hauptschulabschluss.

Wie Tabelle 1 zeigt, sind die zeitlichen Interviewbelastungen vor allem in der Zielgruppe am höchsten, die ein Alphaheft zur Bearbeitung erhielt. Innerhalb dieser Gruppe zeigen sich zudem leichte Unterschiede in der durchschnittlichen Bearbeitungszeit der drei verschiedenen Alphahefte „Joschi“, „Helga“ und „Leschek“. Die durchschnittlich geplante Gesamtinterviewdauer liegt hier bei einer guten Stunde. In Einzelfällen liegen die Interviewzeiten deutlich darüber. Insbesondere für diese für LEO zentrale Gruppe, ist das Risiko, das Interview abubrechen, möglichst gering zu halten. Entsprechend wurde, wie in Kapitel 2.2.1 ausgeführt, der Empfehlung gefolgt, das Hintergrundfragenprogramm zeitlich etwas zu kürzen.

Nachgelagert wurde auf Basis der Pretestergebnisse ein einfaches Editing aller im Pretest bearbeiteten Testhefte vorgenommen und für die Universität Hamburg in Form von Variablen im Pretestdatensatz abgelegt. Herangezogen wurde die sehr umfassende, schriftliche Anleitung aus der Vorgängerstudie LEO 2010. Durch das einfache Editing konnten auch die eingegebenen Testergebnisse der Interviewenden im Pretest überprüft werden.

Beim konkreten Editing der Pretest-Testhefte zeigten sich – erwartungsgemäß – einige Schwierigkeiten. Sie ergaben sich aus der Vielzahl von Einzelfällen, die im Rahmen des Editings zu klären waren (*Ermessensspielraum*). Es war praktisch nicht möglich, für alle Fälle vorab genaue Regeln für die Bearbeitung im Editing der Anleitung zu entnehmen bzw. diese während der Bearbeitung festzulegen. Vor diesem Hintergrund ist eine enge Abstimmung zwischen Editingteam und Projektleitung besonders wichtig, um ein einheitliches Editierverhalten sicherzustellen. Bereits in der Studienkonzeption wurde daher für die Haupterhebung für zehn Prozent der Testhefte ein doppeltes Editing geplant. Nach Vorstellung und Diskussion dieses Pretestergebnisses in der Beiratssitzung im Januar 2018 wurde dieser Ansatz auf 20 Prozent erhöht. Das genaue Vorgehen in der Haupterhebung wird in Kapitel 4 beschrieben.

Im Rahmen des Editings der Pretest-Testhefte fiel zudem ein teils uneinheitliches Vorgehen einzelner Interviewerinnen und Interviewer bei der Bewertung der Ergebnisse in den bearbeiteten Rätselheften in der konkreten Interviewsituation auf. Um eine Verbesserung der Auswertung der Testaufgaben zu erhalten, wurden

im Nachgang des Pretests einerseits die Hinweise für die Interviewerinnen und Interviewer im CAPI angepasst, die Editierarbeiten etwas aufgestockt (s.o.) und andererseits die bereits vorgesehene Schulung für die Interviewerinnen und Interviewer im Format *Train-the-trainer* (Kapitel 2.3.2) inhaltlich planbarer und in ihrer Bedeutung untermauert.

2.2.3 Befragungsunterlagen

Neben dem CAPI-Fragenprogramm wurden zusätzliche Befragungsunterlagen erstellt. Das Rätselheft sowie die drei Alphahefte („Joschi“, „Helga“ oder „Leschek“) werden aufgrund der Geheimhaltungsvorschrift durch die Universität Hamburg nicht in diesem Bericht dokumentiert. Darüber hinaus wurden ein Empfehlungsschreiben der Auftraggebenden, ein Datenschutzblatt und ein Listenheft für die Zielpersonen sowie eine allgemeine Intervieweranleitung und ein sogenanntes Extrablatt für die Interviewerinnen und Interviewer erstellt. Für die Zusatzerhebung, für die personenbezogene Kontaktinformationen vorlagen, wurde das Interview zudem mit einem Anschreiben angekündigt.

Noch im Rahmen der LEO 2010-Studie wurde möglichst auf schriftlastige Befragungsunterlagen für die Zielpersonen verzichtet. In der LEO 2010-Zusatzerhebung wurde auf ein Listenheft verzichtet. Vor diesem Hintergrund wurde auch im Pretest für LEO 2018 ein eher sparsames Listenheft eingesetzt, aber nicht darauf verzichtet. Die Begründung liegt vor allem darin, dass Listen in einem zeitlich eher langen Interviewverlauf für den Durchschnitt der Bevölkerung das Interview beschleunigen. Aus Rückmeldungen der Interviewerinnen und Interviewer ging hervor, dass eine stärkere Visualisierung der Antwortvorgaben (vor allem der Skalen) im Rahmen eines Listenhefts für die Befragung mit Blick auf die Interviewdauer und das leichtere Verständnis hilfreich sei. Die Universität Hamburg unterstützte diesen Vorschlag, weil sie den *cognitive load* in der konkreten Interviewsituation – auch im Falle von Schriftsprachbeeinträchtigungen – dadurch entlastet bzw. nicht mehr belastet sehe. Vor diesem Hintergrund wurde das Listenheft etwas erweitert und umfasst für die Haupterhebung 39 Listen (Anhang A1-1).

Die Interviewerinnen und Interviewer verwiesen im Rahmen einer dem jeweiligen Pretestinterview nachgelagerten Kurzbefragung darauf, dass sich eine Teilnahmebereitschaft der Zielpersonen leichter erwirken ließe, wenn ein Empfehlungsschreiben der Auftraggebenden vorliege. Hierdurch könne die Seriosität des Vorhabens gegenüber den Zielpersonen verdeutlicht werden (*vertrauensbildende Maßnahme*). Vor diesem Hintergrund erstellten die Universität Hamburg und das Bundesministerium für Bildung und Forschung (BMBF) ein mit Blick auf die im Fokus stehende Zielpersonengruppe kurz gehaltenes Empfehlungsschreiben für die Haupterhebung (Anhang A1-2).

Das im Rahmen der Zusatzerhebung versandte Ankündigungsschreiben ist in Anhang A1-3 und das den Zielpersonen übergebene Datenschutzblatt in Anhang A1-4 dokumentiert. Darüber hinaus hielten die Interviewerinnen und Interviewer am Ende der Befragung einen Quittungsblock zur Quittierung der Incentives bereit.

Um sicherzustellen, dass die Literalitätstestung der Zielpersonen weder an einer nicht bereitliegenden Sehhilfe scheitert, wurden die Interviewerinnen und Interviewer nicht nur in den Fragentexten und Hinweisen darauf aufmerksam gemacht, sicherzustellen, dass etwaige Sehhilfen genutzt werden sollen, sondern auch im sogenannten Extrablatt für die Interviewerinnen und Interviewer (Anhang A1-5). Dort wird zudem auf den zur Bearbeitung der Testhefte mit einem „speziellen Stift“ verwiesen, den die Interviewerinnen und Interviewer zur Durchführung der Tests bereitstellten. Hierdurch wurde einerseits sichergestellt, dass eine Bearbeitung der *Paper & Pencil*-Testhefte möglich und andererseits, dass eine einheitliche Lesbarkeit (Farbe und Strichstärke des Geschriebenen) der Antworten für die Editoren zum Vergleich der Schriftbilder gegeben ist.

2.3 Erhöhung der Teilnahmebereitschaft

Es ist davon auszugehen, dass es bei *motivierten* Befragten zu weniger Antwortausfällen (*Item-Nonresponse*) kommt und dass sich diese mehr bemühen, Fragen bzw. Testaufgaben korrekt zu beantworten oder zu bearbeiten. Weil die Teilnahme der Zielpersonen an LEO 2018 freiwillig ist, werden motivationssteigernde Maßnahmen zur Erhöhung der Teilnahmebereitschaft eingesetzt.

Eine Besonderheit der LEO-Grundbildungsstudie liegt darin, dass Personen mit geringen Lese- und Schreibfähigkeiten möglicherweise Vermeidungsstrategien entwickeln, die zu einer hohen Wahrscheinlichkeit von Teilnahmeverweigerungen (*Unit-Nonresponse*) führen können (vgl. hierzu auch Kapitel 5). Demnach könnte eine geringe Teilnahme von Personen mit Problemen beim Lesen und Schreiben die Ergebnisse der Studie verändern (Kapitel 5.2). Ebenso ist aus der Umfrageforschung das Phänomen bekannt, dass Personen mit niedrigen Bildungsabschlüssen eine geringere Bereitschaft zur Umfragebeteiligung aufweisen (*Bildungsbias*, s. Kapitel 5.2). Um die Teilnahmebereitschaft generell zu erhöhen und dadurch das Abbruchrisiko während der Befragung zu verringern, wurden die im Folgenden beschriebenen Schritte unternommen. Eine schematische Zusammenfassung zeigt vorab Abbildung 4.

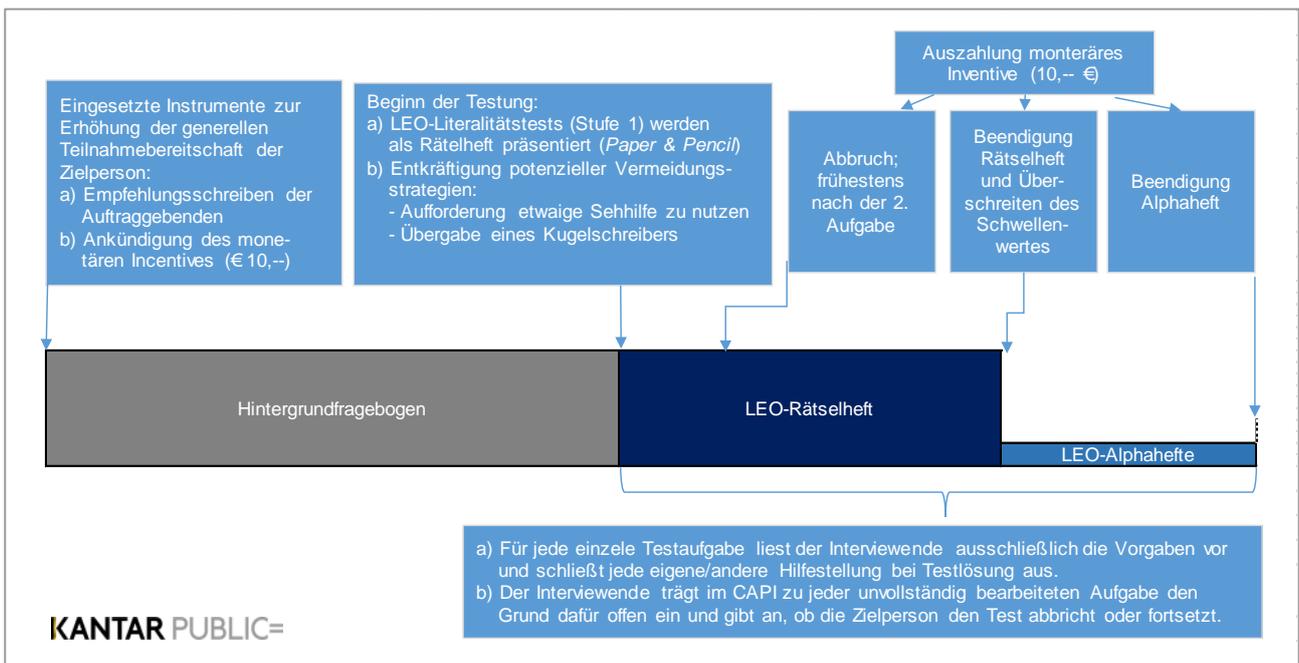


Abbildung 4: Maßnahmen zur Erhöhung der potentiellen Teilnahmebereitschaft und zur Minimierung eines potentiellen Abbruchrisikos im LEO 2018-Interviewverlauf

2.3.1 Kontaktaufnahme zu Beginn des Interviews

Die Interviewerinnen und Interviewer des Kantar-Stabes sind erfahrene *Kontaktprofis* und wurden für LEO 2018 speziell mit Argumenten versorgt, wie die Befragungspersonen zur Teilnahme gewonnen und auch *weiche* Verweigerer von der Wichtigkeit der Untersuchung überzeugt werden können. Als vertrauensbildende Maßnahme im Erstkontakt wurde den Interviewenden ein Empfehlungsschreiben der Universität Hamburg und des BMBF an die Hand gegeben. Im Falle der Zusatzerhebung, für die personenbezogene Kontaktinformationen vorlagen, wurde zudem ein Ankündigungsschreiben im Vorfeld der Befragung versendet (s. Kapitel 2.2.3).

In der Phase der Kontaktaufnahme wurde sowohl im Ankündigungsschreiben im Rahmen der Zusatzerhebung im Empfehlungsschreiben der Auftraggebenden, in den Hinweisen für die Interviewerinnen und Interviewer (*Extrablatt*) als auch in der Ansprache der Zielpersonen durch die Interviewerinnen und Interviewer selbst darauf geachtet, nur allgemein gehaltene Informationen über Art und Ziele der Untersuchung zu geben (Abbildung 5). Es wurde an dieser Stelle des Ablaufs nicht angekündigt, dass am Ende des Interviews einige Testaufgaben zu bearbeiten sind. Ziel dieses Vorgehens war es, die Teilnahmeentscheidung weitgehend themen- bzw. literalitätsneutral und unabhängig von einer eventuellen Test-Aversion zu gestalten.

<p>Auszug aus dem Empfehlungsschreiben der Universität Hamburg und des BMBF:</p> <p>Es geht dabei um Sprache in unterschiedlichen Situationen des Alltags. Hierzu bitten wir Sie um Ihre Mitwirkung.</p> <p>Auszug aus dem Ankündigungsschreiben für die Zusatzerhebung:</p> <p>Die Studie zielt darauf ab, bundesweit repräsentativ zu ermitteln, wie Erwachsene im Alter von 18 bis 64 Jahren mit Sprache in unterschiedlichen Situationen des Alltags umgehen.</p> <p>CAPI-Einleitung des Interviews</p> <table border="1"> <thead> <tr> <th>EINL</th> <th>vorlesen</th> <th>ZP (Zielpersonen): alle</th> </tr> </thead> <tbody> <tr> <td colspan="3"> <p>Kantar Public führt eine repräsentative Befragung für eine wissenschaftliche Untersuchung durch. Auftraggeber sind die Universität Hamburg und das Bundesministerium für Bildung und Forschung. Wir bedanken uns für Ihre Bereitschaft, bei dieser Studie mitzumachen. Für Ihre Teilnahme erhalten Sie 10 Euro als besonderes Dankeschön. Ihre Angaben im Interview werden streng vertraulich behandelt und ausschließlich für die Untersuchung verwendet.</p> <p>***INT: Wenn Befragter fragt, worum es in der Studie geht: Es geht um Sprache in unterschiedlichen Situationen des Alltags. Zunächst möchten wir Sie um einige Angaben zu Ihrer Person und Ihrer Lebenssituation bitten.</p> </td> </tr> </tbody> </table>	EINL	vorlesen	ZP (Zielpersonen): alle	<p>Kantar Public führt eine repräsentative Befragung für eine wissenschaftliche Untersuchung durch. Auftraggeber sind die Universität Hamburg und das Bundesministerium für Bildung und Forschung. Wir bedanken uns für Ihre Bereitschaft, bei dieser Studie mitzumachen. Für Ihre Teilnahme erhalten Sie 10 Euro als besonderes Dankeschön. Ihre Angaben im Interview werden streng vertraulich behandelt und ausschließlich für die Untersuchung verwendet.</p> <p>***INT: Wenn Befragter fragt, worum es in der Studie geht: Es geht um Sprache in unterschiedlichen Situationen des Alltags. Zunächst möchten wir Sie um einige Angaben zu Ihrer Person und Ihrer Lebenssituation bitten.</p>			<p>Auszug aus der Intervieweranleitung:</p> <p>ZIEL DER BEFRAGUNG: Die Studie „Sprache erkennen“ zielt darauf ab, bundesweit repräsentativ zu ermitteln, wie gut Erwachsene im Alter von 18 bis 64 Jahren Sprache erkennen und damit umgehen. Weil manche Personen vielleicht abwehrend reagieren, halten Sie sich bitte in der Erklärung gegenüber den Befragungspersonen an den Eingangstext im Interview. Dort heißt es:</p> <p><i>Kantar Public führt eine repräsentative Befragung für eine wissenschaftliche Untersuchung durch. Auftraggeber sind die Universität Hamburg und das Bundesministerium für Bildung und Forschung. Wir bedanken uns für Ihre Bereitschaft, bei dieser Studie mitzumachen. Für Ihre Teilnahme erhalten Sie 10 Euro als besonderes Dankeschön. Ihre Angaben im Interview werden streng vertraulich behandelt und ausschließlich für die Untersuchung verwendet.</i></p> <p>→ Wenn Befragter fragt, worum es in der Studie geht: Es geht um Sprache in unterschiedlichen Situationen des Alltags.</p>
EINL	vorlesen	ZP (Zielpersonen): alle					
<p>Kantar Public führt eine repräsentative Befragung für eine wissenschaftliche Untersuchung durch. Auftraggeber sind die Universität Hamburg und das Bundesministerium für Bildung und Forschung. Wir bedanken uns für Ihre Bereitschaft, bei dieser Studie mitzumachen. Für Ihre Teilnahme erhalten Sie 10 Euro als besonderes Dankeschön. Ihre Angaben im Interview werden streng vertraulich behandelt und ausschließlich für die Untersuchung verwendet.</p> <p>***INT: Wenn Befragter fragt, worum es in der Studie geht: Es geht um Sprache in unterschiedlichen Situationen des Alltags. Zunächst möchten wir Sie um einige Angaben zu Ihrer Person und Ihrer Lebenssituation bitten.</p>							

KANTAR PUBLIC=
9

Abbildung 5: Literalitätsneutralität in der Kontaktaufnahme

Bereits in der Phase der Kontaktaufnahme wurde allen Zielpersonen ein monetäres Incentive in Höhe von 10,- Euro als Dank für ihre Teilnahme angekündigt (Kapitel 2.3.3).

2.3.2 Maßnahmen im Interviewverlauf

Die ausgewählte Befragungsmethode CAPI (*computer assisted personal interview*) ist für die LEO-Grundbildungsstudie vor allem aus zwei Gründen vorteilhaft.

- Sie erlaubt komplexe Filterführungen (Kapitel 4.1) sowie die Einbindung computergestützter Techniken (z.B. zur randomisierten Auswahl eines von drei Alphaheften oder zur standardisierten Vortragsweise als Audio).
- Sie bindet eine Interviewerin bzw. einen Interviewer ein, die bzw. der die Fragen vorliest und die Antworten eingibt. Für Personen mit geringer Schriftsprachbefähigung bedeutet dies eine Erleichterung in der Interviewsituation. Eine Teilnahmebarriere aufgrund einer geringen Schriftsprachbefähigung ist damit nicht gegeben. Vor diesem Hintergrund ist die Befragungsmethode als *literalitätsneutral* zu bewerten. Dies gilt natürlich für die Literalitätstests nicht, in denen es gerade darum geht das Lese- und Schreibverständnis einer Zielperson zu testen.

Die im Interview präsentierten Instrumente wurden in möglichst einfacher Sprache gehalten (s. Kapitel 2.2.1). Auf Fremdworte – sofern nicht als im Deutschen gängig bewertet – wurde verzichtet.

Mit dem Übergang in die Testsituation wurden den Zielpersonen die ersten elf LEO-Literalitätstests mit weniger schulischem/beruflichem und mehr freizeitleichem Bezug in Form eines *Rätselhefts* präsentiert. Hierdurch sollten prüfungsbedingte Stressoren und Ängste reduziert werden. Dies ist aus unserer Sicht vertretbar, weil die Zielperson keine für sie direkt relevante Bewertung erfährt – weder in einem bildungs- noch berufsbezogenen Zusammenhang.

Um Testvermeidungsstrategien wie der Vorgabe fehlenden Schreibmaterials oder einer Sehhilfe vorzubeugen, wurden die Interviewerinnen und Interviewer angehalten, einen bestimmten Kugelschreiber bereitzustellen und vor Testbeginn die Zielpersonen darum zu bitten, benötigte Sehhilfen zu nutzen. Beide Punkte wurden den Interviewerinnen und Interviewern sowohl in der Schulung (Kapitel 2.4.1), dem Extrablatt, der Intervieweranleitung¹⁴ als auch direkt im Fragenprogramm als Interviewerhinweis (Kapitel 2.2.3) nahegebracht.

2.3.3 Incentivierung

Gemäß unserem Angebot und den Ausführungen auf dem zweiten LEO-Beiratstreffen wurde der Einsatz von Sachgeschenken in LEO 2018 im Vergleich zu LEO 2010 verändert. In LEO 2010 wurde allen Befragungspersonen der Basis- und Zusatzerhebung ein Parker-Kugelschreiber zur Bearbeitung der Testhefte geschenkt. Mithilfe dieses Incentives sollte nicht nur die Teilnahmebereitschaft an dem Literalitätstest erhöht, sondern auch sichergestellt werden, dass Schreibmaterial vorliegt. Darüber hinaus wurde für alle Befragungspersonen der Zusatzerhebung ein monetäres Incentive in Höhe von 10,-- Euro bereitgestellt. Dieses Incentive kam in der Basiserhebung nur für etwa 50 Prozent der Befragungspersonen zum Einsatz, und zwar für all diejenigen, die ein Alphaheft erhielten (etwa 20 Prozent) und darüber hinaus für rund 30 Prozent der Stichprobe zur freien Gestaltung durch die Interviewerinnen und Interviewer.

Zugrunde lag ein Methodenexperiment, das zeigte, dass insbesondere die sogenannten *bildungsfernen* Personen positiv auf monetäre Anreize reagieren (Erfahrungen z.B. aus dem *Panel Arbeitsmarkt und Soziale Sicherheit*, PASS). Daher wurden auch für LEO 2018 monetäre Incentives für die Zielpersonen vorgesehen.

¹⁴ Dort ist formuliert: „Bitte achten Sie vorab darauf, dass die Zielperson etwaige Sehhilfen nutzt und der Kugelschreiber zur Bearbeitung bereitliegt.“

Aktuellere Erfahrungen z.B. aus dem Sozio-oekonomischen Panel (SOEP) verweisen darauf, dass die Übergabe eines Parker-Kugelschreibers keinen erhöhten Teilnahmeanreiz (mehr) mit sich bringt. Vor diesem Hintergrund wurde für LEO 2018 auf ein solches Incentive verzichtet.¹⁵

Aktuellere Erfahrungen aus dem Feld verweisen darauf, dass eine Incentivevergabe nach *Bauchgefühl der Interviewenden* – analog zu LEO 2010 – heute unseriös wirkt. Die Größenordnung des monetären Incentives in Höhe von 10,- Euro scheint nach wie vor angemessen. Ein niedrigerer Betrag trägt erfahrungsgemäß nicht wesentlich zu einer niedrigeren Teilnahmebereitschaft und ein höherer Betrag nicht zu einer höheren Bereitschaft bei. Um zudem einen *Interviewereffekt* möglichst auszuschließen, wurde in LEO 2018 ein Incentive in Höhe von 10,- Euro für *alle* teilnehmenden Zielpersonen vorgesehen. Das Incentive wurde zu Beginn des Interviews angekündigt und nach Beendigung der Testaufgaben ausgezahlt (s.o., Abbildung 4). Um ein *Fair-Play* auch für Personen mit Schriftsprachschwierigkeiten sicherzustellen, gelten bezüglich der Incentivevergabe die Testaufgaben als bearbeitet, wenn mindestens die zweite Aufgabe im Rätselheft (gedanklich) begonnen wurde.

Der Effekt einer proportional zur Bevölkerung abgebildeten Bildungsstruktur in der ungewichteten Basiserhebung, der in LEO 2010 vor allem auf die *Incentivevergabe per Bauchgefühl* zurückgeführt wurde, ist im Rahmen von LEO 2018 entsprechend nicht zu erwarten (genauer s. Kapitel 6.3). Es ist eher von einer überdurchschnittlichen Beteiligung von Personen mit höherem Bildungsniveau auszugehen. Erste Hinweise, die diese Vermutung bestätigen, zeigten sich bereits im Pretest und wurden gemeinsam mit der Universität Hamburg im Rahmen des Workshops diskutiert. Bezüglich der Strukturen ist eine (leichte) Bildungsverzerrung unproblematisch und kann mithilfe einer Strukturgewichtung zur Darstellung der (hochgerechneten) Ergebnisse ausgeglichen werden (s. Kapitel 6). In Bezug auf die Aussagekraft der LEO-Literalitätstests mag sich ein Effekt zeigen, weil die gewichteten Ergebnisse auf eine geringere absolute Personenzahl zurückzuführen sind, die ein Alphaheft bearbeitet haben (Kapitel 5.1.1).

2.4 Aufgaben der Interviewerinnen und Interviewer

Neben den üblichen Intervieweraufgaben, wie z.B. der Kontaktaufnahme mit dem Haushalt, der Zufallsauswahl der zu befragenden Person im Haushalt und der Interviewführung unter Verzicht jeglicher Einflussnahme auf die Zielperson, kam den Interviewerinnen und Interviewern in der LEO-Grundbildungsstudie – wie zuvor bereits in LEO 2010 – die Rolle der Testleitung zu. Im Gegensatz zur sonst üblichen CAPI-Interviewsituation, in der die Interviewerinnen und Interviewer die Antworten nach jeder Frage direkt in das CAPI-Instrument eintragen, sollten sie die Zielpersonen mit Beginn des Literalitätstestblocks zunächst auf Basis der Anweisungen im Erhebungsinstrument durch die Testaufgaben führen. Neben der gebotenen Objektivität war es hier wichtig, die Motivation der Befragten bis zur letzten Testaufgabe aufrechtzuerhalten – und dies sowohl bei Personen, denen die Aufgaben Schwierigkeiten bereiten, als auch bei solchen, die sich durch die Tests unterfordert fühlen. Im Anschluss an die Bearbeitung des Rätselheftes waren die Interviewerinnen und Interviewer aufgefordert, die schriftlich in den Heften niedergelegten Antworten der Zielperson auszuwerten und in das CAPI-Instrument zu übertragen. Im Falle der Bearbeitung eines der drei Alphahefte sollte zudem die jeweils

¹⁵ Davon abgesehen sichergestellt, dass die Interviewerinnen und Interviewer mit einer ausreichenden Menge gut schreibender Kugelschreiber ausgestattet wurden. Dadurch wurde einerseits die Vollständigkeit des Schreibmaterials sichergestellt und andererseits dafür Sorge getragen, dass die Bearbeitung der Tests mit einem farblich lesbaren Stift und diesbezüglich einheitlich erfolgt.

erste Aufgabe am Ende der Befragung erfasst werden. Letzteres diente vor allem im Rahmen des nachgelagerten Editings der einfacheren Zuordnung von Papiertestheften zu realisierten Interviews im Datensatz.

Die geschilderten Aufgaben erfordern von den Interviewerinnen und Interviewern eine Reihe wichtiger Fähigkeiten. Zunächst mussten sie entscheiden, ob die Tests korrekt bearbeitet wurden. Diese Aufgabe wurde dadurch erschwert, dass die Interviewerinnen und Interviewer die Tests nicht in Ruhe auswerten konnten, sondern die Bewertung im Beisein der Zielpersonen vornehmen mussten. Sie waren damit nicht nur dem Druck der Bewertung unter Beobachtung der Befragungspersonen, sondern zudem einem zeitlichen Druck ausgesetzt, um das Interview zügig fortsetzen zu können. Dabei sollten sie dennoch die Rätseltestergebnisse genau und fehlerfrei bewerten und ins CAPI eingeben. Zusätzlich mussten die Interviewerinnen und Interviewer darauf achten, dass es gerade bei Personen mit größeren Problemen beim Lesen und Schreiben durch die in der persönlichen Interviewsituation möglicherweise peinliche Prüfung ihrer Lese- und Schreibkompetenzen auf Basis der Eintragungen im Rätselheft nicht zu einer Verweigerungshaltung bezüglich der Beantwortung des zweiten Heftes – also eines der drei Alphahefte – kam.

Für den Pretest der LEO-Grundbildungsstudie wurden die Interviewerinnen und Interviewer nicht gesondert geschult. Ihnen lagen die Interviewerhinweise im CAPI, die Intervieweranleitung sowie das sogenannte *Extrablatt* als Unterstützung ihrer Rolle als Testleitung vor. Analog zu LEO 2010 verließ man sich für den Pretest darauf, dass mit dem Studiendesign die Interviewsituation so gestaltet war, dass ...

- die schriftlichen Unterlagen zur Intervieweranleitung die Anforderungen der Studie genau beschreiben und für schwierige Passagen (z.B. der der Testleitung) entsprechende Lösungswege aufzeigen und
- der Fragebogen und die Fragen so gestaltet sind, dass die Aufgabe, diese zu beantworten, leicht fällt, und gleichzeitig klar und eindeutig ist, welche Themen die Fragen ansprechen und welche Bedeutung darin verwendete Begriffe haben.

In Verbindung mit der Auswahl der Interviewerinnen und Interviewer und den Standardschulungen bei Kantar sowie der Möglichkeit der Interviewerinnen und Interviewer im Zweifelsfall bei erfahrenen Kontaktinterviewerinnen und -interviewern oder der Studienleitung nachzufragen, konnte im Vorfeld der Befragung davon ausgegangen werden, dass die Interviewerinnen und Interviewer auch zur Durchführung von Kompetenztests in der Lage seien. Wenngleich LEO 2010 den Beweis erbrachte (Bilger, Jäckle, von Rosenblatt & Strauß, 2012), dass genau dies möglich ist, fielen im Pretest von LEO 2018 einzelne Interviewerinnen und Interviewer mit Verbesserungspotenzial auf (Kapitel 2.4.1). Eine einfache Schulung im Format *Train-the-trainer* war von vorn herein im Studiendesign der Universität Hamburg für LEO 2018 vorgesehen und wurde vom LEO-Beirat im Januar 2018 befürwortet. Eine Beschreibung der Schulung der Interviewerinnen und Interviewer erfolgt in Kapitel 2.4.2.

2.4.1 Ergebnisse aus dem Pretest zur Interviewerrolle als Testleitung

Die Interviewerinnen und Interviewer wurden im Rahmen des Pretests gebeten, einen kurzen Fragebogen zur Machbarkeit von LEO bzw. Schwierigkeiten mit Fragenblöcken oder Einzelfragen zu beantworten. Ein Teil der Fragen diente der Erkundung der Schwierigkeiten der Interviewerinnen und Interviewer mit ihrer Rolle als Testleitung. Wenngleich gut neun von zehn der im Rahmen des LEO-Pretests eingesetzten Interviewerinnen und Interviewer Schwierigkeiten verneinten, fiel einem Zehntel die Rolle nach eigenen Angaben eher oder

sehr schwer. Zudem hielten 83 Prozent eine Schulung für die LEO-Grundbildungsstudie für sinnvoll.¹⁶ Im Rahmen der Schulung wünschten sich die Interviewerinnen und Interviewer vor allem eine genauere Erläuterung der Hintergründe der Studie, Hilfestellungen zu Problemen während des Interviews wie z.B. nachlassender Motivation der Zielperson und Umgang mit einer Lese- oder Schreibschwäche einer Zielperson. Eine genaue Besprechung des Vorgehens mit Rätselheft, Übergangssituation und Alphaheften wäre aus Sicht der Interviewerinnen und Interviewer hilfreich. Dieses Ergebnis unterstrich die Bedeutung der vor der Haupterhebung zeitlich verorteten Interviewerschulung und gab erste Anhaltspunkte zur Priorisierung der Schulungsinhalte.

Im Rahmen des Pretests der LEO-Grundbildungsstudie wurden insgesamt 376 Rätselhefte und weitere 86 Alphahefte nach der Erhebung bei Kantar Public editiert. Dieser Projektschritt diente ...

- als Qualitätssicherungsmaßnahme der Angaben der Interviewerinnen und Interviewer,
- im Falle der Rätselheftergebnisse zur Korrektur der im Datensatz abgelegten Bewertungen (Kapitel 4) sowie
- als Umsetzung ins Datenformat im Falle der Alphahefte (Kapitel 4).

Tabelle 2: Ergebnisse der Rätselhefte: Bewertungen der Interviewenden und des nachgelagerten Editings im Vergleich

Aufgaben im Rätselheft		Anteilswerte in %		
		Interviewende im Vergleich zu Editing:		
		gleich bewertet	schlechter bewertet	besser bewertet
Basis: alle Pretestinterviews (n=384)				
T1b	Sehtest 2	95	2	3
T21		88	0	12
T22		83	2	16
T23	Fernweh	84	3	13
T24		85	0	15
T25		73	0	27
T26		93	0	7
T5a	Rahmenwörter 1	96	3	2
T5b	Rahmenwörter 2	94	1	5
T6a	Richtig schreiben 1	65	1	34
T6b	Richtig schreiben 2	51	1	49
T6c	Richtig schreiben 3	81	0	19
Durchschnittswert Gesamt		82	1	17

Bedeutsam ist an dieser Stelle der Vergleich zwischen den Interviewer-eingegebenen und Rater-editierten Rätselheftergebnissen. Aufgrund der oben beschriebenen Drucksituation der Interviewerinnen und Interviewer bei Eingabe der Ergebnisse ist ein gewisser Fehler zu erwarten. In 82 Prozent der Bewertungen der Aufgaben aus dem Rätselheft kamen die Interviewerinnen und Interviewer im Vergleich zum Editingteam zu demselben Ergebnis (Tabelle 2). In einem Prozent der Fälle wurden die Ergebnisse schlechter und in 17 Prozent besser bewertet. Weil eine Besser-Bewertung in der konkreten Interviewsituation dazu führt, dass der für die Bearbeitung des Rätselhefts ermittelte Score steigt und damit die Wahrscheinlichkeit zur Auswahl eines Alphahefts sinkt, ist dieser Fehler von vergleichsweise hoher Tragweite, weil er nachträglich nicht korrigiert werden kann. Ein Prozent der Aufgaben wurde dagegen schlechter bewertet.

¹⁶ Es ist anzunehmen, dass ein solches Ergebnis für nahezu jede Studie im Falle einer solchen Nachfrage erzielt wird.

Als Hinweis darauf, dass und an welchen Stellen Schulungsbedarf seitens der Interviewerinnen und Interviewer für zumindest das Rätselheft gegeben ist, konnten die Ergebnisse der Bewertung der einzelnen Aufgaben genutzt werden. Während die Bewertung der Aufgabe **T1b** durch die Interviewerinnen und Interviewer zu 95 Prozent mit derjenigen der Editoren übereinstimmt, liegt dieser Wert z.B. für Aufgabe **T6b** nur bei 50 Prozent.

Im Rahmen der projektspezifischen Interviewerschulung sollten vor dem Hintergrund dieser Ergebnisse nicht nur Einzelaufgaben der Testhefte genauer besprochen werden. Zudem sollten die Interviewerinnen und Interviewer dazu angehalten werden, bei der Bewertung der Rätselhefte im Zweifelsfalle von einem falschen und nicht von einem richtigen Ergebnis auszugehen.

2.4.2 Generelle, standardisierte Interviewerschulung bei Kantar

Mit den grundlegenden Standards und den besonderen Fertigkeiten persönlich-mündlicher Befragungen (Auftritt, Regeln des erfolgreichen Kontaktens, *refusal avoidance training* bzw. Stichprobenverfahren wie etwa *Address Random* etc.) werden unsere Interviewerinnen und Interviewer zu Beginn ihrer Tätigkeit bei Kantar umfassend vertraut gemacht. Die Schulung erfolgte sowohl schriftlich, per E-Learning und im Rahmen der sogenannten F2F-Akademie¹⁷ vor Ort als auch in Form eines praktischen Trainings zusammen mit einer bzw. einem sogenannten Kontaktinterviewerin bzw. -interviewer, die bzw. der für die Ausbildung und Begleitung in einer bestimmten Region zuständig ist.

Bei den Kontaktinterviewerinnen und -interviewern handelt es sich um Personen mit sehr großer Erfahrung und ausgeprägter Kontakt- und Führungsstärke. Jede Interviewerin und jedem Interviewer steht so auch außerhalb der Geschäftszeiten immer eine Ansprechperson vor Ort zur Verfügung. Die Kontaktinterviewerinnen und -interviewer sind bei der Rekrutierung und Einarbeitung der Interviewerinnen und Interviewer vor Ort ebenso von zentraler Bedeutung wie bei der Schulung. Die Kontaktpersonen halten in fest vorgegebenen Mindestabständen *Projektdurchsprachen* mit der regionalen Einsatzleitung der Münchner Feldorganisation. Sie nehmen in regelmäßigen Abständen an zentralen Veranstaltungen der Feldorganisation oder an projektspezifischen Schulungen teil und dienen damit als Multiplikatoren für die Streuung von wichtigen Informationen und Kenntnissen an die Interviewerinnen und Interviewer.

Das praktische Einüben der Kontaktierungstechniken im Rahmen der konkreten Interviewertätigkeit vor Ort ist für den Schulungserfolg entscheidend. Die Interviewerinnen und Interviewer lernen dadurch, einen eigenen Zugang zu entwickeln, der zu ihrer Persönlichkeit passt und der dann im Rahmen der Tätigkeit weiter verfeinert werden kann. Im Anschluss daran erfolgt die Übernahme in den Kantar-Interviewerstab. Die Kontaktinterviewerinnen und -interviewer stehen der jeweiligen Interviewerin bzw. dem jeweiligen Interviewer insbesondere zu Beginn ihrer bzw. seiner Tätigkeit auch weiterhin persönlich zur Verfügung. Je nach individuellem Bedarf können dann gegebenenfalls weitere gezielte Nachschulungen durchgeführt werden.

Die Interviewereinweisung in F2F-Projekten kann bei Kantar auf dieser Grundlage standardmäßig anhand von schriftlichen Projektunterlagen erfolgen, die von der Interviewerin bzw. dem Interviewer dann zusammen mit dem Fragebogen in Form eines Testinterviews vor dem eigentlichen Start der Feldarbeit durchgearbeitet werden.

¹⁷ F2F: Face-to-face; mündlich persönliches Interview

2.4.3 Projektspezifische Interviewerschulung

Gesonderte projektspezifische Schulungen erfolgen in der Regel dann, wenn z.B. ein besonders komplexes Erhebungsinstrument oder aber besondere Anforderungen an die Interviewerinnen und Interviewer gegeben sind. Eine Schulung hat stets mehrere Ziele. Neben der Vermittlung inhaltlicher oder technischer Kompetenzen zur praktischen Interviewdurchführung erzeugt eine projektspezifische Schulung stets auch Aufmerksamkeit und – ein entsprechendes Schulungsdesign vorausgesetzt – zusätzliche Motivation bei den Interviewerinnen und Interviewern.

Für LEO 2018 wurde aus Gründen der Qualitätssicherung eine Schulung im Format *Train-the-trainer* geplant.¹⁸ Bei diesem Konzept werden in einem ersten Schritt die 28 regional ansässigen und speziell ausgebildeten Kontaktinterviewerinnen und -interviewer in einer persönlichen Schulung zentral geschult. Diese wiederum schulen mit dem bereitgestellten Schulungsmaterial dezentral in kleineren Gruppen persönlich die in ihrer Region eingesetzten Interviewerinnen und Interviewer.

Die überschaubaren Gruppengrößen sowohl bei den zentralen als auch bei den dezentralen Schulungen ermöglichen es, neben den zu vermittelnden Theorieteilen (z.B. Vorgehen als Testleitung) auch Praxiselemente einzubauen und auf Fragen aus der Gruppe einzugehen. Wichtig ist es, neben einer ausgewogenen Mischung aus Theorie und Praxis auch ein Forum für den Austausch der Kontaktinterviewer zu schaffen und die Möglichkeit zur Rückmeldung an das Befragungsinstitut bzw. den Auftraggeber der Studie zu geben.

Die persönlich geschulten Kontaktinterviewerinnen und -interviewer stehen – neben der Projektleitung – während des gesamten Feldverlaufs als Ansprechpersonen für einen eventuellen Nachschulungsbedarf zur Verfügung und sind in der Lage, neu eingesetzte Interviewerinnen und Interviewer zeitnah in die speziellen Projektmodalitäten einzuweisen.

Die halbtägige LEO-Schulung der Kontaktinterviewerinnen und -interviewer erfolgte am 29. Januar 2018 in Eching bei München. Im Vorfeld der Veranstaltung erhielten die Teilnehmenden folgende Unterlagen:

- CAPI-Fragenprogramm der LEO-Grundbildungsstudie
- Rätselheft
- die Alphahefte „Joschi“, „Helga“ und „Leschek“
- Intervieweranleitung
- Listenheft
- Datenschutzerklärung
- Extrablatt
- Empfehlungsschreiben der Universität Hamburg und des BMBF
- Ankündigungsschreiben für die Zusatzerhebung

Die Schulung der Kontaktinterviewer und -interviewerinnen erfolgte durch die Universität Hamburg, namentlich Dr. Klaus Buddeberg und Kirsten Vittali, im Beisein der Projektleitung von Kantar Public und der CAPI-Feldorganisation von Kantar. Zentral waren zwei Tagesordnungspunkte, vorgetragen von der Universität Hamburg: a) Begrüßung und Vorstellung der Studie und b) Aufbau des Fragenprogramms und Besonderheiten in der Befragungssituation. Zentral wurden demnach folgende inhaltliche Punkte behandelt:

¹⁸ Diese vermittelnde Form der Interviewerschulung wurde nicht zuletzt aus ökonomischen Erwägungen gewählt.

- Informationen über den Projekthintergrund und die Ziele der Untersuchung
- Besonderheiten der Studie und der Zielgruppe
- Thematische Ausrichtung der Studie
- Überblick über das Fragenprogramm
- Überblick über die Literalitätstests getrennt nach Rätsel- und Alphaheften
- Vorgehen in der konkreten Interviewsituation in der Rolle einer Testleitung, vorgestellt in Form eines Rollenspiels
- Vorgehen bei der Bewertung der bearbeiteten Testaufgaben aus dem Rätselheft zur CAPI-Eingabe. Zentrale Veränderung gegenüber LEO 2010 war hier, dass die Interviewerinnen und Interviewer aufgefordert wurden, eine Testaufgabe bei der Bewertung im Zweifelsfalle als „nicht richtig gelöst“ einzustufen. Hierdurch sollte die Vollständigkeit von Informationen sichergestellt werden (s.o. Kapitel 2.4), wenngleich im Einzelfalle eine Zielperson, die an sich den Schwellenwert erreicht hätte, dennoch ein Alphaheft bearbeiten musste.

Die Schulung wurde aufgezeichnet. Im Nachgang der Schulung erstellte die Universität Hamburg ein Schulungsvideo¹⁹, das die Kontaktinterviewerinnen und -interviewer für die Schulung der Interviewenden vor Ort neben dem oben genannten Befragungsmaterial nutzten.²⁰ Die Kontaktinterviewerinnen und -interviewer meldeten im Nachgang der je eigenen Schulung die Namen der geschulerten Interviewerinnen und Interviewer. Erst mit einer solchen Schulungsmeldung durften die Interviewerinnen und Interviewer im Rahmen der Haupterhebung der LEO-Grundbildungsstudie von der CAPI-Einsatzleitung für den konkreten Einsatz berücksichtigt werden.

2.5 Anlage der Stichproben

LEO basiert auf zwei Teilstichproben – der repräsentativen Basiserhebung und der Zusatzerhebung unter Erwachsenen mit niedrigem schulischem Bildungsniveau. Die daraus resultierenden Daten können (gewichtet, s. Kapitel 6) gemeinsam und getrennt voneinander genutzt werden.

Die Vergleichbarkeit der LEO-Grundbildungsstudie (LEO 2018) mit der Vorgängerstudie leo. – Level-One (LEO 2010) ist hoch bedeutsam. Vor diesem Hintergrund wurde, nach längeren Diskussionen, eine unveränderte Form von Stichprobenanlagen für LEO 2018 vorgegeben. Die Basiserhebung erfolgte analog zum LEO 2010 als *Random Route nach ADM-Standard*²¹ unter 18- bis 64-Jährigen (Kapitel 2.5.1) und analog ihr kleineres Pendant im Pretest. Für die Zusatzerhebung wurde – analog zur LEO 2010-Zusatzerhebung – eine Stichprobe aus per Kantar-CAPI-Bus zufällig rekrutierten, grundsätzlich befragungsbereiten Personen im Alter von 18 bis 64 Jahren mit höchstens einem Hauptschulabschluss gezogen (Kapitel 2.5.2). Um ausreichend Adressen für diese Stichprobe in der Haupterhebung bereitstellen zu können, wurde dieser Ansatz – ebenfalls analog zu LEO 2010 – im Pretest nicht nachgebildet. Stattdessen wurde im Pretest eine einfache Quotenstichprobe gezogen (Kapitel 2.5.3), um für die Analysen der Pretestdaten eine ausreichende Anzahl von Alphaheften bereitstellen zu können. Nach der (Pretest-)Erhebung wurden je beide Stichproben in einem gemeinsamen Datensatz integriert (Kapitel 2.5.4).

¹⁹ Wiederum aufgrund der Geheimhaltungserklärung gegenüber der Universität Hamburg ist das Schulungsvideo dem Methodenbericht nicht beigelegt.

²⁰ Die Universität Hamburg nahm an einer Schulung durch einen Kontaktinterviewer im Raum Hamburg teil.

²¹ ADM: *Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V.*

2.5.1 Anlage der (Pretest-)Basiserhebung der LEO-Grundbildungsstudie

Die Basiserhebung sowohl im Rahmen der Pretest- als auch der Haupterhebung umfasst in ihrer Grundgesamtheit die in Privathaushalten lebende Wohnbevölkerung der Bundesrepublik Deutschland zwischen 18 und 64 Jahren mit für die Befragung ausreichenden Deutschsprachkenntnissen. Dies sind nach Zahlen der amtlichen Bevölkerungsfortschreibung (Stand: Dezember 2016) 51,6 Millionen Personen. Die Auswahl der Befragungspersonen erfolgte mit Hilfe einer mehrfach geschichteten, dreistufigen Zufallsstichprobe nach ADM-Standard. Der Stichprobenumfang der Erhebung lag bei netto 6.681 Fällen (Pretest: n=272).

Erste Auswahlstufe: Regionale Schichtung der Stichprobe

Auswahlgrundlage der ersten Stufe waren die bundesweit 53.000 zur Verfügung stehenden *Samplepoints*.²² Durch Schichtung nach regionalen Merkmalen (Bundesländer, Regierungsbezirke, BIK-Gemeindetypen, Kreise, Gemeinden und gegebenenfalls Stadtbezirke) wurden Zellen gebildet, die in ihrer Struktur der Verteilung der Privathaushalte entsprechen. Innerhalb der durch die Schichtungsmerkmale definierten Zellen wurde dann jeweils eine Zufallsauswahl von 1.300 Samplepoints mit Wahrscheinlichkeiten proportional zur Zahl der dort lebenden Haushalte gezogen (*Probability proportional to size* [PPS]).

Zweite Auswahlstufe: Auswahl der Haushalte

Innerhalb der Samplepoints (SP) wurden die Befragungshaushalte zufällig ausgewählt. Ausgehend von einer zufälligen Startadresse, die selbst nicht Befragungshaushalt ist, wird dafür auf einem zufälligen Weg (*Random Route*) jeder dritte Haushalt bestimmt. Der zufällig ausgewählte Startpunkt legt aufgrund der vorgegebenen Regeln für den *Random Walk* die Folgehaushalte eindeutig fest. In jedem ausgewählten Haushalt wurde zunächst geprüft, ob darin eine Person der Grundgesamtheit lebt. Wenn ja, gehörte dieser Haushalt zur Stichprobe. Wenn nein, galt er als neutraler Ausfall und wurde am Ende der Random Route durch einen, nach demselben Verfahren ausgewählten, anderen Haushalt ersetzt. Entsprechend wurde auch mit anderen Ausfällen verfahren.

Dritte Auswahlstufe: Auswahl der Zielpersonen

Innerhalb der ausgewählten Haushalte musste die Interviewerin bzw. der Interviewer ermitteln, wie viele Personen der Zielgruppe im Haushalt leben. Leben mehrere potenzielle Zielpersonen im Haushalt, erfolgte die Auswahl der zu befragenden Person mit Hilfe eines systematischen Auswahlchlüssels (hier: *Schweden-schlüssel*, Abbildung 6), der für alle zum Haushalt gehörenden Personen der Grundgesamtheit die gleiche Chance gewährleistet, in die Stichprobe zu gelangen, und die Zielperson eindeutig festlegt.

²² Sample Points sind Gebietsflächen, aufgeteilt anhand der kommunalen, statistischen Bezirke und unter Zuhilfenahme eines geografischen Informationssystems [GIS] zur Straßeneinteilung der Bundesrepublik Deutschland.

Ermitteln Sie die Zielperson, indem Sie alle Personen der Zielgruppe dem Alter nach geordnet - älteste Person zuerst - auflisten und dann laut Auswahl Schlüssel auswählen.

Zielgruppe im Haushalt				Auswahl Schlüssel							
Alter	m/w	Alter	m/w								
1. _____	_____	5. _____	_____	Anz. Pers. i.d. Zielgruppe							
2. _____	_____	6. _____	_____	1 2 3 4 5 6 7 8							
3. _____	_____	7. _____	_____	1 1 2 1 5 4 3 4							
4. _____	_____	8. _____	_____	Schlüssel							

KANTAR PUBLIC= 7

Abbildung 6: Anwendungsbeispiel für den Schwedenschlüssel

Im Rahmen der Basiserhebung wurden 6.681 Interviews realisiert. Informationen über die Ausfallgründe und die Ausschöpfung finden sich in Kapitel 3.2.1.

2.5.2 Anlage der LEO 2018 Zusatzerhebung

Grundgesamtheit der LEO-Zusatzerhebung ist die in Privathaushalten lebende Wohnbevölkerung der Bundesrepublik Deutschland von 18 bis 64 Jahren mit für die Befragung ausreichenden Deutschsprachkenntnissen, die als höchsten Schulabschluss über einen Hauptschulabschluss verfügt, ein niedriger einzustufendes schulisches Qualifikationsniveau hat oder keinen Abschluss vorweisen kann.

Das für die Definition der korrekten Untersuchungseinheiten notwendige Screening wurde aus ökonomischen Erwägungen nicht direkt auf Basis einer Zufallsstichprobe ermittelt, sondern mittelbar auf Basis des bei Kantar regelmäßig durchgeführten CAPI-Busses. Hierbei handelt es sich um eine bevölkerungsrepräsentative Mehrthemenbefragung von in Deutschland in Privathaushalten lebenden Personen ab 14 Jahren. Die Anlage der Stichprobe des Kantar-CAPI-Busses erfolgt mit Ausnahme der ausgewählten Grundgesamtheit (Alter) analog zum in Kapitel 2.5.1 beschriebenen Verfahren der LEO-Basiserhebung. Die Bruttostichprobe für die LEO-Zusatzbefragung weist damit ebenfalls die Eigenschaften einer Zufallsstichprobe auf. Jeweils am Ende der Mehrthemenbefragung wird gefragt, ob die Zielperson mit einer neuerlichen Befragung durch Kantar einverstanden ist und wenn ja, werden die Kontaktinformationen dafür erhoben.

Neben den für verschiedene Auftraggeber und Auftraggeberinnen erhobenen Informationen liegt aus der Mehrthemenbefragung eine sogenannte Standardstatistik vor, die für Zusatzerhebungen spezieller Personengruppen herangezogen werden kann, sofern die Zielpersonen sich damit im Vorfeld einverstanden erklärt haben.

Für die LEO-Zusatzerhebung wurde die o.g. Zielpersonengruppe mithilfe der Merkmale Schulabschluss und Alter für die Bruttostichprobe aus den CAPI-Bus-Erhebungen der letzten 18 Monate vor Feldstart mit einem überhöhten Stichprobenansatz ausgewählt. Bei Ziehung der Bruttostichprobe wurde darauf geachtet, dass die Altersgruppen der 18- bis 34-Jährigen, der 35- bis 49-Jährigen und der 50- bis 64-Jährigen jeweils proportional im Verhältnis zu den Strukturen der Grundgesamtheit verteilt sind. Im Rahmen der Zusatzerhebung wurden 511 Interviews realisiert (Kapitel 3.2.2).

2.5.3 Anlage der Zusatzerhebung im Pretest

Die Zusatzerhebung wurde im Pretest auf Basis einer Quotenstichprobe durchgeführt, um die in Kapitel 2.5.2 vorliegenden Adressen aus dem CAPI-Bus komplett für die Haupterhebung nutzen zu können (analog zu LEO 2010). Die Entscheidung für eine andere, qualitativ weniger hochwertige Stichprobe im Rahmen des Pretests war vertretbar, weil es im Pretest nicht darum ging, das Screening-Verfahren oder Stichprobeneffekte zu prüfen, sondern vor allem darum, ausreichend Fälle mit einem bearbeiteten Alphaheft bereitstellen zu können.

Grundgesamtheit der Zusatzerhebung im Pretest ist die in Privathaushalten lebende Wohnbevölkerung der Bundesrepublik Deutschland zwischen 18 und 64 Jahren mit für die Befragung ausreichenden Deutschsprachkenntnissen, die als höchsten Schulabschluss über einen Hauptschulabschluss verfügt, ein niedriger einzustufendes schulisches Qualifikationsniveau hat oder keinen Abschluss vorweisen kann.

Für das *Quota Sampling* melden die Interviewerinnen und Interviewer des Kantar-Stabes Zielpersonen der Grundgesamtheit nach bestimmten Strukturmerkmalen (Alter, Geschlecht, Region), bei denen sie ein Interview realisieren können. Aus den gemeldeten Adressen wird eine Stichprobe gemäß der vorgegebenen Strukturmerkmale nach Randverteilungen, angepasst an die Merkmale des Mikrozensus bzw. der Bevölkerungsfortschreibung, gezogen. Im Rahmen des Pretests wurden 112 Interviews im Rahmen der Quotenstichprobe realisiert. Eine Rücklauf- bzw. Ausschöpfungsmeldung im Rahmen einer Quotenstichprobe kann nicht sinnvoll berechnet werden und wird daher im Rahmen dieses Berichts nicht erstellt.

2.5.4 Integration beider Stichproben der Haupterhebung

Bei der Auswertung der Daten ist es möglich, sowohl auf die Basis- und Zusatzerhebung getrennt zuzugreifen als auch die beiden Stichproben zu integrieren. Letzteres führt jedoch aufgrund der Definition der Grundgesamtheit in der LEO-Zusatzerhebung zu einer überproportionalen Berücksichtigung von Personen mit niedrigem schulischem Abschluss in der Gesamtstichprobe. Entsprechend wurden für die Arbeiten mit der integrierten Stichprobe eigene Gewichtungsfaktoren erstellt, welche die Verzerrung des Samples hinsichtlich der Bildungsgruppen ausgleichen (genauer s. Kapitel 6.3).

3 Feldarbeit

Die Feldarbeit der LEO-Grundbildungsstudie wurde von Interviewerinnen und Interviewer des Kantar-Stabs durchgeführt. Insgesamt waren 367 für LEO geschulte Interviewerinnen und Interviewer (Kapitel 2.4.3) an der Studie beteiligt. Das folgende Kapitel berichtet über die durchschnittliche Interviewdauer (Kapitel 3.1), die Ausfälle bzw. die Beteiligungsrate (Kapitel 3.2) sowie die standardisierten und für LEO 2018 (gesondert) unternommenen Qualitätskontrollen des Kantar-Interviewerstabes (Kapitel 3.3). In einem nachgelagerten Schritt wurde im Rahmen der Interviews erhoben, ob sich eine Befragungsperson an einer Wiederholungsbefragung durch Kantar beteiligen würde. Darüber wird im Exkurs-Teil am Ende dieses Kapitels berichtet.

3.1 Durchschnittliche Interviewdauer

Die durchschnittliche Interviewdauer des Hintergrundfragenprogramms betrug 36 Minuten, die durchschnittliche Bearbeitungsdauer der Rätselhefte betrug zwölf Minuten und die der Alphahefte („Joschi“, „Helga“ und „Leschek“) im Mittel zwei Minuten. Das gesamte Interview einschließlich Literalitätstest umfasste eine durchschnittliche Dauer von 49 Interviewminuten. Tabelle 3 zeigt die Ergebnisse insgesamt und getrennt nach den beiden Stichproben sowie getrennt nach dem Bildungshintergrund der Zielperson.

Ein Problem, das sich bei der Durchführung der Pretesterhebung zeigte, war die überdurchschnittlich hohe zeitliche Belastung der Personen, die im Rahmen der LEO-Erhebung ein Alphaheft bearbeiteten (s. Kapitel 2.2.2, Tabelle 1). In dieser Gruppe sollte die durchschnittliche Interviewdauer, im Einverständnis mit der Universität Hamburg, nach Möglichkeit in der Haupterhebung eine Zeitstunde nicht übersteigen. Letzteres wurde mithilfe einiger Instrumentenmodifikationen für die Haupterhebung angestrebt (s. Kapitel 2.2.1). In der Haupterhebung wurde das Ziel erreicht. Personen, die das Hintergrundfragenprogramm beantworteten und sowohl ein Rätsel- als auch ein Alphaheft bearbeiteten, wendeten durchschnittlich 58 Minuten dafür auf. Dass die durchschnittlichen Interviewzeiten dennoch von starken Unterschieden in den Einzelinterviews geprägt sind, zeigen die Standardabweichungen (SD).

Tabelle 3: Durchschnittliche Interviewdauer in ganzen Minuten (SD) insgesamt und getrennt nach Stichprobe und Bildungsniveau der Zielpersonen

	Basis: alle 18- bis 64-Jährigen				
	Durchschnittliche Interviewdauer in Minuten (SD)				
	alle Interviews (n=7.192)	Interviews der Basis- stichprobe (n=6.681)	Interviews der Zusatz- stichprobe (n=511)	ZP ¹⁾ : niedrige Bildung ²⁾ (n=2.104)	ZP ¹⁾ : mittlere bis hohe Bildung ²⁾ (n=4.932)
Hintergrundfragen	36 (17)	36 (17)	41 (16)	36 (15)	36 (18)
Rätselheft	12 (20)	11 (20)	15 (6)	13 (17)	11 (21)
Alphaheft	2 (7)	2 (7)	4 (5)	4 (6)	1 (8)
Gesamt	49 (19)	49 (19)	61 (19)	53 (19)	48 (19)
Neue Basis: Personen mit Rätsel- und Alphaheft:	(n=1.689)	(n=1.447)	(n=251)	(n=905)	(n=760)
Joschi	9 (7)	9 (8)	10 (5)	9 (5)	8 (5)
Helga	6 (21)	5 (23)	7 (3)	7 (4)	4 (31)
Leschek	8 (7)	7 (7)	9 (3)	8 (9)	8 (7)
Dauer Alphahefte	7 (14)	7 (15)	9 (4)	8 (6)	6 (19)
Gesamt	58 (19)	56 (18)	66 (20)	58 (19)	56 (18)

¹⁾ ZP: Zielperson

²⁾ Niedrige Bildung: Die Zielperson verfügt höchstens über einen Hauptschulabschluss.

Mittlere bis hohe Bildung: Die Zielperson verfügt über einen höheren Abschluss als einen Hauptschulabschluss.

Fehlende Werte zur Gesamtsumme: Schüler, fehlende Angabe

3.2 In LEO 2018 realisierte Interviews und Ausfälle

Bei der Gegenüberstellung von Brutto- und Nettofallzahlen ist zu berücksichtigen, dass die beiden Stichproben der Basis- und der Zusatzerhebung unterschiedlich angelegt wurden (s. Kapitel 2.5). Erstere basiert auf der Auswahl zufälliger Startpunkte, von welchen aus die Interviewerinnen und Interviewer nach bestimmten Regeln einem *Random Walk* folgten und aus jedem dritten Haushalt eine Zielperson zufällig auswählten. Die LEO-Zusatzerhebung dagegen basiert auf einer bestimmten Anzahl von Bruttoadressen, die aus dem Kantar-CAPI-Bus ausgewählt wurden.

Die Feldarbeit im Rahmen der LEO-Grundbildungsstudie begann am 1. März 2018²³ und wurde am 3. September 2018, also nach sechsmonatiger Dauer, abgeschlossen.²⁴ Der Umfang der realisierten Nettostichprobe beträgt 7.192 Interviews und entspricht damit der Soll-Vorgabe. Von den Interviews wurden 6.681 im Rahmen der Basis- und 511 im Rahmen der Zusatzerhebung realisiert.

²³ Der Feldstart von LEO 2018 lag damit acht Jahre minus einen Tag nach demjenigen von LEO 2010.

²⁴ Das Feldende der Zusatzerhebung erfolgte am 8. Juni 2018 und damit – fallzahlbedingt – deutlich früher als das der Basiserhebung.

3.2.1 Ausschöpfung und Ausfälle der LEO 2018-Basiserhebung

Die Anzahl der vorgegebenen Adressen, die im Rahmen der Feldzeit vom 1. März bis 3. September 2018 angegangen wurden und die unterschiedlichen Ausfälle für die Basiserhebung sind Tabelle 4 zu entnehmen. Wenngleich die Interviewerinnen und Interviewer aufgrund einer Umstellung des Systems vor knapp zwei Jahren stärker als bis dahin angehalten sind, die Ausfallgründe zu benennen, ist dies nicht für alle Adressen möglich. In der Berechnung ist daher ein Schätzwert enthalten.

Tabelle 4: Realisierte Interviews und Ausfälle nach Ausfallgründen in der LEO-Basiserhebung

Zielpersonen: Wohnbevölkerung der Bundesrepublik Deutschland von 18 bis 64 Jahren	Anzahl	Anteilswerte in %
Bruttoansatz	107.331	
keine ZP im Haushalt	10.333	9,6
nicht verwendbare Adressen	10.333	9,6
bereinigtes Brutto	96.998	
nicht abschließend bearbeitet	6.644	6,8
im Haushalt niemand angetroffen	30.985	31,9
Adressen unbekannter Eignung	37.629	38,8
vollständig bearbeitete Adressen	59.369	
ZP während der Feldzeit nicht erreichbar	9.435	15,9
körperlich/geistig nicht in der Lage	1.468	2,5
KP Sprachproblem	1.007	1,7
ZP Sprachproblem	1.845	3,1
qualitätsneutrale Ausfälle	13.755	23,2
verbleibende Adressen (II)	45.614	
nicht zum Interview bereit	38.791	85,0
Verweigerungen	38.791	85,0
Realisierte Interviews (I)¹⁾	6.823	15,0
davon vollständige Interviews	6.785	14,9
davon nicht auswertbar	38	0,1
Summe Interviews zur Auswertung	6.681	14,6
<i>Anzahl Sample Points</i>	1.300	

¹⁾ Cooperation Rate 4 nach AAPOR (American Association for Public Opinion Research)

Schätzung der Bearbeitungsergebnisse für die Adressen für die diese Informationen von den Interviewerinnen und Interviewern nicht angegeben wurden.

Mit unserer Darstellung in Tabelle 4 folgen wir, anders als noch im Methodenbericht für LEO 2010, den internationalen AAPOR-Standards.²⁵ Die AAPOR-Vorgaben sehen verschiedene Arten von Rücklaufindikatoren vor (vier *Response Rates* und vier *Cooperation Rates*, AAPOR 2016). Für die Basiserhebung von LEO 2018 können vier dieser acht Indikatoren mithilfe der in Tabelle 4 stehenden Informationen berechnet werden. Wir weisen hier die *Cooperation Rate 4* nach AAPOR, also die Ausschöpfungsrate, aus. Sie beträgt 15 Prozent.²⁶

²⁵ AAPOR: American Association for Public Opinion Research

²⁶ Die Ausschöpfungsrate wird berechnet, indem alle realisierten Interviews (I) durch die verbleibenden Adressen (II) geteilt werden.

3.2.2 Ausschöpfung und Ausfälle der LEO 2018-Zusatzerhebung

Die Feldarbeit im Rahmen der Zusatzerhebung begann am 1. März 2018 und endete am 8. Juni 2018, also nach etwa dreimonatiger Dauer. Zur Sicherstellung der Stichprobenqualität wurde, wie in Kapitel 2.1 beschrieben, relativ weit vorne im Hintergrundfragenprogramm, das sowohl im Rahmen der Basis- als auch der Zusatzerhebung eingesetzt wurde, mithilfe der Frage nach dem schulischen Abschluss geprüft, ob die befragte Person, tatsächlich Zielperson der LEO-Zusatzerhebung ist. Entsprechend der Vereinbarungen nach Sichtung der Pretestergebnisse wurden Personen, die (doch) nicht zur Zielgruppe gehören, von der Zusatzerhebung ausgeschlossen. Die Nettostichprobe der Zusatzerhebung umfasst insgesamt 511 Fälle. Die Anzahl der Bruttoadressen, die unterschiedlichen Ausfälle sowie die Ausschöpfung der Stichprobe sind Tabelle 5 zu entnehmen. Die Rücklaufquote beträgt 71 Prozent.

Tabelle 5: Realisierte Interviews und Ausfälle nach Ausfallgründen in der LEO-Zusatzerhebung

Zielpersonen: Wohnbevölkerung der Bundesrepublik Deutschland von 18 bis 64 Jahren mit niedrigem schulischen Bildungsniveau			Anzahl	Anteilswerte in %
Bruttoansatz			1.554	
	ZP wohnt nicht (mehr) unter der Adresse		75	4,8
	ZP gehört nicht in die Zielgruppe		7	0,5
	sonstige Nicht-Eignung		93	6,0
nicht verwendbare Adressen			175	11,3
bereinigtes Brutto			1.379	
	nicht abschließend bearbeitet		511	37,1
	Adresse falsch/existiert nicht		16	1,2
	sonstige unbekannter Eignung		0	0,0
Adressen unbekannter Eignung			527	38,2
vollständig bearbeitete Adressen			852	
	ZP während der Feldzeit nicht erreichbar		33	3,9
	niemand im Haushalt erreicht		22	2,6
	ZP verstorben		7	0,8
	körperlich/geistig nicht in der Lage		16	1,9
	KP/ZP Sprachproblem		7	0,8
qualitätsneutrale Ausfälle			85	10,0
verbleibende Adressen (II)			767	
	nicht zum Interview bereit		219	28,6
Verweigerungen			219	28,6
Realisierte Interviews (I)¹⁾			548	71,4
	davon nicht auswertbar		37	4,8
Summe Interviews zur Auswertung			511	66,6

¹⁾ Cooperation Rate 4 nach AAPOR (American Association for Public Opinion Research)

3.3 Qualitätskontrolle Interviewer

Im Rahmen der Erhebung der LEO-Grundbildungsstudie wurden zur Qualitätssicherung einige standardisierte und einige projektspezifische Schritte der Interviewerkontrolle unternommen, die im Folgenden beschrieben werden. Vorab erfolgt eine allgemeine Information über die CAPI-Feldorganisation und den Kantar-Interviewerstab.

3.3.1 Allgemeine Informationen über die Kantar-CAPI-Line sowie den Kantar-Interviewerstab

Der gesamte Prozess der Rekrutierung, Schulung (s. Kapitel 2.4), Organisation und Führung sowie die Kontrolle der Feldperformance der Interviewerinnen und Interviewer ist bei der Feldorganisation von Kantar angesiedelt und ermöglicht ein für *Best-Practice-Verfahren* im Bereich anspruchsvoller Erhebungen unerlässliches *total quality management* (TQM).

Die zentrale Steuerung des Interviewerfeldes von Kantar erfolgt über die sogenannte Feldorganisation in München. Der *Feldorganisation* obliegt die komplette Steuerung des institutseigenen Interviewerstabes. Zur Tätigkeit der Feldorganisation zählen unter anderem

- aufwendige Prozesse der Interviewerrekrutierung
- Aufbau und Pflege der datenbankgestützten Informationssysteme zur Verwaltung des Interviewerstabes
- Aufbau und Pflege der datenbankgestützten Steuerung des Interviewerstabes (*Performance-Rating*)
- Steuerung der Stichprobenbearbeitung und Erstellen von Rücklaufstatistiken
- Kooperation mit der Studienleitung
- Erstellung der Feldunterlagen und gegebenenfalls der Schulungsunterlagen für Interviewerinnen und Interviewer in Absprache mit der Studienleitung

Neben der Feldorganisation ist bei Kantar vor allem der sogenannte regionale Einsatz für eine effektive Interviewerführung zuständig. Fünf regionale Einsatzleiterinnen und -leiter steuern auf Basis einer systematischen Eignungs- und Qualitätsbeurteilung in einer Datenbank den Einsatz mit Interviewerinnen und Interviewern auf Ebene der Einzelprojekte. Die Einsatzleitung nimmt ein kontinuierliches Monitoring der Stichprobenbearbeitung, dokumentiert in Rücklaufstatistiken vor, die der Studienleitung zeitnahe Informationen zum Bearbeitungsfortschritt ermöglicht. Dieses Monitoring ist von entscheidender Bedeutung, um mögliche Auffälligkeiten oder Schwierigkeiten bei der Bearbeitung der Stichprobe frühzeitig im Feldverlauf zu erkennen und gegebenenfalls einzugreifen.

Die regionalen Einsatzleiterinnen und -leiter nehmen bei großen Stichproben unvermeidbare Umbesetzungen von Interviewerinnen und Interviewern vor und organisieren in solchen Fällen gegebenenfalls eine nachträgliche Schulung²⁷ sowie den nachträglichen Versand der Feldunterlagen. Durch regelmäßige Besprechungen der Einsatzleitung mit den Kontaktinterviewerinnen und -interviewern sowie der internetbasierten Kommunikation letzterer mit sämtlichen Interviewerinnen und Interviewern wird eine professionelle Führung des Stabes gewährleistet.

²⁷ Im Falle von LEO 2018 wurden hierfür, in Absprache mit der Universität Hamburg, das Schulungsvideo sowie sämtliche Schulungsunterlagen verwendet (s. Kapitel 2.4.3).

3.3.2 Standardisierte Interviewerkontrolle durch die CAPI-Line

Die Qualität des Kantar-Interviewerstabes bzw. die Arbeit der Interviewerinnen und Interviewer wird standardmäßig durch den Einsatz verschiedener Kontrollverfahren bei mindestens zehn Prozent der durchgeführten Interviews geprüft. Die Auswahl der zu kontrollierenden Studien, Interviews oder Interviewerinnen und Interviewer wird zufällig vorgenommen. Die korrekte Durchführung eines Interviews wird dabei auf unterschiedlichen Wegen festgestellt. Einerseits werden standardisierte Postkarten oder Validierungsbögen direkt an die Zielpersonen geschickt und bestimmte Kriterien zur Interviewdurchführung abgefragt. Andererseits werden telefonische Befragungen bei Zielpersonen zur Durchführung des jeweiligen Interviews vorgenommen. Diese klassische Interviewer- bzw. Qualitätskontrolle wird durch die elektronisch abgespeicherten Daten ergänzt. Hierzu gehören folgende Kontrollen.

- Kontrolle der Existenz von angegebenen Adressen
- Kontrolle der Erfassung von Adressen auf Vollständigkeit und Mehrfachbefragungen
- Kontrolle der Einhaltung von Auswahlvorschriften, wie z.B. des Begehungsweges
- Kontrolle der Einhaltung von Methodenspezifikationen, wie z.B. der Zielpersonenauswahl
- Kontrolle der Interviewdauer und der zeitlichen Lage eines Interviews sowie der zeitliche Abstand zwischen zwei Interviews einer Interviewerin bzw. eines Interviewers
- Kontrolle der Ausfallgrundangaben

Interviewerkontrollen finden zu folgenden Zeitpunkten oder Anlässen statt.

- Regelmäßige stichprobenartige Kontrollen
- Kontrolle bei überdurchschnittlich häufigem Einsatz einer Interviewerin bzw. eines Interviewers
- Kontrolle bei Auffälligkeiten
- Teil- oder Totalkontrolle im Falle einer expliziten Anforderung durch die Auftraggeberin bzw. den Auftraggeber²⁸

Sämtliche Informationen aus allen Kontrollformen werden in einer Datenbank dokumentiert. Dadurch können projektspezifische Informationen oder Informationen über bestimmte Regionen oder auf Ebene einer Interviewerin oder eines Interviewers zügig und umfangreich eingesehen werden.

Im Falle einer auffälligen Interviewerkontrolle wird die betroffene Interviewerin bzw. der Interviewer entsprechend der Auffälligkeit nachgeschult, sei es durch nochmaliges Schulen von grundlegenden Interviewerfähigkeiten (z.B. Ansprache, Methode etc.) oder von projektspezifischen Punkten (z.B. Zielpersonenauswahl). Nicht korrekt durchgeführte Interviews werden nicht ausgewertet, nicht honoriert und nicht ausgeliefert. Im Falle grober Verstöße werden die entsprechenden Interviewdaten unverzüglich gelöscht. Im Falle massiver Häufungen oder schwerer Vergehen seitens einer einzelnen Interviewerin oder eines Interviewers beendet Kantar die Zusammenarbeit.

Darüber hinaus wurde mit dem sogenannten *BUSradar* von Kantar ein einzigartiges Instrument zur Ergänzung der Interviewerkontrolle bzw. der Interviewqualität geschaffen. Im *BUSradar* werden *alle* eingehenden CAPI-Interviews tagesaktuell erfasst und automatisierte Analysen der Befragungs- und Durchführungsdaten durch-

²⁸ Bei einer Auftraggeberseitig gewünschten Kontrolle können spezielle Anforderungen, wie zum Beispiel Kontrollart, Umfang oder Aufnahme bestimmter Fragen auf Karten oder Validierungsfragebögen umgesetzt werden. Diese Form der Kontrolle wird vorab vertraglich vereinbart.

geführt (z.B. werden regelmäßige Prüfungen von Antworten zu bestimmten Hauptfragen eines Langzeitprojektes oder zeitliche Abweichungen von der durchschnittlichen Interviewdauer).²⁹ So wird gewissermaßen ein *Profiling* analog zur modernen Kriminalistik vorgenommen, denn Daten der Markt- und Meinungsforschung verraten einiges darüber, unter welchen Umständen sie generiert wurden. Datenauffälligkeiten treten insbesondere dann auf, wenn Interviews unter starken externen Einflüssen oder durch Fehlverhalten z.B. einer Interviewerin oder eines Interviewers zustande gekommen sind. Im Falle von Auffälligkeiten wird im *BUSradar* ein bestimmter Code hinterlegt und eine Interviewerkontrolle folgt entsprechend des oben beschriebenen Vorgehens. Kantar geht davon aus, dass die alleinige Tatsache von systematischen Interviewerkontrollen auf verschiedenen Wegen nicht nur Vergehen aufdeckt, sondern auch präventiv wirkt.

3.3.3 Projektspezifische Interviewerkontrolle

Wenngleich die LEO-Daten bereits im Rahmen eines standardisierten Prozesses feldseitig geprüft wurden (s. Kapitel 3.3.2), erfolgt ein ähnlicher Schritt nochmals auf andere Weise. Durch die Projektleitung werden alle Interviews auf Vollständigkeit (s. auch Kapitel 4), hinsichtlich einer überdurchschnittlichen Häufung der Antwortvorgabe „keine Angabe“ sowie hinsichtlich der jeweiligen Interviewdauer in Minuten insgesamt und je Fragenblock (s. Kapitel 3.1), auch getrennt nach Interviewenden geprüft.

Ein weiteres Indiz für die Qualität der Interviewarbeiten wird mithilfe des der Erhebung nachgelagerten Editings vorgenommen. Im Rahmen des Editings werden – pauschal ausgedrückt – die Testhefte durchgesehen, bewertet und die Bewertungen in den LEO-Datensatz aufgenommen. Auf eine genaue Beschreibung des Editings wird an dieser Stelle verzichtet und auf Kapitel 4.2 bis 4.3 verwiesen. Bedeutsam ist an dieser Stelle, dass im Rahmen des Editings die von den Befragungspersonen bearbeiteten Testhefte genau durchgesehen werden. Dabei fallen Unzulässigkeiten, wie z.B., zwei unterschiedliche Schriften in einem Heft oder Interviewerkommentare auf. Beide Punkte gehen in die Qualitätsprüfung ein. Im Falle von Kommentaren ist fallweise über eine (Un-)Zulässigkeit zu entscheiden. Im Falle von *zwei Schriften* wurde während der Feldphase eine Nachschulung der entsprechenden Interviewerin bzw. des Interviewers eingeleitet und das Interview gegebenenfalls von der Datenübergabe ausgeschlossen.

Viele ex-post-Verfahren zur Prüfung von Datenqualität, egal für welche Erhebungsmethode, basieren auf Vergleichen einfacher Verteilungen (s.o. und Kapitel 3.3.2). Befragungsdaten, die bewusst gefälscht wurden, fallen nicht immer in einer solchen Betrachtung auf. Wenngleich gefälschte Daten äußerst selten vorliegen, ist es wichtig, ein Prüfverfahren einzusetzen, das Fälschungen schnell identifizieren kann.³⁰

Die LEO-Daten bzw. die Interviewerinnen und Interviewer wurden zusätzlich mit dem von Kantar eigens dafür entwickelten Verfahren überprüft, das im standardisierten Qualitätsprüfungsprozess (noch) nicht inbegriffen ist, dem sogenannten *Similarity Check*³¹. Das Prüfprogramm basiert auf der Hypothese, dass Interviewerinnen und Interviewer, die Interviews fälschen, Daten mit geringerer Varianz erzeugen als solche, die nach den vorgegebenen Regeln Interviewdaten erheben. Dabei spielt weder die Erhebungsart noch die detaillierte Kenntnis des Fragenprogramms eine große Rolle. Das R-basierte³² Prüfprogramm berechnet die *Ähnlichkeit* zwischen jeweils zwei Fällen in einem Datensatz und kann so Fallpaare identifizieren, die besonders viele

²⁹ Mithilfe des *BUSradars* wird die Verteilung verschiedener Merkmale eines Interviews (z.B. Interviewdauer) aggregiert einer Interviewerin bzw. eines Interviewers geprüft. Es erfolgt ein Vergleich mit definierten Schwellenwerten oder bekannten Verteilungen, um Auffälligkeiten (z.B. Kurz-Interviews, Nachtinterviews, An- und Abfahrten von einem zum nächsten Interviewtermin und Mehrfachbefragungen einer Person) identifizieren zu können.

³⁰ Dies ist nicht nur für die Datenqualität, sondern auch für den Ruf des Hauses von Bedeutung.

³¹ Wir danken Rita Reimer der Abteilung AMS (*Applied Marketing Science*) für Ihre Unterstützung bei der Durchführung und Beschreibung des *Similarity Checks*.

³² R: Statistisches Programm zur Datenanalyse

Ähnlichkeiten aufweisen, also in nur wenigen Variablen unterschiedliche Ausprägungen haben. Folgende Formel bringt dies zum Ausdruck:

$$\frac{\sum_1^k \text{sgn} |x_{il} - x_{jl}|}{k} \quad \begin{array}{l} k = \text{Anzahl der Variablen} \\ l = \text{Zähler über die Variablen} \end{array}$$

Jedes Fallpaar x_i und x_j wird dabei nur einmal betrachtet, die Reihenfolge von i und j spielt keine Rolle.

Die *Größe* der Unterschiede je Merkmal hat dabei keinen Einfluss, weil alle Variablentypen und Skalenniveaus gleichbehandelt werden. Anstelle alphanumerischer Variablen werden gegebenenfalls die nach einer Ver-codung erzeugten numerischen Variablen verwendet (s. Kapitel 4.1). Daraus resultiert eine vereinfachte Distanz-Matrix, d.h. einer paarweisen Berechnung von Differenzen zwischen allen Fällen: „Wie viele Variablen sind zwischen diesen zwei Fällen unterschiedlich?“. Damit die Distanzwerte nicht verzerrt sind, werden vor der Analyse Variablen aus dem Datensatz entfernt, die a) nur filterabhängig beantwortet werden, b) per Definition innerhalb einer Analyseschicht ähnlicher sein müssen als andere Variablen oder c) von denen bekannt ist, dass sie einen relativ hohen Anteil fehlender Werte (auch „keine Angabe“) enthalten.

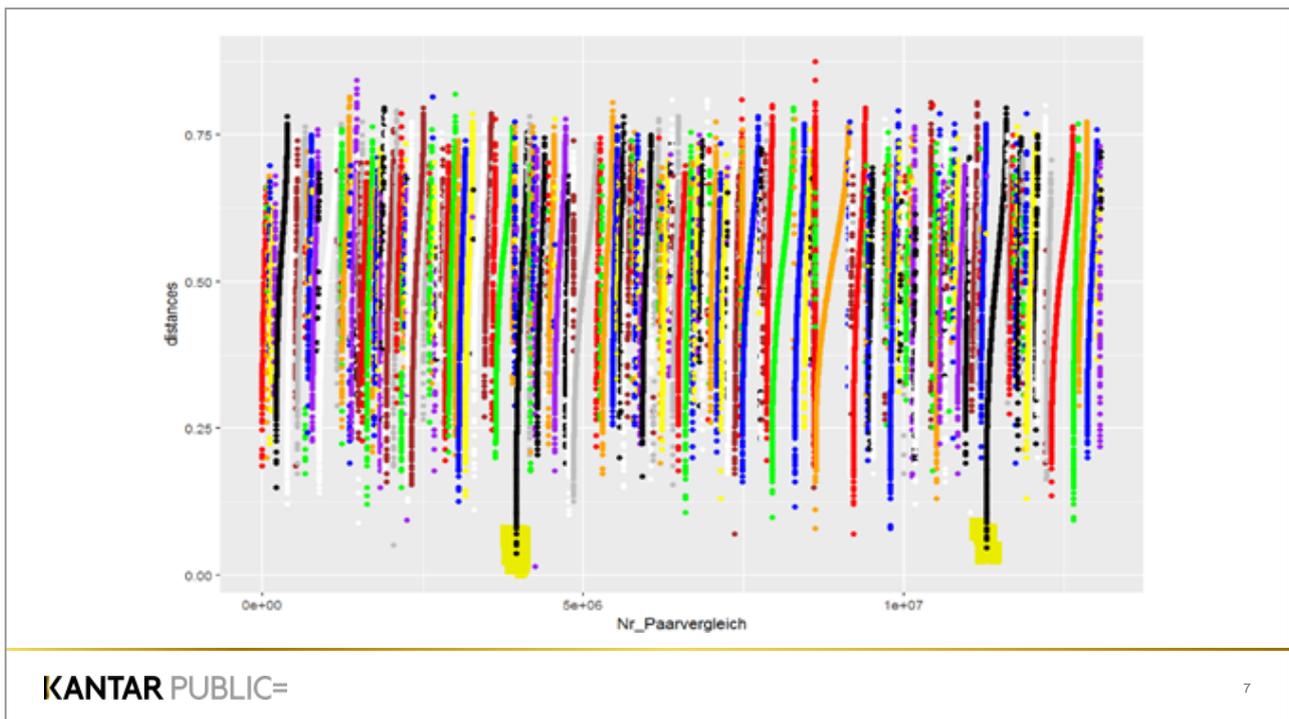


Abbildung 7: Beispiel für eine gruppierte Betrachtung von Distanzwerten (nicht LEO 2018)

Aus der o.g. Distanz-Matrix lassen sich (normierte) Distanzwerte zwischen 0 und 1 ableiten. Dabei entspricht der Wert „0“ keiner Distanz, was bedeutet, dass alle Variablen gleich also sogenannte *Dubletten* sind. Der Wert „1“ entspricht einer maximalen Distanz, die anzeigt, dass keine der Variablen gleich ist. Der zu interpretierende Distanz- bzw. Ähnlichkeitswert liegt demnach zwischen den Werten „0“ und „1“. Durch vielfache Visualisierungen können Fallpaare identifiziert und weiter analysiert werden, um zu entscheiden, ob ihnen ein Fehlverhalten seitens der Interviewerin oder des Interviewers zugrunde liegt. Anschließend werden die Distanzwerte nach verschiedenen Analyseschichten (hier Interviewende) betrachtet. Beurteilt wird nun, ob und in welcher Größenordnung eine Häufung von sehr geringen Distanzen innerhalb derselben Schicht im Vergleich

zu Distanzen zwischen verschiedenen Schichten besteht. Abbildung 7 zeigt einen derart auf Ebene der Interviewerinnen und Interviewer ermittelten Paarvergleich beispielhaft entnommen aus einer anderen Studie. Hier zeigt sich, dass die Interviews einiger Interviewerinnen und Interviewer deutlich ähnlicher sind als im Durchschnitt (gelbe Markierungen).

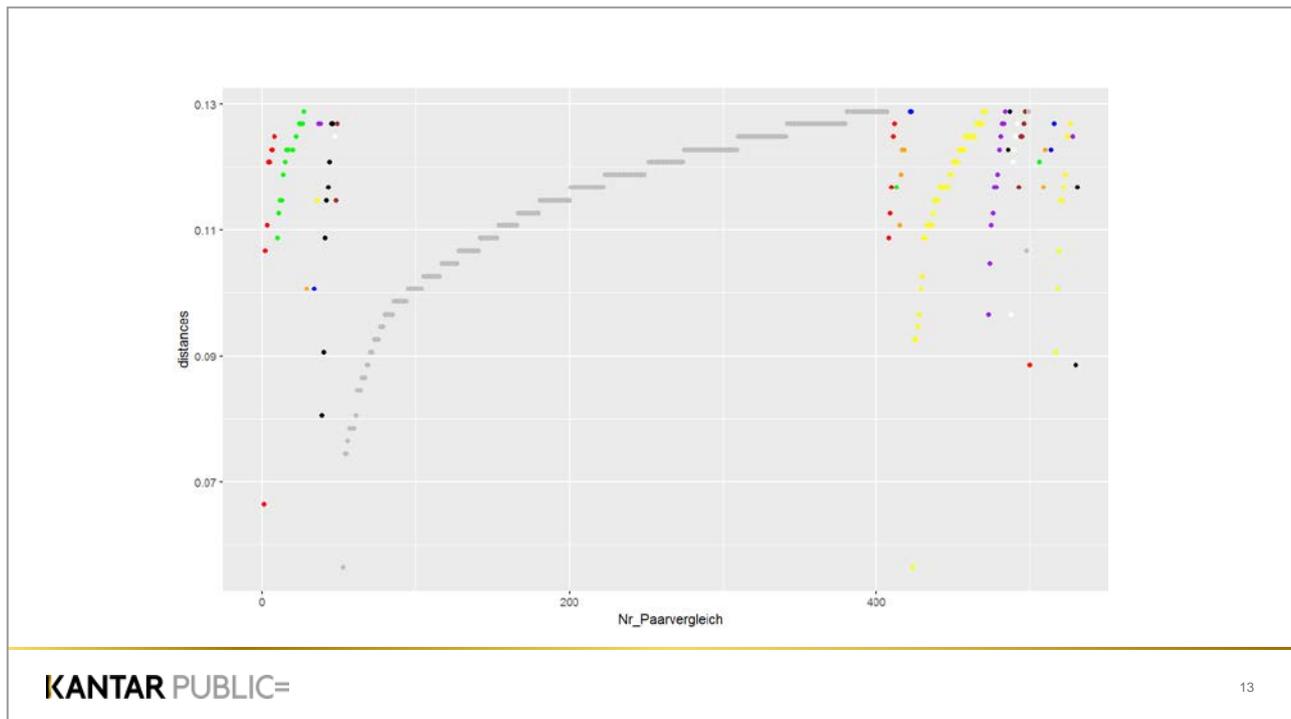


Abbildung 8: Gruppierter Betrachtung von Distanzwerten beschränkte auf Interviewer/innen mit Auffälligkeiten in den LEO 2018-Interviews

Abbildung 8 zeigt eine andere Visualisierungsart, beschränkt auf Interviewerinnen und Interviewer mit Auffälligkeiten, entnommen aus geprüften Vorabdaten der LEO-Haupterhebung. Für die Interviewerinnen und Interviewer, für die darin die Distanzwerte in grüner, gelber und grauer Farbe abgezeichnet sind (eindeutige Muster), lagen zudem (teils) Hinweise auffälliger Art aus anderen Prüfroutinen vor. In einem Fall hat sich zwar nicht der Verdacht der Fälschung, wohl aber einer eigenmächtigen Zielpersonenauswahl erhärtet. Alle Interviews dieses Interviewers (s. Kapitel 3.2.1 und 3.2.2 und „nach Prüfung nicht in auswertbar“ in den Tabellen 4 und 5) wurden aus dem LEO 2018-Datensatz entfernt.³³

³³ Diese realisierten Fälle wurden in den Ausschöpfungstabellen 4 und 5 in diesem Bericht als „nicht auswertbar“ kategorisiert.

Exkurs: Teilnahmebereitschaft der LEO-Befragungspersonen für eine Wiederholungsbefragung

Weil auftraggeberseitig der Wunsch einer (auszugsweisen) Nachbefragung (einzelner) Befragungspersonen der LEO 2018-Erhebung geäußert wurde, hat sich die Kantar Public erlaubt, am Ende des LEO-Interviews die Frage nach einer potenziellen Teilnahmebereitschaft an einer weiteren Kantar-Studie zu erfragen und gegebenenfalls personenbezogene Kontaktinformationen zu erheben.

Die Hälfte (50 Prozent) aller Befragungspersonen gab eine Teilnahmebereitschaft an (Tabelle 6). Dabei lag der Wert der Befragungspersonen, die im Rahmen der Basiserhebung gefragt wurden mit 47 Prozent deutlich niedriger als derjenige von Personen, die im Rahmen der Zusatzerhebung um Auskunft gebeten wurden (88%). Dies ist insofern plausibel, als die Befragungspersonen, die sich an der Zusatzerhebung freiwillig beteiligten, bereits im Rahmen der dem Screening für die Zusatzerhebung vorgelagerten Kantar-CAPI-Bus-Befragung ihre Teilnahmebereitschaft angegeben hatten.

Tabelle 6: Bereitschaft für eine anschließende Befragung durch Kantar

Basis: 18- bis 64-Jährige (ungewichtet)	alle (n=7.192)		Basiserhebung (n=6.801)		Zusatzerhebung (n=511)	
	n	%	n	%	n	%
Zustimmung zu einer Kantar-Wiederholungsbefragung						
ja	3.592	49,9	3.142	47,0	450	88,1
nein	3.600	50,1	3.539	53,0	61	11,9
gesamt	7.192		6.681		511	
getrennt nach Literalitätstestbearbeitung in LEO 2018						
nur Rätselheft bearbeitet:						
ja	2.625	47,5	2.399	45,5	226	89,0
nein	2.896	52,5	2.868	54,5	28	11,0
gesamt	5.521		5.267		254	
Rätselheft und ein Alphaheft:						
ja	927	59,6	719	54,6	208	87,8
nein	628	40,4	599	45,4	29	12,2
gesamt	1.555		1.318		237	
darunter: "Joschi":						
ja	302	61,5	235	56,6	67	88,2
nein	189	38,5	180	43,4	9	11,8
gesamt	491		415		76	
darunter: "Helga":						
ja	323	59,5	246	54,1	77	87,5
nein	220	40,5	209	45,9	11	12,5
gesamt	543		455		88	
darunter: "Leschek":						
ja	302	58,0	238	53,1	64	87,7
nein	219	42,0	210	46,9	9	12,3
gesamt	521		448		73	

Personen, für die aus der Bearbeitung des Rätselhefts im CAPI ein Score größer oder gleich dem vorgegebenen Schwellenwert berechnet wurde und die somit kein Alphaheft erhielten, geben zu 48 Prozent eine Teilnahmebereitschaft an (Fälle aus der Basiserhebung: 46% und Fälle aus der Zusatzerhebung 89%). Für Personen, die aufgrund einer Unterschreitung des Schwellenwerts ein Alphaheft bearbeiteten, liegt der Vergleichswert mit 60 Prozent sogar noch höher. Immerhin 927 Personen würden nach eigener Aussage auch noch an einer Folgestudie teilnehmen.

Ein weiteres Merkmal zur Identifikation von Befragungspersonen mit Schriftsprachschwierigkeiten auf den Alpha-Levels 1 bis 3 kann zwar nicht mit den *Plausible Values*, auf Wunsch gegebenenfalls aber mit den *Weighted Likelihood Estimates* (WLE) herangezogen werden, sofern diese analog zu LEO 2010 neuerlich berechnet und bereitgestellt werden.

4 Datenaufbereitung und Auswertung der Testaufgaben

Die LEO-Grundbildungsstudie wurde überwiegend computergestützt erhoben. Der Hintergrundfragebogen ist als standardisierte CAPI-Befragung komplett programmiert und durch ein Listenheft ergänzt (s. Kapitel 2.2). Die nachgelagerte Messung der literalen Kompetenzen (*Testteil*) wurde als Methodenmix konzipiert. Die Aufgaben werden durch die Interviewerin oder den Interviewer vorgelesen und durch die Testpersonen selbst auf Papier, also in den vorgesehenen Heften, bearbeitet (s. Kapitel 2.1). So liegen die Testantworten nach der Befragung auf Papier vor und können ex-post ausgewertet werden. Im vorliegenden Kapitel werden zunächst die generellen Vorteile einer computergestützten Befragung beschrieben (Kapitel 4.1) und anschließend der komplexe Auswertungsprozess der Testaufgaben im Rahmen des sogenannten *Editings* (Kapitel 4.2 bis 4.3).

4.1 Aufbereitung der im CAPI erhobenen Daten

Der Vorteil von computergestützten Befragungen liegt darin, verschiedene Fehlerquellen bei der Beantwortung der Fragen bereits durch die Programmierung des Erhebungsinstruments auszuschließen. Die Filterführung erfolgt automatisch. Die zulässigen Wertebereiche sind festgelegt. Ungültige Werte können nicht eingegeben werden. Ferner werden bestimmte Konsistenz- bzw. Plausibilitätsprüfungen in den programmierten Fragebogen integriert. Diese Prüfungen geben im Falle von inkonsistenten oder nicht plausiblen Antworten im CAPI Fehlermeldungen aus, die bei der Interviewerin oder bei dem Interviewer eine Nachfrage an die Zielperson oder gegebenenfalls eine Fehlerkorrektur veranlassen. Um die Fehlerfreiheit des CAPI-Programms sicherzustellen, wurde es vor und direkt nach Feldstart umfangreich getestet. Der Test nach dem Feldstart konnte auf Basis erster Interviewdaten durchgeführt werden.

4.1.1 Prüfung auf Filterfehler

Zusätzlich zu dem umfangreichen Test des CAPI-Fragenprogramms wurde mittels eines Prüfprogramms sichergestellt, dass das CAPI-Programm hinsichtlich der Filterführung korrekt umgesetzt wurde und die Daten korrekt erhoben werden. Diese Prüfung erfolgte in den ersten Wochen nach Feldstart mehrfach auf Basis von *Vorabdaten*. Zur Überprüfung der korrekten Filterführung wurde in einer SPSS-Syntax die Filterführung des Erhebungsinstrumentes erneut programmiert (*Redundanz*) und mit den Daten abgeglichen. Sofern in einer Variablen nicht die korrekte Filterführung abgebildet wird, erzeugt das Prüfprogramm einen Fehlercode. Im Falle eines Fehlercodes wird die Ursache des Fehlers identifiziert und im CAPI-Programm berichtigt. Dieses

Vorgehen bewährte sich auch für die LEO-Grundbildungsstudie. Im Zuge der Vorabdatenprüfung konnte eine fehlerhafte Filterführung im CAPI-Programm kurz nach Feldstart identifiziert und korrigiert werden.

4.1.2 Vollständigkeit der Interviews und Umgang mit Antwortverweigerungen

Ein Interview im Rahmen der LEO-Grundbildungsstudie gilt dann als vollständig, wenn im CAPI-Fragenprogramm der Testteil begonnen wurde – unabhängig davon ob ein Testheft bei Kantar Public eingegangen ist.

Das CAPI-Interview ist so programmiert, dass die Interviewerinnen und Interviewer jede Frage stellen und die Antworten der Zielpersonen eintragen müssen. Damit entfällt die Möglichkeit einer Nicht-Beantwortung einzelner Fragen. Erst nach einem Eintrag in die vorgegebenen Antwortkategorien im CAPI kann die nächste Frage im Rahmen der Interviewsituation bearbeitet werden. Insofern sind beendete (und nicht abgebrochene) CAPI-Interviews vollständig.

Tabelle 7: Übersicht über „weiß nicht“- und „keine Angabe“ mit Anteilswerten von wenigstens fünf Prozent gültiger Antworten

Variablenbezeichnung	Frage	Anteilswerte in %	
		"weiß nicht"	"keine Angabe"
Dig025	Zutrauen mit Online-Stellenbörsen zurecht zu kommen	5,6	1,2
Dig026	Zutrauen mit Online-Partnerbörsen zurecht zu kommen	13,2	3,9
Dig027	Zutrauen mit Online-Wohnungsbörsen zurecht zu kommen	7,2	1,7
Arb008	Wunschberuf	-	5,1
Arb005	Anerkennung Berufstätigkeit: Erhalt von verdienter Anerkennung vom Vorgesetzten	-	5,8
Arb006	Anerkennung Berufstätigkeit: für erbrachte Leistungen und Anstrengungen angemessene Karriere	-	8,8
Arb010	Marktwert: Schwierigkeit vergleichbare Stelle zu finden auf freiem Markt	3,7	2,1
F022	Bruttolohn	-	23,4
Fin020	Zutrauen bestes Angebot der Altersvorsorge zu finden	5,9	1,7
Fin021	Vorteile und Risiken von Geldgeschäften beurteilen: Online-Banking	5,7	1,1
Fin022	Vorteile und Risiken von Geldgeschäften beurteilen: Käufe mit Ratenzahlung	5,3	1,7
Ges017	Zutrauen sich Unterstützung zu finden z.B. bei psychischen Problemen	4,3	0,8
Pol003	Sonntagsfrage	-	6,6
Pol007	Liebste Zeitung	-	5,3
Pol033	Politische Entscheidungen: Beurteilung, ob Beteiligung an Unterschriftenliste	6,2	2,7
Pol036	Politische Entscheidungen: Beurteilung, ob Teilnahme an Demo	5,1	2,4
Pol034	Politische Entscheidungen: Beurteilung, ob Partei vertritt, was wichtig ist	7,1	3,2
Wb016	Grund für Freistellung für Bildung	-	5,9
Wb027	Wunschthema für Weiterbildung	-	11,6
Wb033	Weiterbildungsempfehlung durch Vorgesetzten	-	6,5

In CAPI-Erhebungen ist es von großer Bedeutung ausschließlich Frageninstrumente einzusetzen, die eine inhaltliche Antwort ermöglichen. Die Beantwortbarkeit der im Rahmen von LEO 2018 eingesetzten Erhebungsinstrumente wurde im Rahmen des Instrumententests im Pretest eingehend geprüft und für die Haupterhebung leicht angepasst (Kapitel 2.2). Insofern ist davon auszugehen, dass alle im CAPI-Fragenprogramm enthaltenen Fragen der LEO-Grundbildungsstudie spontan von den Zielpersonen beantwortet werden konnten, gegebenenfalls mit Hilfe erklärender Interviewerhinweise.

Tabelle 7 gibt einen Überblick über die Fragen von LEO 2018 mit den höchsten Anteilswerten von „weiß nicht“ und „keine Angabe“.³⁴ Aufgeführt sind alle Variablen, die von der befragten (Teil-)Gruppe zu wenigstens fünf Prozent mit „weiß nicht“ oder „keine Angabe“ beantwortet wurden.

4.1.3 Datenaufbereitung und Codierung

Durch die dem CAPI-Programm immanenten Filterführungen, Vorgaben zulässiger Werte, Plausibilitäts- und Konsistenzprüfungen verringert sich der Aufwand der Datenprüfung nach Feldende. Nicht alle möglichen Fehlerquellen lassen sich allerdings ausschließen. Daher erfolgt auch bei CAPI-Interviews ein gewisser Teil der Datenprüfung, nach Abschluss der Feldarbeit.

Häufig nicht im Rahmen des CAPI-Programms machbar sind Codierungen offener Nennungen, die z.B. im Rahmen geschlossener Fragen mithilfe eines Statements wie „andere, und zwar (bitte angeben)“ erhoben werden. Diese offenen Nennungen werden hinsichtlich ihrer Zuordenbarkeit in die geschlossenen Antwortvorgaben überprüft und gegebenenfalls zugeordnet. Ebenfalls vercodet wurden offene, textliche Begründungen der Interviewerinnen und Interviewer zur fehlenden oder unvollständigen Bearbeitung von Aufgaben der Befragungspersonen in den Rätselheften (RH) und Alphaheften (AH; Kapitel 5.3). Nähere Informationen zum Umgang mit fehlenden oder unvollständigen Einträgen in den *Paper-&-Pencil*-Testheften werden nicht hier, sondern in den Kapiteln 5.2 bis 5.4 berichtet.

Darüber hinaus enthält das Hintergrundfragenprogramm eine Reihe von Informationen, die nach internationalen Klassifikationssystemen vercodet werden sollen. Dies gilt für ...

- die Berufe: Vercodung nach der International Standard Classification of Occupations (ISCO 08),
- die Wirtschaftszweige: Vercodung nach der Nomenclature des statistiques des Activités économiques de la Communauté Européenne (NACE rev. 2) und
- die Themen der Weiterbildungsaktivitäten: Vercodung nach dem *International Standard Code of Education* (ISCED-Fields) vercodet wurden.

Die Standardtexte zum Vorgehen bei der Vercodung nach ISCO, NACE und ISCED von Kantar Public wurden in Anhang A2 dokumentiert, auf die an dieser Stelle verwiesen wird. Sowohl die Befragungsdaten selbst als auch die generierten Informationen wurden für die Universität Hamburg aufbereitet als anonymisierter SPSS-Datensatz übergeben.

³⁴ Bei besonders schwierig zu beantwortenden Fragen oder Fragen zu sensiblen Themen (z.B. Gesundheit oder Einkommen) wurde neben der nicht von den Interviewerinnen und Interviewern vorzulesenden Antwortoption „keine Angabe“ zusätzlich die Antwortoptionen „weiß nicht“ eingefügt.

4.2 Bewertung der Testaufgaben für den Datensatz (*Editing*)

Kernstück der LEO-Grundbildungsstudie (LEO 2018) sind die am Ende des Fragenprogramms verorteten Literalitätstests. Erst mit ihrer Analyse ist die *Bestimmung einer Größenordnung funktionaler Analphabeten* möglich. Für den Vergleich mit der leo. – Level-One-Studie (LEO 2010) und damit einer Darstellung der Entwicklung der Größenordnung ist eine direkte Vergleichbarkeit sowohl der Instrumente in den *Paper & Pencil*-Testheften (Rätsel- und Alphahefte) als auch ihrer Bewertung während der Übertragung in den Datensatz für weitere Analysen unerlässlich (s. Kapitel 2.2). Für LEO 2018 sind die Qualitätsansprüche im Vergleich zur Vorgängerstudie generell gestiegen, denn LEO 2018 hat im Gegensatz zu LEO 2010 keinen Pilotcharakter. Für die Bewertung der Testaufgaben und die parallele Übertragung der Ergebnisse in eine Datenbank und schließlich in den LEO-Datensatz waren sowohl die Erfahrungen aus LEO 2010 einzubeziehen als auch die Kriterien als Mindestmaß zu beachten, die im *Programme of International Assessment of Adult Competencies* (PIAAC) berücksichtigt wurden.

Wichtig für eine gelungene Bewertung der Testaufgaben ist eine hohe Übereinstimmung der *Rater*. Die Voraussetzung dafür ist ein einheitliches Verständnis für richtige und falsche Testantworten. Um ein gemeinsames Grundverständnis herzustellen, wurde im März 2018 eine zwei-tägige Schulung aller an der Auswertung beteiligten Personen bei Kantar Public gemeinsam mit der Universität Hamburg durchgeführt. Auf Grundlage des Regelwerks von LEO 2010 sowie der Pretest-Bewertungen der Literalitätstests wurde für LEO 2018 das Regelwerk in Abstimmung mit der Universität Hamburg überarbeitet. Die an die Schulung anschließende Einarbeitungsphase der *Rater* wurde zeitlich nicht beschränkt, sondern dauerte an, bis eine zufriedenstellende Übereinstimmung der Bewertungen der einzelnen *Rater* gegeben war. Zu diesem Zeitpunkt ging auch der Diskussionsbedarf über die Entscheidungen der Bewertungen in Einzelfällen stark zurück.

Die Auswertung der Testaufgaben wurde – analog zu LEO 2010 ex post bei Kantar Public durchgeführt. Die Arbeitsplätze der vier zu diesem Zweck eingestellten *Rater* lagen räumlich direkt neben denen der beiden Projektleitungen von Kantar Public. Die räumliche Nähe stellte eine durchgängig einfache Kommunikation zwischen *Ratern* und Projektleitung sicher. Der gesamte Bewertungsprozess der Testaufgaben wurde in drei Tranchen aufgesetzt, um die mit der Universität Hamburg vereinbarten (Zwischen-)Datenlieferungstermine gewährleisten zu können. Der erste Datensatz mit 1.000 Fällen für den 13. April 2018 geplant und realisiert. Entsprechend wurde Mitte März die *Raterschulung* durchgeführt und eine Bewertung der bis zu diesem Zeitpunkt vorliegenden Testhefte (Tranche 1) vorgenommen. Die zweite Datenlieferung mit einem Umfang von 5.000 Fällen erfolgte Mitte August 2018 und die Enddatenlieferung am 9. Oktober 2018. Bis dahin wurden nahezu alle Testhefte abschließend bewertet. Einzelne nach diesem Termin bei uns eingegangene Testhefte wurden nachträglich bewertet und der Universität Hamburg übergeben.

Im Folgenden wird der gesamte Bewertungsprozess genauer beschrieben, und zwar in chronologischer Abfolge. In einem ersten Schritt werden Informationen über die Schulungs- und Einarbeitungsphase (Kapitel 4.2.1) bereitgestellt, danach solche über die Haupt-Bewertungsphase (Kapitel 4.2.2). Dabei werden unter Qualitätsaspekten die prozentuale Übereinstimmung zwischen den *Ratern* und über alle Testitems hinweg sowie die *Interrater-Reliabilität* berichtet. Um die Vergleichbarkeit zur Vorgängerstudie genau beurteilen zu können, wurden rund fünf Prozent der Testhefte aus LEO 2010 auch durch das *Editingteam* der aktuellen Studie bewertet. Informationen dazu sind in Kapitel 4.3 dokumentiert.

4.2.1 Schulung der Rater

Für die Bewertung der Testaufgaben wurden vier Personen als *Rater* angeworben. Das Team setzte sich aus drei Studierenden und einer Rentnerin zusammen. Alle vier Personen sind deutschsprachig aufgewachsen und haben Abitur gemacht. Zwei der Rater studierten deutsche Linguistik. Die Rentnerin brachte Kenntnisse über ältere Schriftbilder ein. Alle vier Rater waren zeitlich so flexibel, dass eine Bewertung der Literalitätstests in Tranchen möglich war. Mit der Einstellung von vier Ratern war ausreichend personelle Flexibilität gegeben und ein intensives Monitoring der Arbeiten dieser Personen durch eine intern benannte Moderatorin sowie die Projektleitung möglich. Die Moderatorin, namentlich Elisabeth Baier, war die erste inhaltliche Ansprechpartnerin für die Bewertung der Testaufgaben. Sie leitete die Rater-Briefings im gesamten Prozess und war die letzte Entscheidungsinstanz für zweifelhafte Einzelfälle. Dafür hielt sie Rücksprache mit der Projektleitung bei Kantar Public oder in Einzelfällen mit dem Projektteam der Universität Hamburg.

Über die vier Rater hinaus wurde unsere Auszubildende (im ersten Lehrjahr) in die Raterschulung sowie die Einarbeitungsphase einbezogen. Hierdurch wurde sichergestellt, dass im Falle sich verzögernder Feldarbeiten eine Bewertung der Literalitätstests auch nach Ende der befristeten Rateranstellungen möglich gewesen wäre und dass zeitliche Engpässe innerhalb der drei Bewertungstranchen gegebenenfalls hätten ausgeglichen werden können. Die Testaufgaben wurden also von einem Team von fünf Ratern ausgewertet.

Zentrales Ziel der zwei-tätigen Raterschulung war, ein einheitliches Verständnis für die Bewertung der Testaufgaben zu erzeugen. Erreicht werden sollte eine Übereinstimmung der Bewertungen der einzelnen Rater in mindestens 95 Prozent der Fälle über alle Items hinweg. Analog zu PIAAC (Tamassia et al., 2016) sollte darüber hinaus im gesamten Team für jedes einzelne Item eine zumindest 85-prozentige Übereinstimmung der Bewertungen vorliegen. Im Rahmen der Schulung machten sich die Rater mit der Datenbank vertraut, die zur Eingabe der Bewertungsergebnisse eigens für LEO 2018 von Kantar Public programmiert wurde.³⁵

Für die Raterschulung erstellte die Universität Hamburg ausführliche Schulungsunterlagen, die anschließend zu einem Regelwerk zusammengestellt und jedem Rater als Nachschlagewerk zur Verfügung gestellt wurden. Grundlage dafür war das Regelwerk von LEO 2010. Darüber erstellte das LEO-Team der Universität Hamburg auf Basis der Auswertungen der 384 Pretestfälle Musterlösungen. Diese Fälle wurden im Rahmen der Schulung zur Anschauung bereitgestellt, dienten als Vergleichsfälle und wurden daher in die Datenbank eingelesen. Im ersten Teil der Schulung wurde die LEO-Grundbildungsstudie vorgestellt. Anschließend wurde in die Bewertung der Testaufgaben und die Datenbank eingewiesen. Nach ausführlicher Erläuterung der Regeln anhand von zahlreichen Richtig- bzw. Falsch-Beispielen folgten drei Schulungsrunden. Die Dauer einer solchen Runde wurde im Vorfeld unterschätzt. Es zeigte sich, dass größerer Diskussionsbedarf seitens der Rater bestand als erwartet.

Eine Schulungsrunde begann mit der zufälligen Auswahl einer bestimmten Anzahl von Rätsel- und zugehörigen Alphaheften. Jedes Heft sollte dabei von jedem Rater sowie der Moderatorin bewertet werden, um uneinheitliche Bewertungen im Anschluss möglichst schnell und diskursiv anpassen zu können. Dafür wurde jeder Fall einzeln betrachtet und jeder codierte Wert im gesamten Team mit der Musterlösung abgeglichen. Jede Uneinigkeit wurde einzeln diskutiert und dabei ein gemeinsames, regelbasiertes Verständnis für *richtig* und *falsch* entwickelt. Die drei Schulungsrunden werden im Folgenden stichpunktartig beschrieben.

³⁵ Dafür danken wir dem Datenbankentwickler Alfred Weber.

Runde 1 (Tag 1, 13:15 – 15:00 Uhr)

- Auswerten von fünf Fällen je Rater, sodass jeder Rater die gleichen Fälle bearbeitete. Die Fälle wurden so ausgewählt, dass mindestens jedes Alphaheft einmal enthalten war.
- Benennung und Diskussion der Schwierigkeiten bei der Bewertung
- Diskussion von Einzelfällen, bei denen sich das Raterteam uneinig war und gegebenenfalls auch Korrektur der Musterlösung
- Das *Fleiss' Kappa*³⁶ betrug zu diesem Zeitpunkt immerhin 0,959³⁷.
- Zusätzlich wurde die jeweilige prozentuale Übereinstimmung mit der korrigierten Musterlösung berechnet (Tabelle 8). Diese Übereinstimmung liegt in der ersten Runde insgesamt zwischen 73 und 100 Prozent.

Tabelle 8: Prozentuale Übereinstimmung mit der korrigierten Musterlösung³⁸ in Schulungsrunde 1

	Anteilswerte in %			
	Rätselheft (n=5)	Joschi (n=1)	Helga (n=1)	Leschek (n=2)
Rater 1*	96	100	100	97
Rater 2	99	96	100	100
Rater 3	91	87	100	77
Rater 4	96	96	100	100
Rater 5	88	91	100	73
Rater 11**	94	100	100	100
Min. – Max.	88 - 99	87 - 100	100 - 100	73 - 100

* Rater 1: Auszubildende im 1. Lehrjahr

** Rater 11: Moderatorin

Runde 2 (Tag 1, 15:15 – 17:00 Uhr)

- Auswerten von jeweils acht weiteren Rätselheften je Rater, sodass jeder Rater die gleichen Fälle bearbeitete. Die Fälle wurden so ausgewählt, dass mindestens jedes Alphaheft zweimal enthalten war.
- Benennung und Diskussion der Schwierigkeiten bei der Bewertung
- Diskussion und gegebenenfalls auch Korrektur von Fällen, bei denen sich das Raterteam uneinig war
- Das *Fleiss' Kappa* beträgt zu diesem Zeitpunkt 0,947 (s. Fußnote 36).
- Die prozentuale Übereinstimmung mit der korrigierten Musterlösung liegt in der zweiten Runde insgesamt zwischen 74 und 100 Prozent (Tabelle 9).

Tabelle 9: Prozentuale Übereinstimmung mit der korrigierten Musterlösung in Schulungsrunde 2

	Anteilswerte in %			
	Rätselheft (n=8)	Joschi (n=2)	Helga n=3	Leschek n=2
Rater 1*	87	96	82	100
Rater 2	95	93	85	97
Rater 3	92	100	91	97
Rater 4	96	93	94	100
Rater 5	91	74	91	97
Rater 11**	94	89	77	93
Min. – Max.	87 - 96	74 - 100	77 - 94	93 - 100

* Rater 1: Auszubildende im 1. Lehrjahr

** Rater 11: Moderatorin

³⁶ Kennwert zur Beurteilung der Übereinstimmung mehrerer Rater unter Berücksichtigung der zufällig zu erwartenden Übereinstimmung

³⁷ Ab einem *Fleiss' Kappa*-Wert von 0,8 ist die Übereinstimmung zwischen Ratern als nahezu vollkommen einzuordnen (Fleiss & Cohen, 1973).

³⁸ Im laufenden Schulungsprozess wurde in Einzelfällen gemeinsam mit dem Projektteam der Universität Hamburg entschieden von der Musterlösung abzuweichen. In diesen Fällen wurde die Musterlösung nachträglich korrigiert und die Berechnung der Kennwerte erneut durchgeführt.

Runde 3 (Besprechung Tag 2, 10:00 – 12:00 Uhr)

- Da die Diskussion über die Nicht-Übereinstimmungen am ersten Schultag viel länger als erwartet dauerte, wurde das Auswerten weiterer 20 Hefte auf die darauffolgende Woche verschoben. Jede Raterin und jeder Rater sollte innerhalb einer Woche bei freier Zeiteinteilung 20 weitere Fälle auswerten, sodass jede Raterin und jeder Rater die gleichen Fälle bearbeitete. Die Fälle wurden so ausgewählt, dass mindestens jedes Alphaheft fünfmal enthalten war.
- Benennung und Diskussion der Schwierigkeiten bei der Bewertung
- Diskussion und gegebenenfalls auch Korrektur von Fällen, bei denen sich das Editingteam uneinig war
- Besprechung und Nachschulung von Aufgaben, die mehrfach uneinheitliches Verständnis für die Bewertung erzeugten aus der ersten und zweiten Schulungsrunde
- Das Fleiss' Kappa beträgt zu diesem Zeitpunkt 0,948 (s. Fußnote 37).
- Die prozentuale Übereinstimmung mit der korrigierten Musterlösung liegt in der dritten Runde insgesamt zwischen 91 und 99 Prozent (Tabelle 10).

Tabelle 10: Prozentuale Übereinstimmung mit der korrigierten Musterlösung in Schulungsrunde 3

	Anteilswerte in %			
	Rätselheft (n=20)	Joschi (n=5)	Helga (n=5)	Leschek (n=5)
Rater 1*	95	95	96	99
Rater 2	98	96	96	99
Rater 3	96	94	95	95
Rater 4	98	97	95	99
Rater 5	96	91	94	99
Rater 11**	98	97	95	99
Min. – Max.	95 - 98	91 - 97	94 - 96	95 - 99

* Rater 1: Auszubildende im 1. Lehrjahr
 ** Rater 11: Moderatorin

4.2.2 Bewertung der Testaufgaben

Die ersten 791 Fälle wurden doppelt bearbeitet, also jeweils von zwei unabhängigen Ratern ausgewertet, um eine möglichst hohe Übereinstimmung zwischen den Ratern im Rahmen der Einarbeitungsphase zu erzielen. Damit wurden elf Prozent der insgesamt erhobenen 7.192 Fälle bereits zu Beginn doppelt editiert. Die doppelt bearbeiteten Fälle wurden anhand der sogenannten *Rücklaufpaginiernummern*, die jedes Rätselheft eindeutig kennzeichnen³⁹, ausgewählt.

Bei der Sortierung und gegebenenfalls Zuordnung der Rätsel- und Alphahefte zueinander wurden zugleich auf den Paginiernummern basierende Nummernkreise festgelegt. Diese Nummernkreise wurden für eine klare Aufgabenteilung zwischen den Ratern in einem randomisierten Verfahren den Ratern zugeordnet. Mit dieser Zufallsauswahl wurde keine exakte Gleichverteilung von Rätsel- und Alphaheften auf die Rater angestrebt. Insofern variiert die bearbeitete Anzahl der Testhefte getrennt nach Rater. Ebenfalls per Zufallsverfahren wurden einzelne Hefte ausgewählt, um 20 Prozent aller Testhefte doppelt zu vercoden. Dieses Ziel wurde am 11. September 2018, mit 1.835 doppelt vercodeten Heften von insgesamt 8.634 Heften (21%) erreicht. Tabelle 11 zeigt die Verteilungen der einzeln und doppelt codierten Testhefte getrennt nach Rater und Testheftart.

³⁹ Zugehörige Alphahefte erhielten die Rücklaufpaginiernummer des Rätselhefts.

Tabelle 11: Anzahl der einfach und doppelt ausgewerteten Hefte pro Rater

	Rätselhefte			Alphahefte		
	einfach	doppelt	Summe	einfach	doppelt	Summe
Rater 1*	18	132	150	1	50	51
Rater 2	1.165	818	1.983	284	209	493
Rater 3	1.603	671	2.274	342	180	522
Rater 4	1.069	665	1.734	260	144	404
Rater 5	1.771	620	2.391	307	145	452

* Rater 1: Auszubildende im 1. Lehrjahr

Für die Auswertung der Testaufgaben wurde von Kantar Public eine Datenbank konzipiert und programmiert, in die wöchentlich die neuen Interviewfälle eingelesen wurden. Jede Datenbanknutzerin und jeder -nutzer hatte einen personalisierten, passwortgeschützten Zugang und nutzungsprofilspezifische Rechte. Es gibt grundsätzlich zwei unterschiedliche Profile. Mit jedem Profil konnte in der Datenbank die *Editingmaske* aufgerufen werden. Sowohl die Moderatorin als auch die Projektleitung erhielten erweiterte Rechte a) zum Aufruf verschiedener Masken zum Monitoring, b) zum Aufruf aller Fälle in allen Bearbeitungsstadien, c) zum Aufruf von Fällen, die auf Wiedervorlage gesetzt sind, d) zur Bearbeitung aller Fälle und e) zum Auslesen und Einlesen von Fällen.

Die Rater konnten bei Aufruf eines Falls zunächst die Auswertung der Interviewerinnen und Interviewer in den Rätselheften und der ersten Aufgabe der Alphahefte sehen und bei Bedarf korrigieren. Bei den Aufgaben zwei bis vier in den Alphaheften wurde durch die Rater die erste Codierung vorgenommen, weil hier eine Bewertung seitens der Interviewerinnen und Interviewer im Rahmen des CAPI-Interviews nicht verlangt war.

Wenn ein Fall in der Datenbank aufgerufen wird, wird eine Kopie des Falls angelegt. Jede Raterin und jeder Rater kann einen Fall nur einmal aufrufen, es sei denn, er oder sie nutzt die Möglichkeit der Wiedervorlage. Die Wiedervorlagefunktion wurde eingefügt, um den Raterinnen und Ratern eine Besprechung nicht eindeutiger Bewertungsvorgaben einzuräumen und hier nach Klärung den richtigen Code einzugeben. Bei Fällen, die doppelt bearbeitet werden sollten, wird beim zweiten Aufruf des Falls durch eine zweite Raterin bzw. einen Rater wiederum eine Kopie analog zum ersten Aufruf angelegt.⁴⁰ Hierdurch wird sichergestellt, dass im zweiten Codierungsschritt ausschließlich die von der Interviewerin bzw. vom Interviewer eingegebenen Informationen vorliegen, nicht aber die Bewertung durch die erste Raterin bzw. den ersten Rater. Daraus ergeben sich in der Regel für jedes Heft und darin für jedes Item zwei dichotome Einträge (1 „richtig“ oder 2 „falsch“ bzw. bei den Multiple Choice-Items: 1 „ausgewählt“ oder 0 „nicht ausgewählt“). Wenn eine Aufgabe nicht bearbeitet wurde, wurde diese Information als dritter möglicher Wert aufgenommen (9 „nicht bearbeitet“). Nicht-Übereinstimmungen innerhalb des Teams wurden nach der zweiten Codierung für Moderatorin und Projektleitung erkennbar. Die Moderatorin entschied dann in der Regel gemeinsam mit den Raterinnen und Ratern in einer weiteren Kopie des Falls über ungleiche Werte von Rater 1 und 2 und schloss damit die Bearbeitung des Falls ab.

Das Ausmaß der Übereinstimmung der Rater wird anhand von *Cohen's Kappa (Interrater-Reliabilität)*⁴¹ und der prozentualen Übereinstimmung berichtet (s.o.). Aufgrund eines einfachen und schnell möglichen Ausleseprozesses der Informationen aus der Datenbank erfolgte ein kontinuierliches und zeitnahes Monitoring des Editingprozesses. Cohen's Kappa wurde für jedes Rater-Paar berechnet. Die entsprechenden Werte der zehn

⁴⁰ Der Zeitpunkt des Speicherns eines bearbeiteten Falls wird in der Datenbank festgehalten.

⁴¹ *Cohen's Kappa (Interrater-Reliabilität)* ist ein Kennwert zur Erfassung der Übereinstimmung zweier Beurteiler unter Berücksichtigung der zufällig zu erwartenden Übereinstimmung.

Rater-Paare reichen von 0,915 bis 0,961 (s. Fußnote 36). Zusätzlich wurde die jeweilige prozentuale Übereinstimmung unter den einzelnen Ratern berechnet. Diese Übereinstimmung lag zu Feldbeginn insgesamt zwischen 93 und 100 Prozent (Tabelle 12).

Tabelle 12: Prozentuale Übereinstimmung der Rater zum Feldstart

	Anteilswerte in %			
	Rätselheft (n=791)	Joschi (n=77)	Helga (n=91)	Leschek (n=82)
Rater 1*Rater 2	97	96	97	96
Rater 1*Rater 3	98	96	99	100
Rater 1*Rater 4	96	96	93	98
Rater 1*Rater 5	97	100	95	n.v.
Rater 2*Rater 3	98	98	95	100
Rater 2*Rater 4	97	98	97	99
Rater 2*Rater 5	97	98	97	99
Rater 3*Rater 4	98	98	99	99
Rater 3*Rater 5	96	97	97	100
Rater 4*Rater 5	99	93	100	100
Min. – Max.	96 - 99	93 - 100	93 - 100	96 - 100

Dieses Ergebnis konnte gehalten werden, sodass die Übereinstimmung auch zu Feldende, unter Einbezug aller doppelt editierten Fälle, noch im selben Intervall liegt (Tabelle 13).

Tabelle 13: Prozentuale Übereinstimmung der Rater gesamt

	Anteilswerte in %			
	Rätselheft (n=1.453)	Joschi (n=123)	Helga (n=141)	Leschek (n=118)
Rater 1*Rater 2	97	96	97	98
Rater 1*Rater 3	98	96	99	100
Rater 1*Rater 4	96	96	93	98
Rater 1*Rater 5	97	100	95	n.v.
Rater 2*Rater 3	98	98	96	100
Rater 2*Rater 4	98	99	98	99
Rater 2*Rater 5	97	99	98	99
Rater 3*Rater 4	98	97	99	99
Rater 3*Rater 5	96	97	96	100
Rater 4*Rater 5	97	97	98	100
Min. – Max.	96 - 98	96 - 100	93 - 99	98 - 100

Regelmäßig wurden Briefings einerseits zur Klärung von Rückfragen zu Einzelfällen (Stichwort *Wiedervorlage*, s.o.) und zur item-bezogenen Nachschulung durchgeführt. Eine uneinheitliche Bewertung auf Ebene der Einzelitems ließ sich anhand einer Liste der Items mit ihrer jeweiligen Quote der Abweichung identifizieren. Im Falle von Quoten der Abweichung, die auf eine vergleichsweise geringe Übereinstimmung unter den Raterinnen und Ratern schließen ließen, wurden die Betreffenden in den entsprechenden Testaufgaben durch die Moderatorin im *Briefing* nachgeschult.

Die prozentuale Übereinstimmung insgesamt über alle Items beträgt 97 Prozent (Tabelle 14). Die jeweiligen prozentualen Übereinstimmungen für jedes Item übersteigen über alle 1.453 Fälle hinweg den Richtwert von 85 Prozent (Tabelle 15, 16, 17, 18). Damit wurden die Ziele und die Richtwerte aus der PIAAC-Studie (s.o.) erreicht. Bei der Betrachtung der Quoten der Abweichung jedes einzelnen Items ist zu beachten, dass den

Rätsel- und Alphaheften nicht dieselben Fallzahlen zugrunde liegen. Weil bei der Berechnung der prozentualen Übereinstimmung der Alphahefte deutlich weniger Fälle einbezogen werden, hat ein abweichendes Item in einem Fall einen vergleichsweise großen Einfluss auf die Quote der Abweichung des jeweiligen Items insgesamt.⁴²

Tabelle 14: Zusammenfassung der prozentualen Übereinstimmung (Feldende)

	Anteilswerte in %			
	Phase 1 (n=787)	Phase 2 (n=345)	Phase 3 (n=322)	Gesamt (n=1.453)
Rätselheft: Min. – Max.	93 - 100	92 - 100	93 - 100	93 - 100
Joschi: Min. – Max.	86 - 97	82 - 100	100 - 100	89 - 100
Helga: Min. – Max.	87 - 100	84 - 100	99 - 100	89 - 100
Leschek: Min. – Max.	96 - 100	100 - 100	100 - 100	97 - 100
Gesamt (alle Hefte)	97	97	99	97

Tabelle 15: Prozentuale Übereinstimmung für jedes Item im Rätselheft

	Anteilswerte in %			
	Phase 1 (n=787)	Phase 2 (n=345)	Phase 3 (n=322)	Gesamt (n=1.453)
T71	99	99	98	99
T721	98	98	98	98
T722	97	95	97	96
T723	96	95	96	96
T724	96	96	95	96
T725	93	97	95	94
T726	98	98	98	98
T75A1	100	100	100	100
T75A2	100	100	100	100
T75A3	100	100	100	100
T75A4	100	100	100	100
T75A5	100	100	100	100
T75B	98	98	100	99
T76A	93	95	93	94
T76B	93	92	93	93

⁴² Einzig in Phase 2 im Alphaheft „Joschi“ unterschreitet ein Wert die Zielgröße, allerdings liegt hier aufgrund der geringen Fallzahl zwischen 85 Prozent (A544b) und 82 Prozent (A543b) nur eine Nicht-Übereinstimmung vor.

Tabelle 16: Prozentuale Übereinstimmung für jedes Item im Alphaheft „Joschi“

	Anteilswerte in %			
	Phase 1 (n=77)	Phase 2 (n=33)	Phase 3 (n=13)	Gesamt (n=123)
A511	97	100	100	100
A512	95	100	100	98
A513	96	100	100	99
A514	97	100	100	100
A515	97	100	100	100
A516	97	100	100	100
A517	97	100	100	100
A518	96	100	100	99
A521	97	100	100	100
A522	95	100	100	98
A523	95	97	100	97
A531	94	94	100	96
A532	96	97	100	98
A533	94	91	100	95
A534	96	97	100	98
A541a	97	100	100	100
A542a	97	100	100	100
A543a	95	100	100	98
A544a	95	100	100	98
A541b	86	91	100	89
A542b	91	100	100	96
A543b	90	82	100	89
A544b	92	85	100	93

Tabelle 17: Prozentuale Übereinstimmung für jedes Item im Alphaheft „Helga“

	Anteilswerte in %			
	Phase 1 (n=92)	Phase 2 (n=31)	Phase 3 (n=18)	Gesamt (n=141)
B51	99	97	100	98
B521	97	100	100	98
B522	97	97	100	97
B531a	99	100	100	99
B532a	96	90	100	95
B533a	96	90	100	95
B534a	98	100	100	99
B535a	97	100	100	98
B531b	96	94	99	94
B532b	87	94	99	89
B533b	90	84	100	90
B534b	93	100	100	96
B535b	92	97	100	94
B541	100	94	100	99
B542	99	100	100	99
B543	100	100	100	100
B544	99	100	100	99
B545	99	100	100	99
B546	98	100	100	99
B547	100	97	100	99
B548	100	100	100	100
B549	99	100	100	99

Tabelle 18: Prozentuale Übereinstimmung für jedes Item im Alphaheft „Leschek“

	Anteilswerte in %			
	Phase 1 (n=82)	Phase 2 (n=22)	Phase 3 (n=14)	Gesamt (n=118)
C511	100	100	100	100
C512	100	100	100	100
C513	100	100	100	100
C514	100	100	100	100
C521	100	100	100	100
C522	100	100	100	100
C523	96	100	100	97
C524	99	100	100	98
C525	100	100	100	100
C526	100	100	100	99
C527	100	100	100	100
C53	98	100	100	98
C541	100	100	100	100
C542	96	100	100	97
C543	98	100	100	98

4.3 Erneute Bewertung ausgewählter Testhefte aus LEO 2010 durch das Raterteam von LEO 2018

Um die Ergebnisse von LEO 2018 mit LEO 2010 genau vergleichen zu können, beauftragte die Universität Hamburg Kantar Public zusätzlich mit einer Neuvercodung ausgewählter Testhefte aus LEO 2010 durch das in LEO 2018 eingesetzte Raterteam. Vorgelegt wurden insgesamt 409 Fälle mit insgesamt 488 Testheften.

Für diesen Arbeitsschritt wurde das Raterteam über die Unterschiede in den Testheften von LEO 2010 und LEO 2018 informiert. Die insgesamt vier Änderungen wurden genau besprochen, auch hinsichtlich des Vorgehens in der Bewertung. Die ansonsten gleich gebliebenen Testaufgaben sollten nach den Regeln bewertet werden, die auch im Rahmen von LEO 2018 angewendet wurden und auf denen von LEO 2010 aufbauen. Entsprechend wurde auch hier die Datenbank genutzt, die auch für das Editing der LEO 2018-Testhefte zum Einsatz kam (s. Kapitel 4.2). Sowohl die Rätsel- als auch die Alphahefte wurden gleichmäßig auf die vier Raterinnen und Rater verteilt. Jede Raterin und jeder Rater codierte entsprechend etwa 100 Rätselhefte und je sieben zugehörige Alphahefte aus LEO 2010.

Zur Überprüfung des Grades der Übereinstimmung der Originalcodierungen von LEO 2010 und der neu durchgeführten wurde auch die *Interrater-Reliabilität* berechnet (s.o.).⁴³ Das Ausmaß der Übereinstimmung wird anhand des *Cohen's Kappa* (s.o.) und der prozentualen Übereinstimmung berichtet. Das Cohen's Kappa wurde für jede Raterin und jeden Rater berechnet. Die Werte der vier Raterinnen und Rater reichen von 0,857 bis 0,927 (s. Fußnote 36). Zusätzlich wurde die jeweilige prozentuale Übereinstimmung zwischen der Codierung der Jahre 2010 und 2018 berechnet. Wie Tabelle 19 zeigt, liegen die Werte der Übereinstimmung zwischen 94 und 99 Prozent pro Heft.

Tabelle 19: Prozentuale Übereinstimmung zwischen den Codierungen 2010 und 2018

	Anteilswerte in %			
	Rätselheft (n=409)	Joschi (n=27)	Helga (n=25)	Leschek (n=27)
Rater 2	97	95	98	99
Rater 3	96	98	97	94
Rater 4	96	96	98	99
Rater 5	96	98	98	98
Min. – Max.	96 – 97	95 – 98	97 - 98	94 - 99

Wie Tabelle 20 zeigt, liegt die Übereinstimmung über alle Raterinnen und Rater hinweg insgesamt zwischen 84 und 100 Prozent. Bei der Interpretation der Werte ist zu beachten, dass den Rätselheften eine höhere Fallzahl zugrunde liegt als jeweils den Alphaheften. Folglich hat bei den Alphaheften ein abweichendes Item in einem Fall einen vergleichsweise großen Einfluss auf die Abweichungsquote des jeweiligen Items.

⁴³ Für die Berechnungen wurden die Codierungen aus dem Jahr 2010 und dem Jahr 2018 dichotomisiert. Dieser Schritt war notwendig, weil die Regel zur Nutzung des dritten Codes („9“) seit dem Jahr 2010 verändert wurde. Die LEO 2010-Werte wurden nach denselben Regeln aufbereitet, wie diejenigen in LEO 2018. Folglich waren nach diesem Schritt die Codierungen beider Jahre vergleichbar und konnten für Berechnungen der Reliabilität verwendet werden.

Tabelle 20: Prozentuale Übereinstimmung für jedes Item über alle Rater

Item	Rätselheft	Alphaheft "Joschi"		Alphaheft "Helga"		Alphaheft "Leschek"	
	prozentuale Übereinstimmung	Item	prozentuale Übereinstimmung	Item	prozentuale Übereinstimmung	Item	prozentuale Übereinstimmung
T71	99	A511	100	B51	100	C511	100
T721	97	A512	100	B521	100	C512	100
T722	94	A513	100	B522	100	C513	100
T723	95	A514	100	B531a	100	C514	100
T724	93	A515	100	B532a	100	C521	100
T725	93	A516	100	B533a	100	C522	100
T726	97	A517	100	B534a	96	C523	100
T75A1	99	A518	93	B535a	92	C524	100
T75A2	100	A521	100	B531b	84	C525	96
T75A3	98	A522	100	B532b	96	C526	96
T75A4	99	A523	100	B533b	88	C527	96
T75A5	99	A531	100	B534b	100	C53	100
T75B	95	A532	93	B535b	100	C541	100
T76A	94	A533	93	B541	100	C542	93
T76B	92	A534	93	B542	100	C543	85
		A541a	100	B543	100		
		A542a	100	B544	100		
		A543a	100	B545	100		
		A544a	96	B546	100		
		A541b	85	B547	100		
		A542b	93	B548	100		
		A543b	93	B549	96		
		A544b	93				
Min.- Max.	92 – 100		85 – 100		84 – 100		85 – 100

5 Testteilnahme und Testverweigerungen

Ziel der LEO-Grundbildungsstudie (LEO 2018) ist es, eine empirisch fundierte Schätzgröße für die Zahl von Analphabeten in Deutschland im Vergleich zur Vorgängerstudie der leo. – Level-One-Studie (LEO 2010) zu ermitteln. Methodische Grundlage ist eine repräsentative Bevölkerungsstichprobe, in der standardisierte Tests der Lese- und Schreibfähigkeiten durchgeführt werden. Der Anspruch der Untersuchung ist, quantitative Aussagen über die literale Kompetenz der erwachsenen Bevölkerung auf Basis systematischer Literalitätstests zu machen, wobei das Hauptinteresse dem unteren Kompetenzbereich gilt (*level one*). Die Teilnahme an der Befragung und den Tests beruht auf dem Prinzip der Freiwilligkeit. Es liegt auf der Hand, dass nicht alle zufällig ausgewählten Personen bereit sind, sich an einer solchen Untersuchung zu beteiligen.

Testverweigerung ist ein generelles Problem jeder survey-gestützten empirischen Forschung. Für ein Vorhaben wie der LEO-Grundbildungsstudie liegt es jedoch noch in besonderer, verschärfter Weise vor. Testverweigerung kann hier eine „besondere“ Strategie der Zielpersonen sein, ihre Lese- und Schreibschwäche nicht offenzulegen. Damit sind die Testverweigerungen als eine kritische Größe im Messverfahren zu betrachten und stellen im Extremfall den Wert oder zumindest die Verallgemeinerbarkeit der Ergebnisse in Frage. Wenn es beispielsweise so wäre, dass nur Personen mit guten Lese- und Schreibkenntnissen zur Teilnahme an den Tests bereit gewesen wären, wären Aussagen über die Größenordnung des Analphabetismus in Deutschland auf der gegebenen Datenbasis kaum möglich.

Im Folgenden wird beschrieben, welche methodischen Verfahren im Rahmen von LEO 2018 angewendet wurden, um zu vertretbaren Lösungen dieses Problems zu gelangen. Einschätzungen der Ergebnisse auf Basis einer ersten Datenanalyse werden vorgestellt. Es wird sicherlich eine Aufgabe der Universität Hamburg sein, im Rahmen ihrer eigenen Analyse und Interpretation der Daten auch das Problem der Testverweigerungen zu berücksichtigen.

Mit Blick auf die Beschreibung von Ausfällen der Teilnahme sind Personen zu unterscheiden, die a) zur Teilnahme an der gesamten Befragung nicht bereit waren, also überhaupt nicht Bestandteil der realisierten Nettostichprobe sind (*Unit-Nonresponse*), und b) diejenigen, die zwar an der Befragung teilgenommen haben, für die jedoch gar keine oder nur unvollständige Antworten zu den Testaufgaben vorliegen (*Item-Nonresponse*).

5.1 Nichtteilnahme an der Befragung (*Unit-Nonresponse*)

Idealerweise sind Ausfälle in der Befragung, die es aufgrund von Nicht-Antreffbarkeit, Verweigerung oder anderen Teilnahmebarrieren gibt, in Hinblick auf die zentralen Untersuchungsthemen zufallsverteilt (*missing at random*) und keine Beeinträchtigung für die Qualität der Stichprobe. Inwieweit diese Bedingung erfüllt ist, ist im konkreten Fall meist schwer feststellbar, weil über Ausfälle keine Information vorliegen. Meist wird angenommen, dass die *Missing-at-random*-Bedingung umso besser erfüllt ist, je höher die Ausschöpfung der Stichprobe ist. Dies muss jedoch nicht der Fall sein. Dies mag ein Beispiel illustrieren. Im Beispiel wird davon ausgegangen, dass die Zielgruppe der Analphabeten zehn Prozent der Bevölkerung umfasst. Wenn zuträfe, dass diese Zielgruppe sich jeglicher Teilnahme an einer Befragung entzieht, dann wäre sie in der realisierten Bevölkerungsstichprobe nicht enthalten – unabhängig davon, ob die Stichprobenausschöpfung 30, 60 oder 90 Prozent beträgt. Worauf es vielmehr ankommt, ist eine möglichst geringe *Selektivität* des Teilnahmeverhaltens hinsichtlich der Kernthemen der Untersuchung.

Der Frage, wie wahrscheinlich es ist, dass Personen mit geringer Schriftsprachbefähigung sich an einer repräsentativen Bevölkerungsbefragung beteiligen, wurde bereits im Rahmen von LEO 2010 genauer nachgegangen. Die dortigen Ergebnisse zusammengefasst, wurde bis zur Beweisführung durch LEO 2010 (in der Literatur) angenommen, dass die Zielgruppe ein hohes Maß an Abwehr gegenüber jeglichen Befragungen und Tests zeigt.⁴⁴ Die Gültigkeit solcher Aussagen wurde (und wird noch immer) als von den zugrunde liegenden Bedingungen und methodischen Verfahren (*Settings*) abhängig bewertet. In praktischer Sicht kommt es darauf an, die Befragungsmethodik so zu gestalten, dass für Personen mit Lese- und Schreibschwächen möglichst keine oder nur niedrige Teilnahmebarrieren gegeben sind. Maßgebliche Qualitätskriterien für eine Stichprobe, die den Anteil von Personen mit geringer Schriftsprachkompetenz in der Bevölkerung korrekt wiedergeben soll, sind a) eine *Zufallsauswahl* der Stichprobe (s. Kapitel 2.5) und b) die *Literalitätsneutralität* der Kontaktaufnahme und der Ansprache der Zielpersonen und Interviewmethoden (s. Kapitel 2.3). Beide Anforderungen sind in der vorliegenden Untersuchung im gegebenen Rahmen bestmöglich erfüllt, wie in den entsprechenden Kapiteln zuvor beschrieben ist.

Zur Prüfung der *Literalitätsneutralität* der realisierten Stichprobe sind vor allem zwei Punkte gegeben:

- a) die Struktur der Stichprobe nach Bildungsniveau und
- b) das faktische Vorhandensein von Personen mit Lese- und Schreibschwäche unter den Befragungspersonen, zunächst belegt durch die Tatsache, dass Alphahefte im Rahmen der Befragung bearbeitet wurden.

Beide Punkte werden im Folgenden geprüft.

5.1.1 Die Bildungsstruktur der Stichprobe

In repräsentativen Umfragen ist in der Regel von einer gewissen Positivselektivität der Befragungsteilnehmenden nach (schulischem) Bildungsniveau auszugehen. Erfahrungsgemäß beteiligen sich Personen mit höherem Bildungsgrad eher an Umfragen und sind in der realisierten Stichprobe also überrepräsentiert, während

⁴⁴ Auch die Studie *AlphaPanel*, die damals zeitlich parallel zu LEO 2010 durchgeführt wurde, zeigte ein anderes Ergebnis. Hier konnten in der Regel die Kursleitungen als *Gate Keeper* identifiziert werden. Sofern die Kursleitungen eine Vorstellung der Interviewerinnen und Interviewer in ihren Kursen zur schriftsprachlichen Grundbildung zuließen, lag die Teilnahmebereitschaft der Kursteilnehmenden im Rahmen der Befragung vergleichsweise hoch (Bilger 2011; Rosenblatt & Bilger 2011).

umgekehrt Personen niedrigen Bildungsgrads weniger zur Teilnahme bereit sind und im Ergebnis mehr oder weniger unterrepräsentiert sind. Diesem Umstand konnte in LEO 2010 durch eine damals wirksame – und heute eher als unzulässig erachtete – Art der Incentivierung der Probanden nach *Bauchgefühl* der Interviewerinnen und Interviewer entgegengewirkt werden. Ein monetäres Incentive wurde auch im Rahmen der LEO 2018-Erhebung entrichtet, allerdings nicht nach *Bauchgefühl* der Interviewerinnen und Interviewer an einen Teil der Teilnehmenden, sondern an alle. Hierdurch konnte zwar ein gewisser Interviewereffekt in der Zielpersonenauswahl ausgeschlossen werden, der sonst in repräsentativen Erhebungen gegebene Bildungsbias wurde allerdings bereits in der Pretesterhebung beobachtet (s. Kapitel 2.3.3). Eine gewisse Strukturverzerrung kann ohne weiteres per Gewichtung ausgeglichen werden, um repräsentative Aussagen zu machen. Voraussetzung ist dafür, dass für alle gewichtungsrelevanten Gruppen ausreichend Informationen vorliegen, worauf in Kapitel 5.1.2 eingegangen wird. Hier wird gezeigt, in welchem Ausmaß eine Bildungsverzerrung in der LEO 2018-Basisstichprobe überhaupt gegeben ist.

Tabelle 27 zeigt verschiedene Repräsentativstichproben bezogen auf die Gruppe der 18- bis 64-Jährigen, getrennt nach höchstem schulischem Abschluss im Vergleich. Die Bildungsstrukturen, an denen diejenigen der LEO-Basiserhebung jeweils zu messen sind, sind dem Mikrozensus entnommen – für LEO 2010 war das der Mikrozensus 2008 (MZ) und für LEO 2018 der MZ 2016. Wie beschrieben, erfolgte die LEO-Basiserhebung des Jahres 2010 im Verein mit der Erhebung des Adult Education Survey (AES 2010). Für den direkten Vergleich der LEO-Grundbildungsstudie (LEO 2018) wurden hier zudem die Strukturen des AES 2016⁴⁵ dargestellt und entsprechend die für die Gewichtung herangezogenen Strukturinformationen aus dem MZ 2014.

Tabelle 27: Ausgewählte Repräsentativstichproben im Vergleich hinsichtlich ihrer (schulischen) Bildungsstruktur, gewichtet und ungewichtet

Zielpersonen: 18- bis 64-Jährige der jeweiligen Repräsentativstichprobe	MZ 2008 (zum Vgl.)	LEO/AES 2010 (n=7.035)		MZ 2014 (zum Vgl.)	Anteilswerte in % AES 2016 (n=7.102)		MZ 2016 (zum Vgl.)	LEO 2018 (n=6.681)	
		ungewichtet	gewichtet		ungewichtet	gewichtet		ungewichtet	gewichtet
Niedrige Bildung	34,8	35,0	34,1	29,6	27,3	29,7	28,2	23,9	28,2
Mittlere Bildung	33,3	33,5	32,7	34,3	35,3	34,2	33,3	32,6	33,4
Höhere Bildung	29,7	30,2	31,6	34,9	36,6	34,9	36,9	41,3	37,0
Keine Angabe	0,8	0,6	0,7	0,2	0,3	0,3	0,4	0,2	0,2
Noch Schüler	1,3	0,7	0,9	1,0	0,5	1,0	1,2	2,1	1,3
Gesamt	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0

MZ: Mikrozensus

LEO 2010 bildet in den ungewichteten Daten – begründet durch die Incentivierung der Teilnehmenden – die Sollstrukturen aus dem Mikrozensus 2008 nahezu richtig ab. In der ungewichteten LEO 2018-Stichprobe zeigt sich erwartungsgemäß eine etwas andere Struktur. Personen mit niedrigerem schulischem Bildungsniveau sind, gemessen am MZ 2016, etwas unter- und Personen mit höherem Bildungsniveau etwas überrepräsentiert (Differenz jeweils vier Prozentpunkte). Die Verzerrung der Bildungsstrukturen in LEO 2018 ist etwas ausgeprägter als im AES 2016. Dort liegt eine Abweichung von der Sollstruktur des MZ 2014 von je zwei Prozentpunkten vor. Analog zum AES 2016 wird eine Anpassung der Strukturen per Gewichtung vorgenommen (s. Kapitel 6).

Aus Tabelle 27 wird zudem ein anderes Phänomen deutlich. Beim Vergleich der Anteilswerte von Personen mit niedrigem schulischem Bildungsabschluss wird aus dem Mikrozensus deutlich, dass sich der Anteilswert seit dem Jahr 2008 (35%) deutlich verringert hat (2014: 30%; 2016: 28%). Insofern ist hier eine Anhebung des Bildungsniveaus in der erwachsenen Bevölkerung zu erkennen. Wir gehen dabei nicht davon aus, dass es

⁴⁵ Das sind die zum Zeitpunkt der Berichterstattung aktuellsten Informationen des deutschen AES.

sich hierbei um eine individuelle Anhebung des Bildungsniveaus handelt, sondern interpretieren das Ergebnis vielmehr als einen Kohorteneffekt.

5.1.2 Die Bearbeitung der Alphahefte im Vergleich von LEO 2010 und LEO 2018

Wie viele Befragungspersonen im Rahmen der LEO-Erhebungen die zehn (LEO 2010) bzw. elf Testaufgaben (LEO 2018) so bearbeiteten, dass im CAPI ein Score unterhalb des Schwellenwerts von acht bzw. neun Punkten erzielt wurde (s. Kapitel 2.2.2) und somit ein Alphaheft zur Bearbeitung vorgelegt wurde, dient als zweites überprüfbares Kriterium der *Stichprobenqualität* bzw. der *Literalitätsneutralität*. In Tabelle 28 sind für die LEO-Erhebungen 2010 und 2018 jeweils die Basis- und Zusatzstichprobe bezüglich der ungewichteten Bildungsstruktur ausgewiesen (jeweils erste Spalte als Spaltenprozent). Wie im letzten Abschnitt beschrieben, ist der geringe Anteilswert von Personen mit niedrigem Schulbildungsniveau in der Basiserhebung des Jahres 2018 zwar auf eine Verzerrung der Stichprobe, vor allem aber auf eine veränderte Zusammensetzung in der Bevölkerung zurückzuführen. In der jeweils zweiten Spalte der Tabelle 28 wird der Anteilswert von Personen, die ein Alphaheft in der jeweiligen Bildungsgruppe während der Erhebung bearbeiteten ausgewiesen. Bei der Interpretation der Ergebnisse ist zu beachten, dass aufgrund der in LEO 2018 ergänzten Testaufgabe keine exakte Vergleichbarkeit der Ergebnisse der Jahre 2010 und 2018 gegeben ist (s. Kapitel 2.2.2).

In LEO 2018 haben mit 22 Prozent etwas mehr Personen als noch in LEO 2010 (21%) ein Alphaheft zur Bearbeitung erhalten. Dies dürfte vor allem auf die neue Vorgabe, Rätselheftaufgaben in der konkreten Bewertungsphase während des Interviews im Zweifel als „falsch“ zu bewerten, zurückzuführen sein (s. Kapitel 2.4.3). In den drei Bildungsgruppen finden sich in LEO 2018 entsprechend durchgängig leicht höhere Anteilswerte von Personen, die ein Alphaheft zur Bearbeitung erhielten als in den Vergleichsgruppen in LEO 2010. Unter Personen mit mittlerem Bildungsniveau liegt die Differenz mit sechs Prozentpunkten allerdings überdurchschnittlich hoch (LEO 2010: 15%; LEO 2018: 21%), gefolgt von derjenigen unter Personen mit niedrigem Bildungsabschluss (vier Prozentpunkte; LEO 2010: 37%; LEO 2018: 41%). Wenngleich die Ergebnisse nicht exakt vergleichbar sind, zeigen sie, dass grundsätzlich davon auszugehen ist, dass auch in LEO 2018 kein robuster Widerstand von Personen mit zumindest gewissen Schriftsprachbeeinträchtigungen hinsichtlich ihres generellen Teilnahmeverhaltens bestand. Wir schließen auf dieser Basis entsprechend einen systematischen Fehler aus.

Weder die Überprüfung der Bildungsstrukturen noch die Tatsache, dass Alphahefte getrennt nach Bildungsgruppen etwa analog zu LEO 2010 bearbeitet wurden, verweisen auf einen empirischen Grund zur Annahme einer systematischen Nichtteilnahme an der LEO-Grundbildungsstudie. Dennoch bleibt – wie in jeder umfragegestützten Forschung – eine gewisse Unsicherheit über mögliche Selektivitätseffekte eines *Unit-Non-response*. Gleichwohl gilt bis zur Vorlage gegenteiliger Indizien die Arbeitshypothese, dass die Ergebnisse der Stichprobenerhebung für die Grundgesamtheit verallgemeinerbar sind.

Tabelle 28: Ungewichtete Anteilswerte von Personen, die im Rahmen der LEO-Erhebungen 2010 und 2018 ein Alphaheft bearbeiteten getrennt nach schulischer Bildung

Basis: 18- bis 64-Jährige	ungewichtete Basiserhebungen im Vergleich			
	LEO/AES 2010 (n=7.035)		LEO 2018 (n=6.681)	
	Spalten- prozentwerte	<i>darunter: Anteil von Personen, die ein AH bearbeiteten</i>	Spalten- prozentwerte	<i>darunter: Anteil von Personen, die ein AH bearbeiteten</i>
Niedrige Bildung	35,0	36,9	23,9	41,2
Mittlere Bildung	33,5	15,0	32,6	21,2
Höhere Bildung	30,2	8,7	41,3	10,8
Keine Angabe	0,6	23,3	0,2	23,1
Noch Schüler	0,7	9,8	2,1	19,3
Gesamt	100,0	20,8	100,0	21,7
	ungewichtete Zusatzerhebungen im Vergleich			
	LEO/AES 2010 (n=1.401)		LEO 2018 (n=511)	
	Spalten- prozentwerte	<i>darunter: Anteil von Personen, die ein AH bearbeiteten</i>	Spalten- prozentwerte	<i>darunter: Anteil von Personen, die ein AH bearbeiteten</i>
Niedrige Bildung	99,6	38,6	99,4	48,8
Mittlere Bildung	-	-	-	-
Höhere Bildung	0,1	100,0	-	-
Keine Angabe	0,4	60,0	-	-
Noch Schüler	-	-	0,6	100,0
Gesamt	100,0	38,8	100,0	49,1
	Gewichtete Gesamtstichproben im Vergleich			
	LEO/AES 2010 (n=8.436)		LEO 2018 (n=7.192)	
	Spalten- prozentwerte	<i>darunter: Anteil von Personen, die ein AH bearbeiteten</i>	Spalten- prozentwerte	<i>darunter: Anteil von Personen, die ein AH bearbeiteten</i>
Niedrige Bildung	34,6	37,4	27,7	41,2
Mittlere Bildung	32,3	14,8	33,5	20,1
Höhere Bildung	31,4	9,0	37,2	13,8
Keine Angabe	0,8	30,2	0,2	24,1
Noch Schüler	0,9	11,0	1,4	23,3
Gesamt	100,0	20,9	100,0	23,7

AH: Alphaheft

5.2 Nichtteilnahme an Testaufgaben (*Item-Nonresponse*)

Eine Person, die an der freiwilligen Befragung teilnimmt, kann die Bearbeitung aller oder einzelner Testaufgaben zur Lese- und Schreibkompetenz verweigern. Die dahinterstehenden Gründe einer (Einzel-)Testverweigerung können ebenso wie die Verweigerung einer Teilnahme an der Befragung insgesamt sehr vielfältig sein. Sie können von Unlust, über zeitliche Begrenzung, Störung durch eine dritte Person über eine konkrete Test-Aversion oder gar eine Vermeidungsstrategie aufgrund schriftsprachlicher Beeinträchtigungen reichen. Vor allem die zuletzt genannten Gründe sind mit Blick auf die LEO-Grundbildungsstudie als nicht qualitätsneutral zu bewerten.

Methodisch kann ein solches Problem als eine Form von *Item-Nonresponse* angesehen werden, also als eine fehlende Information für eine bestimmte Frage oder Variable. Dies gilt selbst dann, wenn die Nichtbeantwortung sich auf den gesamten Block von Testfragen bezieht. Entscheidend ist, dass die Person an dem vorausgehenden Interview teilgenommen hat und über sie dadurch Informationen aus dem Interview vorliegen. Diese können genutzt werden, um Determinanten des item-bezogenen Nonresponse zu analysieren und gegebenenfalls fehlende Werte über Schätzverfahren zu imputieren.

Im Rahmen der Interviewsituation hat das Bemühen um möglichst vollständige Informationen im Rahmen eines Interviews durch die Interviewerinnen und Interviewer Vorrang. In der LEO-Grundbildungsstudie gehört hierzu auch die möglichst vollständige Testbearbeitung. Grundsätzlich gilt, dass auch die Bearbeitung der Testaufgaben freiwillig ist. Die Befragungsperson muss also die Möglichkeit haben, an jedem Punkt des Ablaufs ihre Bearbeitung zu beenden. Gleichwohl wird im Rahmen der LEO-Grundbildungsstudie einem sofortigen Testabbruch entgegengewirkt (s. Kapitel 2.3). Im Falle von Verweigerungen, eine Testaufgabe vollständig zu bearbeiten, werden darüber hinaus jeweils die Gründe dafür durch die Interviewerinnen und Interviewer erfasst, sofern diesen von der Zielperson benannt wurden (s. Kapitel 2.2.2). Konkret wurde folgendes Vorgehen umgesetzt.

- Der Überleitungstext vom Befragungsteil zum Testteil des Interviews ist knapp und einfach formuliert. Er vermeidet Begriffe, die Ängste oder Abwehr auslösen könnten, wie „Test“ oder „Aufgaben“ oder „Lesen und Schreiben“ (Stichwort: Testneutralität, s. Kapitel 2.3.1 und 2.3.2).
- Die erste Testaufgabe (**T1**: Sehtest) dient in einfacher Form als beispielhaft für die folgenden Tests. Sie dient auch dazu, mit den Testaufgaben vertraut zu werden und führt sowohl das Lesen (**T1a**) als auch das Schreiben (**T1b**) in einfacher Form ein. Erst danach folgen die für LEO 2018 relevanten Aufgaben (**T2**, **T5** und **T6**).
- Nach jeder (Teil-)Aufgabe kann die Befragungsperson die Testbearbeitung beenden. Informationen der jeweils vorangegangenen Testaufgabe bleiben im Falle eines Testabbruchs im CAPI erhalten und gehen in den Datensatz ein. Demnach liegen auch im Falle eines Abbruchs zumindest einige Informationen vor.
- Sofern eine Testaufgabe nicht oder nicht vollständig beantwortet wurde, wird die Interviewerin bzw. der Interviewer an der entsprechenden Stelle im CAPI gebeten, die Abbruchgründe der Befragungsperson zu erfassen. Diese Zusatzinformationen sind Bestandteil des Datensatzes. Sie können im Falle eines Abbruchs hilfreich sein, die vorliegenden Testergebnisse richtig einzuschätzen.

Item-Nonresponse kann demnach datenseitig in zwei Formen auftreten:

1. Eine Testaufgabe wurde gar nicht oder nur unvollständig bearbeitet. Sofern dies zutrifft, sollten a) Gründe durch die Interviewerin bzw. den Interviewer eingetragen sein und b) angegeben sein, ob die Zielperson die Bearbeitung abgebrochen oder fortgesetzt hat.

2. Eine Testaufgabe wurde nicht bearbeitet, weil die Befragungsperson die Testbearbeitung bereits davor abgebrochen hat. Die Basis (Fallzahl) für die Bearbeitung der Aufgabe ist entsprechend im Datensatz vermindert.

Tabelle 29 zeigt für die Aufgaben der ersten Teststufe (*Rätselheft*), wie viele Befragungspersonen die jeweiligen Testaufgaben bearbeitet haben und wie viele Fälle von *Item-Nonresponse* und von Abbrüchen es bei den einzelnen Aufgaben gab.

Tabelle 29: Item-Nonresponse und Abbrüche auf der ersten Teststufe

Testaufgabe im Rätselheft		Basis gesamt (n)	Item-Nonresponse							
Variable	Label		nicht (voll) bearbeitet (n)	% von Basis	Begründung liegt vor ¹⁾		Abbruch		Abbrüche kumulativ	
				ja (n)	% von Basis	nein (n)		% von Basis	(n)	% von Basis
T1a	Sehtest 1	7.192	143	2,0	84	1,2	59	0,8	56	0,8
T1b	Sehtest 2	7.136	83	1,2	33	0,5	50	0,2	70	1,0
T2	Fernweh	7.122	320	4,5	194	2,7	126	0,6	110	1,5
T5a	Rahmenwörter 1	7.082	290	4,1	208	2,9	82	0,4	141	2,0
T5b	Rahmenwörter 2	7.051	388	5,5	193	2,7	195	0,2	152	2,1
T6a	Richtig schreiben 1	7.040	150	2,1	94	1,3	56	0,3	171	2,4
T6b	Richtig schreiben 2	7.021	82	1,2	33	0,5	49	0,1	177	2,5
T6c	Richtig schreiben 3	7.015	64	0,9	35	0,5	29	0,1	186	2,6
Summe (n)		56.659	1.520	2,7	874	1,5	646	0,3	1.063	

¹⁾ Die Berechnungen wurden auf Basis der geschlossenen Fragen, ob ein Grund genannt wurde vorgenommen: alle Gründe aus **T1a21, T1a22 und T1a23** nicht genannt, alle Gründe aus **T1a22, T1a23** nicht genannt; sowie value "nein, ohne Begründung aus den Variablen **T1b2, T2a2, T5a2, T5b2, T6a2, T6b2 und T6c2**

Die erste Aufgabe (**T1a**) wurde von allen Befragungspersonen bearbeitet, also von 7.192 Personen. Die letzte Aufgabe (**T6c**) wurde immerhin von noch 7.015 Personen bearbeitet (97,5%). Davon ausgehend, dass alle 7.192 Befragungspersonen alle elf (Teil-)Aufgaben in LEO 2018 hätten theoretisch bearbeiten sollen, hätten insgesamt 57.536 Testbearbeitungen vorliegen müssen. Real wurden 56.659 Testaufgaben zur Bearbeitung vorgelegt (98,5%), weil in einigen Fällen davor ein Abbruch im Rätselheft erfolgte.

Unter diesen zur Bearbeitung vorgelegten Testaufgaben wurden 1.520 gar nicht oder unvollständig bearbeitet. Das sind bezogen auf alle zur Bearbeitung vorgelegten Tests 2,7 Prozent. Dieser Wert liegt etwa so hoch wie in LEO 2010 mit 3,5 Prozent. Für mehr als die Hälfte dieser Item-Nonresponses in LEO 2018 liegen Gründe oder Erläuterung der (teilweisen) Nichtbearbeitung vor. Mit jedem achten Fall (n=186; 12,2 %) einer (teilweisen) Nichtbearbeitung einer Testaufgabe geht zugleich auch ein Abbruch der Rätselheftbearbeitung einher. Wenngleich dieser Wert vergleichsweise hoch klingt, so liegt er bezogen auf die Personenebene und damit die Grundgesamtheit der Befragungspersonen mit 2,6 Prozent eher niedrig und etwa so hoch wie in LEO 2010 mit 3,0 Prozent. D.h. umgekehrt, dass immerhin 97,4 Prozent der befragten Erwachsenen in LEO 2018 alle (Teil-)Aufgaben der Teststufe 1, also im Rätselheft, vollständig beantwortet haben.

Insgesamt wurden 2,7 Prozent der Testaufgaben nicht oder nur unvollständig bearbeitet. Überdurchschnittlich häufig tritt eine unvollständige Bearbeitung bei den drei Rätselheftaufgaben **T5b** „Rahmenwörter 2“ (5,5%) **T2** „Fernweh“ (4,5%) und **T5a** „Rahmenwörter 1“ (4,1%) auf. Immerhin zwei Drittel (65,7%) der unvollständigen Bearbeitungen sind auf diese drei Aufgaben zurückzuführen.

Die Verteilung der Abbrüche im Rätselheft bei Einzelaufgaben ist damit korreliert, aber nicht identisch. Überdurchschnittliche Werte liegen für die drei (Teil-)Aufgaben **T1a** „Sehtest 1“ (0,8%), **T2** „Fernweh“ (0,6%) und

T5a „Rahmenwörter 1“ (0,4%) vor. D.h., die Abbruchgründe dürften nicht nur mit der Schwierigkeit der Aufgaben, sondern auch mit situationsspezifischen Gründen verkoppelt sein. Dabei sind zwei Muster zu erkennen.

1. Mit zunehmendem Fortschritt im Rätselheft fallen die Abbruchanteile niedriger aus. Die Abbrüche im Rätselheft erfolgen mehrheitlich während der Bearbeitung der beiden ersten Testaufgaben **T1** und **T2** (59,1%), also zu einem früheren Zeitpunkt im Rätselhefttestverlauf. Dabei kommt Personen, die bereits in Aufgabe **T1a** (30,1% der Abbrechenden) die weitere Testbearbeitung verweigerten, hinsichtlich der Literalitätsneutralität des Abbruchs besondere Bedeutung zu. Für diese Personen liegen (nahezu) keine Informationen hinsichtlich ihrer literalen Kompetenz vor (s. Kapitel 5.3.1).
2. Wenn eine Aufgabe begonnen und beendet wurde, werden vergleichsweise häufig auch die damit zusammenhängenden, nachgelagerten Teilaufgaben bearbeitet.

Früh Abbrechende (T1-T2) sind von den im Testverlauf *später Abbrechenden* dadurch zu unterscheiden, dass für letztere immerhin eine Schätzung basierend auf zwei Testaufgaben (drei Teilaufgaben) erfolgen kann. Bezogen auf alle Erwachsenen liegt der Anteilswert der *früh Abbrechenden* in LEO 2018 mit 1,5 Prozent genauso hoch wie in LEO 2010 (1,5%).

■
Insgesamt wurde für alle Befragungspersonen, die das Rätselheft ohne Testabbruch bearbeiteten (n=7.006) nach der Bewertung der Rätselheftaufgaben durch die Interviewerinnen und Interviewer und ihre Eingabe ins CAPI, das Rätselheftergebnis ermittelt und mit dem vorgegebenen Schwellenwert verglichen. Sofern die Befragungsperson den Schwellenwert von 9 unterschritt, sollte sie ein weiteres Testheft – eines von drei Alphaheften – zur weiteren Bearbeitung erhalten. Das war bei 24,2 Prozent (n=1.698) der theoretisch testbereiten Personen der Fall. Der im Pretest ermittelte direkte Vergleichswert lag mit 24 Prozent genauso hoch (s. Kapitel 2.2.2).

Um auf der zweiten Teststufe eine größere Zahl von Testaufgaben einbeziehen zu können, wurden drei Splits mit jeweils vier Testaufgaben gebildet, die per Zufallsauswahl auf die Befragungspersonen verteilt wurden. Diese je vier Testaufgaben je Split wurden wiederum in den drei Alphaheften A „Joschi“, B „Helga“ und C „Leschek“ von den ausgewählten Personen bearbeitet. Die Zufallsauswahl führte zu einer recht guten Verteilung der drei Alphahefte. 31,9 Prozent erhielten das Alphaheft A „Joschi“, 34,4 Prozent das Alphaheft B „Helga“ und 33,6 Prozent das Alphaheft C „Leschek“. Analog zu Tabelle 29 wird in Tabelle 30 gezeigt, in welchem Umfang es bei diesen zusätzlichen Testaufgaben zu Item-Nonresponses und Abbrüchen kam.

Unter den im Alphaheft bearbeiteten Aufgaben lag der Anteilswert unvollständiger Bearbeitungen mit 14,6 Prozent (Tabelle 30) deutlich höher als unter den Rätselheftaufgaben (2,7%, s.o. Tabelle 29). Der Wert liegt zudem in LEO 2018 etwas höher als noch in LEO 2010 (11,3%). Je Alphaheft wurden verschiedene hohe Anteilswerte Abbrechender ermittelt. Während der Bearbeitung des Testhefts *Joschi* brachen 15,5 Prozent die Alphaheftbearbeitung ab, während des Testhefts *Leschek* 12,8 Prozent und während des Testhefts *Helga* 8,4 Prozent. Auch hier liegen die ermittelten Werte deutlich über demjenigen auf Teststufe 1 von 2,6 Prozent (s.o.). Die höheren Raten sowohl von Item-Nonresponse als auch von Abbrechenden dürften auf die ausgewählte Zielpersonengruppe mit Alphaheftbearbeitung zurückzuführen sein. Sie hat per definitionem höhere Schriftsprachschwierigkeiten als die Gruppe der 18- bis 64-Jährigen insgesamt und dürfte zudem gegen Ende der durchschnittlich rund 50-minütig andauernden Interview- bzw. Testsituation (s. Kapitel 3.1) mit Ermüdungserscheinungen gegebenenfalls gepaart mit Konzentrationsschwierigkeiten (*cognitive load*) gekämpft haben.

Tabelle 30: Item-Nonresponse und Abbrüche auf der zweiten Teststufe (eines von drei Alphaheften)

Basis: alle mit Unterschreitung des Scores, die testbereit sind (n=1.698) Testaufgabe in Alphaheften A bis Varia Label		Basis gesamt (n)		nicht (voll) bearbeitet (n) % von Basis		Item-Nonresponse		Abbruch (n) % von Basis		Abbrüche kumulativ (n) % von Basis	
						Begründung liegt vor					
						ja	nein				
						(n)	% von Basis	(n)	% von Basis		
Alphaheft A "Joschi"											
A1	Hafenbilder	543	99	18,2	41	7,6	58	58	10,7	58	10,7
A2	Short News	485	24	4,9	16	3,3	8	6	1,2	64	11,8
A3	Pflanzen besprechen	479	51	10,6	29	6,1	22	20	4,2	84	15,5
A4	Überschriften zuordnen	459	40	8,7	19	4,1	21	-	-	84	15,5
Summe "Joschi" (n)		1.966	214	10,9	105	5,3	109	84	4,3	84	15,5
Alphaheft B "Helga"											
B1	Überschriften zuordnen	584	48	8,2	16	2,7	32	36	6,2	36	6,2
B2	Köchin und Küken	548	18	3,3	7	1,3	11	3	0,5	39	6,7
B3	Lageristin Helga	545	67	12,3	27	5,0	40	10	1,8	49	8,4
B4	Bundeswehr	535	19	3,6	7	1,3	12	-	-	49	8,4
Summe "Helga" (n)		2.212	152	6,9	57	2,6	95	49	2,2	49	8,4
Alphaheft B "Leschek"											
C1	Suchbilder	571	61	10,7	30	5,3	31	49	8,6	49	8,6
C2	Leschek-Martina	522	48	9,2	23	4,4	25	13	2,5	62	10,9
C3	Staubraub	509	391	76,8	330	64,8	61	11	2,2	73	12,8
C4	Leschek	498	48	9,6	26	5,2	22	-	-	73	12,8
Summe "Leschek" (n)		2.100	548	26,1	409	19,5	139	73	3,5	73	12,8
Gesamt											
Summe Gesamt (n)		6.278	914	14,6	571	9,1	343	206	3,3	206	12,1
Summe Gesamt in %		100,0	14,6		9,1		5,5	3,3		3,3	

Auch auf Teststufe 2 lassen sich einzelne Testaufgaben identifizieren, die hinsichtlich der Raten an Item-Nonresponse und Testabbrüchen hervorstechen. Fehlende oder unvollständige Testantworten gibt es bei je einer Aufgabe pro Alphaheft: **C3** „Staubraub“ (76,8%) und **A1** „Hafenbilder“ (18,2%) und **B3** „Lageristin Helga“ (12,3%). Die Abbrüche finden, ähnlich wie auf Teststufe 1, am häufigsten (69%) innerhalb der ersten Testaufgabe je Alphaheft statt.

Für alle Testteilnehmenden auf der zweiten Teststufe liegen Testergebnisse der ersten Stufe vor. Insofern ist für eine personenbezogene Analyse das Problem der Testverweigerungen auf Stufe 2 ein kaum gravierendes Problem.

5.3 Unvollständige Testbearbeitung: Gründe und Typen

Wenn eine Befragungsperson eine Testaufgabe nicht oder nur unvollständig bearbeitet hat, kann dies unterschiedliche Gründe haben. Gründe können einerseits Scheitern an einer Testaufgabe anzeigen. In dem Fall wäre die unvollständige Bearbeitung in der Analyse wie eine falsche Testantwort zu werten. Sie können aber andererseits auch, insbesondere wenn es mit einem Testabbruch einhergeht, lediglich ein Ende der Kooperationsbereitschaft anzeigen. Letzteres kann z.B. aus Zeitgründen oder aus zunehmenden Bedenken gegen die Tests im Allgemeinen oder Bedenken gegen eine spezifische Testaufgabe erfolgen. In diesem Fall wäre der *Item-Nonresponse* als neutraler Ausfall zu werten und nicht als Scheitern an der Testaufgabe (s. auch Kapitel 5.2).

Fehlende Angaben können ohne weitere Zusatzinformationen nicht einer der beiden Deutungen zugeordnet werden. Das Erhebungsinstrument wurde daher für LEO 2010 und analog für LEO 2018 so gestaltet, dass im Fall fehlender oder unvollständiger Testantworten die Interviewerin oder der Interviewer angehalten ist, eventuelle Begründungen seitens der Befragungsperson in kurzen Stichworten zu dokumentieren (s. Kapitel 2.2.2). Diese Texte sind Bestandteil des Datensatzes und können für weitere Analysen genutzt werden. In der Analyse können mit den *Gründen* nicht nur für die Bewertung einzelner Testaufgaben im Verein mit den individuellen Schwierigkeiten der Befragungspersonen hilfreiche Zusatzinformationen herangezogen werden, sondern auch zur Erklärung von Abbrüchen.

Zunächst wird ein Überblick über die Art der Aussagen, die aus diesem Textmaterial vorliegen gegeben (Kapitel 5.3.1). Anschließend werden die Abbruchinformationen auf Personenebene in eine Typologie umgesetzt, mit Hilfe derer eine personengruppenbezogene Analyse der Abbruchgründe vereinfacht werden kann (Kapitel 5.3.2).

Die ausführlichsten Zusatzinformationen liegen für die Einstiegsfrage in die Testaufgaben im Rätselheft vor (Frage **T1a**: „Sehtest“). Zur Bearbeitung der Aufgabe erhält die Interviewerin bzw. der Interviewer drei Nachfragen, die nicht an die Zielperson zu richten sind.

- A) Zunächst wird gefragt, ob die Zielperson den Aufgabentext entweder richtig (1), nur teilweise (2), mit einzelnen falschen Worten (3) oder gar nicht (4) vorgelesen habe (**T1a1**).
- B) Sofern die Frage mit einer der Antwortkategorie „nicht (ganz) richtig“ beantwortet wurde, wird eine Nachfrage im CAPI zum Grund der unvollständigen Bearbeitung eingeblendet. Die Gründe werden, sofern nicht eine der beiden geschlossenen Antwortvorgaben „Augenprobleme/Brille“ und „Kann nicht gut genug lesen“ zutrifft, als offene Texteingabe durch die Interviewerin bzw. den Interviewer eingegeben (**T1a21** bis **T1a24**, **T1a2x**).
- C) Für dieselbe Personengruppe soll die Interviewerin bzw. der Interviewer in einer weiteren Frage beantworten, ob die Zielperson „zur weiteren Teilnahme“ bereit ist.

Diese drei Nachfragen wurden durchgängig nach jeder einzelnen Rätselheftaufgabe vorgesehen, je im Wortlaut an die entsprechende Aufgabe angepasst (genauer s. Anhang A1.1). Entsprechend liegt für jede Aufgabe die Information vor, ob eine Aufgabe (un-)vollständig bearbeitet wurde (Item-Nonresponse), gegebenenfalls ein Grund für die unvollständige Bearbeitung und darüber hinaus die Information, ob an dieser Stelle das Interview abgebrochen oder fortgesetzt wurde.

5.3.1 Gründe der unvollständigen Testbearbeitung

Insgesamt wurden auf der ersten LEO-Teststufe – also im Rahmen des Rätselhefts – 1.520 Einzelitems von insgesamt 912 Personen unvollständig bearbeitet (Tabelle 31). Bezogen auf alle bearbeiteten Aufgaben sind demnach 2,7 Prozent als Item-Nonresponse einzustufen. Diese Tatsache gibt noch keine Auskunft darüber, ob die jeweils unvollständig oder gar nicht bearbeiteten Aufgaben a) aus literalitätsneutralen Gründen, wie z.B. Zeitmangel oder b) aus nicht-neutralen Gründen, wie z.B. eine Leserechtschreibschwäche (LRS) erfolgte. Dies kann mit Hilfe der oben genannten Gründe für nicht vollständig bearbeitete LEO-Aufgaben sowohl auf der ersten als auch der zweiten Teststufe ein Stück weit differenziert werden. Es ist darauf hinzuweisen, dass eine vollständig bearbeitete Aufgabe inhaltlich offenlässt, ob diese richtig oder falsch gelöst wurde. Wenn ein Testabbruch erfolgte, dann nach einer unvollständigen oder gar nicht bearbeiteten Aufgabe (n=186). Bezogen auf alle Befragungspersonen haben 2,6 Prozent den LEO-Test im Rahmen des Rätselhefts abgebrochen, darunter 1,5 Prozent zu einem früheren Zeitpunkt (in Aufgaben **T1** oder **T2**).

Tabelle 31: Bearbeitete Rätselheftaufgaben insgesamt, als Item-Nonresponse und getrennt nach Abbruch bzw. weiterer Testbearbeitung

Basis: alle 18- bis 64-Jährigen (n=7.192) bzw. die von ihnen im Rätselheft bearbeiteten Aufgaben (n=56.659)		Auf Teststufe 1 (Rätselheft) bearbeitete Aufgaben ...					
		alle	unvollständig/nicht bearbeitet	darunter: unvollständig/nicht bearbeitet	darunter: mit direkt darauf folgendem Testabbruch	darunter: früh (T1/2)	darunter: spät (T5/6)
Aufgaben	(n)	56.659	1.520	186	110	76	1.334
	% von alle	100,0	2,7	0,3	0,2	0,1	2,4
Personen	(n)	7.192	912	186	110	76	780
	% von alle	100,0	12,7	2,6	1,5	1,1	10,8
durchschnittliche bearbeitete Aufgaben pro Person	(m)	7,9	1,7	1,0	1,0	1,0	1,7

Die in LEO 2018 offen genannten Gründe wurden in insgesamt zehn Kategorien zusammengefasst (Tabelle 32). Diese Kategorien wurden bereits in LEO 2010 entwickelt. Die Variablen sind im Datensatz enthalten und sind in chronologischer Datenansicht jeweils hinter den Variablen mit den offenen Textnennungen abgelegt. Namentlich wurde die Bezeichnung der offenen Variablen (z.B. **T1a2x**) übernommen und um „_code“ ergänzt (z.B. **T1a2x_code**).

Die Kategorien der Gründe für eine unvollständige Bearbeitung der Testaufgaben sind in Tabelle 32 dargestellt. Eine getrennte Betrachtung ist für Gründe a) unvollständig bearbeiteter Aufgaben, auf die unmittelbar ein Testabbruch erfolgte (Spalten 2 bis 4) und b) solchen, die keine Barriere für die weitere Testbearbeitung darstellten (letzte Spalte) gegeben. Die Aufgaben für die Gründe auf deren (Nicht-)Benennung ein unmittelbarer Testabbruch folgt, können im Rätselheft chronologisch früher (Aufgaben T1 bis T2) oder später (Aufgaben T5 bis T6) bearbeitet worden sein und werden in der Tabelle zusätzlich getrennt ausgewiesen.

Für 45 Prozent der unvollständig bearbeiteten Aufgaben wurden keine näheren Gründe benannt bzw. keine Angabe in der entsprechenden offenen Variablen gemacht. Dabei finden sich diese Nichtbenennungen leicht häufiger unter den unvollständig bearbeiteten Aufgaben von Personen, die danach die Testbearbeitung fortsetzten (45%) als unter solchen, die diese direkt anschließend abbrachen (43%). Unter den verbleibenden Gründen finden sich am häufigsten die beiden Gründe „kann (die Wörter) nicht (gut) lesen“ (20%) und „Hat die Frage/Aufgabe nicht verstanden“ (16%). Der letztgenannte Grund wird häufiger von Personen genannt, die danach die Testbearbeitung fortsetzten (17%) als von solchen, die direkt danach den Test abbrachen (7%).

Mit „Zeitmangel und andere situative Gründe“ (5%) sowie „Kein Interesse, keine Lust und andere persönliche Gründe“ (5%) werden direkt vor dem Testabbruch überdurchschnittlich häufig zwei eher literalitätsneutrale Gründe genannt. Unter den Personen, die das Rätselheft abbrachen, finden sich unter den zuletzt erfragten Gründen der unvollständigen Bearbeitung allerdings auch solche, die nicht als literalitätsneutrale Gründe einzustufen sind. Für den direkten Vergleich werden analog zu LEO 2010 folgende sechs Gründe auch in LEO 2018 darunter verstanden.

- „Kann (die Wörter) nicht (gut) lesen“ (18%),
- „Kann nicht (gut) schreiben“ (3%),

- „Kann nicht genug Deutsch/Sprachprobleme“ (4%),
- „Hat Frage/Aufgabe nicht verstanden“ (7%)
- „Aufgabenspezifische Schwierigkeiten“ (7%) und
- „Sehprobleme“ (2%).

Tabelle 32: Gründe für Item-Nonresponse auf der ersten Teststufe (Rätselheft)

Basis: Personen, die wenigstens eine Rätselheftaufgabe höchstens unvollständig bearbeitet haben (n=912)	Gründe für Item-Nonresponse ⁴⁾				
	darunter:				
	gesamt (n=912)	direkt vor Testabbruch ⁵⁾ (n=186)	darunter: früher Abbruch (T1-T2) (n=110)	später Abbruch (T5-T6) (n=76)	vor weiterer Testbear- beitung (n=780)
Gründe der unvollständigen Aufgabenbearbeitung					
Kann (die Wörter) nicht (gut) lesen	19,8	18,3	12,7	26,3	20,0
Kann nicht (gut) schreiben	2,0	3,2	1,8	5,3	1,9
Kann nicht genug Deutsch/Sprachprobleme	1,4	3,8	6,4	0,0	1,1
Hat Frage/Aufgabe nicht verstanden	15,6	7,0	7,3	6,6	16,8
Aufgabenspezifische Schwierigkeiten	10,5	7,0	5,5	9,2	11,0
Sehprobleme	1,5	2,2	3,6	0,0	1,4
Generelle Testablehnung	1,4	4,8	7,3	1,3	0,9
Zeitmangel und andere situative Gründe	1,1	5,4	6,4	3,9	0,4
Kein Interesse, keine Lust und andere persönliche Gründe	1,9	5,4	6,4	3,9	1,4
Verweigerung ohne nähere Begründung	43,9	39,2	40,0	38,2	44,6
keine Angabe	0,8	3,8	2,7	5,3	0,4
gesamt Antworten	100,0	100,0	100,0	100,0	100,0
n Antworten	1.523 ³⁾	187 ¹⁾	110	77 ¹⁾	1.337 ²⁾
Antworten pro Person	1,7	1,0	1,0	1,0	1,8

¹⁾ Darunter eine Personen je zwei Nennungen ("nicht gut lesen" und "Sehprobleme") in T1a

²⁾ Darunter zwei Personen je zwei Nennungen ("nicht gut lesen" und "Sehprobleme") in T1a

³⁾ Darunter drei Personen je zwei Nennungen ("nicht gut lesen" und "Sehprobleme") in T1a

⁴⁾ Anders als in Tabelle 29 wurden für die Ergebnisdarstellung alle gegebenen Informationen aus den offenen und geschlossen Variablen genutzt. Insofern liegt auch der Wert "Verweigerung ohne nähere Begründung leicht über dem in Tabelle 29 berichteten Wert.

⁵⁾ Einbezogen wurden ausschließlich die Gründe oder Antwortverweigerungen bezogen auf die (Teil-)Aufgabe, die unmittelbar vor dem Testabbruch unvollständig bearbeitet wurde.

Vor allem für Personen, die den LEO-Test auf Teststufe 1 eher früh abbrachen, sind die hier genannten vermeintlichen Abbruchgründe für eine Einschätzung, ob die Zielperson aus literalitätsneutralen Gründen abbrach oder nicht relevant.⁴⁶ Die sechs genannten Gründe wurden von 76 Personen, also von 41 Prozent der Abbrechenden genannt. Getrennt nach Personen mit früherem (**T1** bis **T2**) und späterem (**T5** bis **T6**) Abbruch wurden sie unter Personen, die eher früh abgebrochen haben mit 37 Prozent weniger häufig genannt als unter solchen, die später abbrechen (46%). In der Gemengelage der (vermeintlichen) Abbruchgründe liegen einige weitere Unterschiede zwischen den beiden letztgenannten Gruppen. Die deutlichsten Unterschiede zeigen sich bei den drei Einzelgründen „Kann (die Wörter) nicht (gut) lesen“ (früh: 13% vs. spät: 26%), „Kann nicht genug Deutsch/Sprachprobleme“ (früh: 6% vs. spät: 0%) und „Generelle Testablehnung“ (früh: 7% vs. spät: 1%).

⁴⁶ Inhaltlich können gegebenenfalls auch weitere inhaltliche Gründe des Item-Nonresponse als vermeintliche Abbruchgründe herangezogen werden. Im Falle von Mehrfachnennungen wäre dann zu entscheiden, welcher Grund oder Begründungskomplex für einen Abbruch heranzuziehen ist. Hier wurden lediglich die je zuletzt erfragten Gründe direkt vor dem Abbruch – auch wenn dort „keine Angabe“ o.ä. gemacht wurde – herangezogen. Im Datensatz steht zudem eine Variable **agc**code die den letztgenannten inhaltlichen Grund im Falle eines Abbruchs anzeigt.

Wie bereits erwähnt, wurden im Anschluss an das Rätselheft die Bearbeitungsergebnisse in das CAPI übertragen, dort der Score berechnet und für all diejenigen, die den Testscore von „9“ unterschritten, zufällig eines von drei Alphaheften mit Bitte um weitere Bearbeitung ausgewählt. Laut CAPI unterschritten insgesamt 1.883 Zielpersonen den Score. Abzüglich der Befragungspersonen, die den Test im Rahmen des Rätselhefts abbrachen (n=185)⁴⁷ erhielten damit 1.698 Personen eines von drei Alphaheften.

Tabelle 33: Bearbeitete Alphaheftaufgaben insgesamt, als Item-Nonresponse und getrennt nach Abbruch bzw. weiterer Testbearbeitung

Basis: alle 18- bis 64-Jährigen (n=7.192) bzw. die von ihnen im Rätselheft bearbeiteten Aufgaben (n=56.659)	Auf Teststufe 2 (Alphaheft) bearbeitete Aufgaben ...			
	alle	darunter: unvollständig/nicht		darunter:
		bearbeitet	mit direkt darauf folgendem Testabbruch	mit weiterer Testbe- arbeitung ¹⁾
Alphaheft "Joschi"				
Aufgaben (n)	1.966	214	84	90
% von alle	100,0	10,9	4,3	4,6
Personen (n)	543	166	84	70
% von alle	100,0	30,6	15,5	12,9
durchschnittlich bearbeitete Aufgaben pro Person (m)	3,6	1,3	1,0	1,3
Alphaheft "Helga"				
Aufgaben (n)	2.212	152	49	84
% von alle	100,0	6,9	2,2	3,8
Personen (n)	584	128	49	73
% von alle	100,0	21,9	8,4	12,5
durchschnittlich bearbeitete Aufgaben pro Person (m)	3,8	1,2	1,0	1,2
Alphaheft "Leschek"				
Aufgaben (n)	2.100	275	73	154
% von alle	100,0	13,1	3,5	7,3
Personen (n)	571	225	73	141
% von alle	100,0	39,4	12,8	24,7
durchschnittlich bearbeitete Aufgaben pro Person (m)	3,7	1,2	1,0	1,1
alle Alphahefte				
Aufgaben (n)	6.278	641	206	328
% von alle	100,0	10,2	3,3	5,2
Personen (n)	1.698	519	206	284
% von alle	100,0	30,6	12,1	16,7
durchschnittlich bearbeitete Aufgaben pro Person (m)	3,7	1,2	1,0	1,2
¹⁾ Hier wurden ausschließlich die Informationen der ersten drei der insgesamt vier Aufgaben je Alphaheft berücksichtigt.				

⁴⁷ Eine Person, hat den Test abgebrochen, aber dennoch den Schwellenwert erreicht.

Unter allen Alphaheftaufgaben wurden zehn Prozent unvollständig bearbeitet. Unter den Personen, die die Alphahefte bearbeiteten, wurde wenigstens eine (Teil-)Aufgabe in einem der drei Alphahefte von knapp einem Drittel unvollständig bearbeitet (31%; Tabelle 33). Zwölf Prozent der auf Stufe 2 Teilnehmenden brachen diese Teststufe ab. Wie bereits erwähnt, liegt dieser Wert plausibler Weise deutlich höher als auf Stufe 1 (s. Kapitel 5.2). Die Annahme, dass die Alphahefte unterschiedlich schwer zu bearbeiten sind, wird sowohl durch die verschiedenen hohen Prozentwerte der unvollständig bearbeiteten Testaufgaben („Joschi“: 11%; „Helga“: 7%, „Leschek“: 13%) als auch durch die Quoten der direkt nach der unvollständig bearbeiteten Aufgabe Abbrechenden („Joschi“: 15% [gerundet]; „Helga“: 8%, „Leschek“: 13%) indiziert.

Im Falle einer unvollständigen Alphaheftaufgabenbearbeitung wurden in 57 Prozent der Fälle keine näheren Gründe dafür angegeben. Die wichtigsten genannten Gründe sind „aufgabenspezifische Schwierigkeiten“ (17%) und Aufgabenverständnisprobleme (9%). Diese beiden Gründe werden zudem überdurchschnittlich häufig als Begründung direkt vor einem Testabbruch genannt (22% bzw. 12%). Gründe wie „Zeitmangel“ oder „Interessenlosigkeit“ werden dagegen als Grund eines Item-Nonresponse überdurchschnittlich häufig vor einer Fortsetzung der Alphaheftbearbeitung (15% bzw. 9%) genannt. Nicht literalitätsneutrale Gründe (s.o.) sind im Rahmen der Alphahefte direkt vor dem Testabbruch in den Alphaheften deutlich seltener als Gründe benannt (insgesamt 13%) als in den Rätselheften (41%, s.o.).

Tabelle 34: Gründe für Item-Nonresponse auf der zweiten Teststufe (Alphahefte)

Basis: Personen, die wenigstens eine Alphaheftaufgabe unvollständig bearbeitet haben (n=519)	Anteilswerte in %											
	Alphaheft "Joschi"			Alphaheft "Helga"			Alphaheft "Leschek"			Alphahefte insgesamt		
	alle mit Item-Nonresponse	direkt vor Testabbruch	vor weiterer Bearbeitung	alle mit Item-Nonresponse	direkt vor Testabbruch	vor weiterer Bearbeitung	alle mit Item-Nonresponse	direkt vor Testabbruch	vor weiterer Bearbeitung	alle mit Item-Nonresponse	direkt vor Testabbruch	vor weiterer Bearbeitung
Gründe der unvollst. Aufgabenbearbeitung	(n=166)	(n=84)	(n=91)	(n=128)	(n=49)	(n=81)	(n=225)	(n=73)	(n=155)	(n=519)	(n=206)	(n=327)
Kann (die Wörter) nicht (gut) lesen	3,7	5,8	1,1	3,3	4,9	0,0	2,5	3,0	1,4	3,1	4,2	0,9
Kann nicht (gut) schreiben	0,9	1,7	0,0	0,7	1,0	0,0	0,7	1,0	0,0	0,8	1,2	0,0
Kann nicht genug Deutsch/Sprachprobl.	0,0	0,0	n=9; 10,7%	0,0	0,0	n=4; 8,2%	1,1	1,0	n=14; 19,2%	0,5	0,5	n=27; 13,1%
Hat Frage/Aufgabe nicht verstanden	12,1	20,0	2,1	10,5	12,6	6,1	6,2	7,4	2,7	9,2	12,2	3,2
Aufgabenspez. Schwierigkeiten	11,2	15,0	6,4	7,2	9,7	2,0	26,9	31,7	13,7	17,0	21,6	7,9
Sehprobleme	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
Generelle Testablehnung	0,0	0,0	0,0	2,0	1,9	2,0	0,0	0,0	0,0	0,5	0,5	0,5
Zeitmangel & and. situat. Gründe	9,3	0,8	20,2	3,3	0,0	10,2	4,0	1,0	12,3	5,6	0,7	15,3
Kein Interesse, keine Lust und pers. G.	5,1	0,8	10,6	5,3	2,9	10,2	2,5	1,5	5,5	4,1	1,6	8,8
Verweigerung ohne nähere Begründung	54,7	52,5	57,4	66,4	66,0	67,3	54,5	52,0	61,6	57,4	55,5	61,1
keine Angabe	2,8	3,3	2,1	1,3	1,0	2,0	1,5	1,5	1,4	1,9	1,9	1,9
gesamt Antworten	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0
n Antworten	214	84	130	152	49	103	275	73	202	641	206	435

Wie bereits erwähnt, sind hier Personen, die im Rahmen des Rätselhefts des Literalitätstest abbrechen aufgrund der zur Abschätzung der literalen Kompetenz zur Verfügung stehenden Informationen im Fokus der Betrachtungen. Zu unterscheiden sind dabei Personen, für die wenigstens einige Testinformationen vorliegen – also die später im Rätselheft Abbrechenden – und für die (nahezu) keine Testinformationen vorliegen – also die früher Abbrechenden. In einem nächsten Schritt ist zu prüfen, welche Informationen aus den Gründen, die direkt vor dem Abbruch benannt wurden, vorliegen. Unterschieden werden können hierbei drei Kategorien: a)

Verweigerung ohne Benennung von Gründen (inkl. „keine Angabe“), b) den sechs sogenannten nicht literalitätsneutralen Gründen und c) den literalitätsneutralen Gründen. Die Ergebnisse werden in Tabelle 35 gezeigt.

Tabelle 35: Abbrüche auf Teststufe 1 (Rätselheft) nach Grund und Zeitpunkt

Basis: Abbrechende in den Rätselheftaufgaben	Abbrechende insgesamt		früh Abbrechende (T1/T2)		spät Abbrechende (T5/T6)	
	n	% von allen 18- bis 64- Jährigen	n	% von allen 18- bis 64- Jährigen	n	% von allen 18- bis 64- Jährigen
Informationen aus den Abbruchgründen						
keine	802	11,2	47	0,7	33	0,5
nicht literalitätsneutrale	77	1,1	41	0,6	36	0,5
literalitätsneutrale	29	0,4	11	0,2	7	0,1
Abbrüche gesamt	186	2,6	110	1,5	76	1,1
<i>alle 18- bis 64-Jährigen</i>	7.192		7.192	100,0	7.192	100,0

Wenigstens einige Informationen liegen für alle spät Abbrechenden vor (1,1 %). Bei den früh Abbrechenden (1,5%) liegt dann eine qualifizierte Information hinsichtlich einer potenziellen Schriftsprachschwäche vor, wenn ein nicht literalitätsneutraler Grund zuletzt genannt wurde (0,6%). Analog zu den generellen Gründen der Nichtteilnahme an der LEO-Grundbildungsstudie (s. Kapitel 5.1) können die literalitätsneutralen Gründe dahingehend interpretiert werden, dass sie zu keinen systematischen Ausfällen führen und bis zur Erbringung des Gegenteils als *bevölkerungsdurchschnittlich* hinsichtlich der Testbearbeitung zu bewerten sind. Das ist im Falle der LEO-Grundbildungsstudie bei 0,2 Prozent der Fall. Für 0,7 Prozent der Befragungspersonen liegen allerdings (nahezu) keine Informationen aus den LEO-Tests vor und keine Angabe zu den Gründen des letzten Item-Nonresponse vor dem LEO-Testabbruch. Mit Blick auf die Bewertung erscheint diese Gruppe am schlechtesten zu gruppieren zu sein. Mit Blick auf das Gesamtergebnis ist dies allerdings als Erfolg zu bewerten, weil es sich dabei nur um aufgerundet ein Prozent der Gesamtstichprobe handelt. Aufgrund der leicht veränderten Gruppierung der Gründe, der niedrigen Fallzahlen und der leicht veränderten Rätselheftaufgaben wird an dieser Stelle auf einen direkten Vergleich mit Informationen aus LEO 2010 verzichtet.

6 Gewichtung und Hochrechnung

Die im Rahmen von LEO 2018 erhobenen Daten sollen es ermöglichen, eine empirische Schätzgröße für die Zahl der *funktionalen Analphabeten* in Deutschland zu ermitteln. Die dabei festgestellten Ergebnisse basieren auf einer Stichprobe, sollen aber verallgemeinerbare Aussagen über die Grundgesamtheit ermöglichen. Voraussetzung hierfür ist eine Stichprobe, die die Grundgesamtheit verzerrungsfrei abbildet. Grundsätzlich ist jedoch davon auszugehen, dass stichprobenbasierte Untersuchungen aufgrund verschiedener Ursachen von der Struktur der Grundgesamtheit abweichen. Dass dies für die Bildungsstrukturen in LEO 2018 der Fall ist, wurde in Kapitel 5.1.2 ausgeführt. Alle (weiteren) strukturellen Abweichungen können durch eine faktorielle Gewichtung korrigiert werden.

Die Stichprobe der LEO-Grundbildungsstudie setzt sich aus den Teilen der Basis- und der Zusatzerhebung zusammen (s. Kapitel 2.1). Die Auswahlwahrscheinlichkeiten der Personen und damit die Proportionalität der Abbildung sind in beiden Stichproben von Designkomponenten abhängig, die in der vorliegenden Untersuchung an verschiedenen Stellen zum Tragen kommen (s. Kapitel 2.5).

Darüber hinaus gibt es einen weiteren Ursachenkomplex für mögliche Strukturabweichungen: Zum einen ist bei Zufallsstichproben ein Zufallsfehler unvermeidlich, zum anderen sind bei jeder auf Freiwilligkeit basierenden Befragung *non-response*-bedingte Selektivitäten zu erwarten. Letztere stellen dann kein Problem dar, wenn die Nicht-Teilnahme zufällig über alle Bruttofälle verteilt ist – insofern also keine Korrelation mit einer der für die Auswertung relevanten Variablen vorliegt. Aufgrund der inhaltlichen Ausrichtung der Studie könnten sich jedoch an diesem Punkt Zweifel ergeben. Falls Personen, die ihre Lese- und Schreibschwäche verbergen wollen, aus eben diesem Grund die Teilnahme an der Studie verweigern, ergibt sich ein direkter Zusammenhang zwischen der ermittelten Größenordnung funktionaler Analphabeten und der Testverweigerung. Mithin wäre es aus theoretischer Sicht denkbar, dass die Höhe der *Illiteracy* durch die Studie unterschätzt wird. Erste Überprüfungen der wichtigsten Indizien für diese Hypothese in Kapitel 5.1 zeichnen jedoch ein anderes Bild: Die beiden dort verwendeten Prüfkriterien – die Struktur der Stichprobe nach Bildungsniveau und das faktische Vorhandensein von Personen mit Lese- und Schreibschwäche – bieten keine empirische Bestätigung für die systematische Nichtteilnahme von Personen mit geringer Lese- und Schreibkompetenz an der Befragung.

Im Rahmen der LEO 2018-Grundbildungsstudie wurden für zwei Stichproben Gewichtungsfaktoren berechnet:

- für die Basiserhebung und
- für die integrierte Gesamtstichprobe zusammengesetzt aus Basis- und Zusatzerhebung⁴⁸

⁴⁸ Auch für die Zusatzstichprobe wurde ein eigener Gewichtungsfaktor berechnet, aber nicht ausgeliefert, weil eine ausschließliche Auswertung der Daten der Zusatzerhebung inhaltlich wenig Beitrag zur Erkenntniserweiterung mit sich bringen dürfte. Auf Wunsch liefern wir den Faktor gerne nach.

Die bei Kantar durchgeführte Berechnung der Gewichtungsfaktoren lässt sich in zwei grundlegende Gewichtungsarten unterteilen. Zum einen wurden Designgewichte berechnet, die sich im Rahmen der auf Haushaltsebene realisierten Basisstichprobe aus den a priori festgelegten Auswahlwahrscheinlichkeiten der Personen ergeben (*Designgewichtung I*; Kapitel 6.1). Zum anderen wurden Designgewichte für die integrierte Gesamtstichprobe berechnet, um die über die unterschiedliche Zielgruppenvorgaben (Bildung) designbedingten Disproportionalitäten auszugleichen (*Designgewichtung II*; Kapitel 6.3). Dieser Schritt erlaubt bei der Gewichtung mit den so berechneten Faktoren eine Berücksichtigung des Stichprobendesigns. Eine Hochrechnung auf die jeweiligen Grundgesamtheiten ist damit aber noch nicht möglich. Hierzu dient die jeweils anschließend vorgenommene Kalibrierung, bei welcher die Fallzahlen und Verteilungen der Stichproben an die Referenzstatistiken bzw. Strukturen der Grundgesamtheit angeglichen werden. Das chronologische Vorgehen bei der Gewichtung wird in den Kapiteln 6.1 bis 6.4 nacheinander beschrieben.

6.1 Umwandlung der Basisstichprobe (*Designgewichtung I*)

Um unverzerrte Aussagen über die Grundgesamtheit zu erhalten, ist es notwendig, *design-bedingte Disproportionalitäten der Bruttostichproben* zu korrigieren. Auf der ersten Gewichtungsstufe werden zur Gewichtung der LEO 2018-Basisstichprobe Abweichungen vom ursprünglichen *haushaltsproportionalen Sample-Ansatz* korrigiert, indem Zellen nach verschiedenen regionalen Merkmalen (Bundesländer, Regierungsbezirke, BIK-Gemeindetypen, Kreise, Gemeinden und gegebenenfalls Stadtbezirke) gebildet wurden (s. Kapitel 2.5.1, „Erste Auswahlstufe“). In diesem Gewichtungsschritt werden die sich durch die Schichtung nach den genannten Merkmalen ergebenden unterschiedlichen Auswahlwahrscheinlichkeiten wie folgt ausgeglichen.

1. Regierungsbezirk (28 Zellen)
2. Bundesland (16 Gruppen) x BIK-Typ (10 Gruppen) (106 Zellen⁴⁹)
3. West-Ost (2 Gruppen) x BIK-Typ (10 Gruppen) (20 Zellen)

Streuung der Faktoren: 0,349 bis 3,989
Effektivität: 81,17%

Das in Kapitel 2.5.1 („Zweite Auswahlstufe“) beschriebene Auswahlverfahren führt zu einer *haushaltsrepräsentativen Stichprobe*, weil in jedem *Samplepoint* jeder Haushalt die gleiche Chance hatte, ausgewählt zu werden. In jedem der ausgewählten Haushalte wurde durch ein systematisch (gleiche Auswahlchancen innerhalb eines Haushalts) generierendes Verfahren nur eine Person als Zielperson ausgewählt, unabhängig davon, wie viele zur Grundgesamtheit gehörende Personen in dem betreffenden Haushalt leben (s. Kapitel 2.5.1, „Dritte Auswahlstufe“). Die Chancen für die in Privathaushalten lebenden Personen der Grundgesamtheit, als Befragungsperson ausgewählt zu werden, sind demnach umgekehrt proportional zur Anzahl der zur Grundgesamtheit gehörenden Personen in ihren Haushalten.

Diese ungleiche Auswahlwahrscheinlichkeit auf Personenebene wird durch einen *Umwandlungsfaktor* rechnerisch korrigiert. Dadurch wird die erstellte Haushaltsstichprobe in eine Personenstichprobe transformiert (*Designgewichtung*), in der jede Person der Grundgesamtheit stichprobentheoretisch die gleiche Auswahlchance hat. Dieser Schritt wird in unserem Gewichtungsprogramm selbst gemeinsam mit der Anpassung

⁴⁹ Darunter wurden einige Zellen aufgrund der vorliegenden Fallzahlen zusammengefasst.

der Personenstrukturen durchgeführt.⁵⁰ Diese zwei Schritte der Gewichtung wurden aufgrund der Stichprobenanlage als Personenstichprobe bei der Zusatzstichprobe nicht durchgeführt.

6.2 Anpassung an Bevölkerungsstrukturen (Kalibrierung)

Auf einer weiteren Gewichtungsstufe werden die Basis- und Zusatzstichprobe jeweils einzeln an die aus der amtlichen Statistik (hier: Bevölkerungsfortschreibung 2016 und Mikrozensus 2016) bekannten Sollstrukturen angepasst. Die Sollstrukturen für die Zusatzerhebung wurden in einer eigens dafür beim Statistischen Bundesamt angeforderten Berechnung aus dem Mikrozensus (2016) bereitgestellt. Die in Kapitel 6.1 beschriebenen Designgewichte werden im Falle der Basiserhebung als Eingangsfaktoren für die Anpassung verwendet.

Die Kalibrierung erfolgt bei Kantar in einem *iterativen Randsummenverfahren* (vgl. z.B. Wauschkuhn 1982; Deming, Stephan 1940). In diese Optimierungsbemühungen fließt auch die Begrenzung der Gewichtungsfaktoren ein, die verhindert, dass einzelnen Fällen in der Auswertung zu großem Gewicht beigemessen wird. Nach Festlegung der Variablenkombinationen wird das Modell als Gewichtung mit mehrdimensionalen Randverteilungen realisiert. D.h., es wird ein iterativer Gewichtungsprozess aufgesetzt, dessen Ergebnis Gewichtungsfaktoren sind, die die realisierte Stichprobe an alle vorgegebenen Sollverteilungen mit vordefinierter Genauigkeit und minimaler Varianz anpassen. Durch das iterative Vorgehen werden aktuelle Verschiebungen der Strukturen in späteren Iterationsschritten wieder korrigiert. Dadurch wird die Anpassung an alle Ränder sukzessive solange verbessert, bis ein vorher definiertes *Konvergenzkriterium* erreicht ist.

Die Gewichtungsfaktoren werden dabei auf ein vorgegebenes geschlossenes Intervall begrenzt, es wird also eine maximale Faktorenspannbreite zugelassen (für LEO 2018 erlaubte Faktoren: 0,2 bis 5,0). Diese Faktorenbegrenzung wird innerhalb jeder Iteration nach jeder Anpassung an eine Randverteilung wirksam. D.h., nach jeder Berechnung eines neuen Gewichtungsfaktors werden diejenigen Faktoren auf die vorgegebenen Grenzen gesetzt, welche diese unter- oder überschreiten. Somit wird die Kappung bei denjenigen Randverteilungen respektive Variablen wirksam, welche hohe und niedrige Gewichte erfordern. Bei Randverteilungen, die per se nur moderate Gewichte erfordern, wird durch diese Methodik keine Kappung angewandt. Die Kappung wird also soweit wie möglich in den Gewichtungsprozess integriert und nicht willkürlich im Nachhinein angewandt.

Als Resultat werden alle vorgegebenen Sollverteilungen so weit wie möglich realisiert, bis das Konvergenzkriterium erfüllt ist. Kriterium für alle Randverteilungen ist, dass in Summe maximal eine ein-prozentige Abweichung der einzelnen Zellen toleriert wird. Diese Vorgabe kann manchmal nicht erreicht werden. Die Iteration wird dann abgebrochen, wenn die Anpassung an die einzelnen Randverteilungen nicht mehr verbessert wird.

Aus dem sogenannten Gewichtungsprotokoll wird ersichtlich, ob die Faktoren in einzelnen Zellen zu hoch oder zu niedrig werden oder die Zellenbesetzung gleich *null* ist. In diesen Fällen können dann Zusammenfassungen mit geeigneten Nachbarzellen vorgenommen oder weitere Prüfungsschritte unternommen werden. Zur Ermitt-

⁵⁰ Allein berechnet läge die Effektivität dieser Anpassung bei 83,38%.

lung eines endgültigen Gewichtungsmodells werden verschiedene Varianten mit unterschiedlicher Kombination und Reihenfolge der Sollverteilungen erprobt, um ein Optimum zwischen Faktorenvarianz und -variation einerseits und dem Anpassungserfolg andererseits zu erreichen.

Die Gewichtungsfaktoren werden so berechnet, dass die ausgewiesene Gesamtfallzahl gewichtet genauso hoch ist wie vor der Gewichtung. In einzelnen Untergruppen können sich die gewichtete und die ungewichtete Fallzahl unterscheiden.

Die Gewichtungsmodelle für die vorliegende Studie sahen folgende Ränder bzw. Eckwerte vor:

Strukturanpassung auf Personenebene: Basisstichprobe LEO 2018

(einschließlich Umwandlung der Haushalts- in eine Personenstichprobe, s. Kapitel 6.1)

1.	West/Ost x Schulbildung (5 Gruppen)	(9 Zellen)
2.	Altersgruppen (5 Gruppen) x Schulbildung (5 Gruppen)	(21 Zellen)
3.	Bundesland (4 Gruppen) x Altersgruppen (9 Gruppen)	(36 Zellen)
4.	Bundesland (4 Gruppen) x Geschlecht	(8 Zellen)
5.	Bundesland	(16 Zellen)
6.	West/Ost x Geschlecht x Altersgruppen (9 Gruppen)	(36 Zellen)
7.	West/Ost x Deutsch/nicht Deutsch (3 Gruppen)	(5 Zellen)
8.	Erwerbsumfang - Vollzeit/Teilzeit (3 Gruppen)	(3 Zellen)
9.	West/Ost x Erwerbstätigkeit	(4 Zellen)

Streuung der Faktoren: 0,200 bis 4,997

Effektivität: 51,79%

Strukturanpassung auf Personenebene: Zusatzstichprobe LEO 2018

10.	Deutsch/nicht Deutsch	(2 Zellen)
11.	Bundesland	(15 Zellen)
12.	Geschlecht	(2 Zellen)
13.	Alter	(9 Zellen)
14.	Haushaltsgröße	(4 Zellen)
15.	Erwerbsumfang - Vollzeit/Teilzeit (3 Gruppen)	(3 Zellen)
16.	Erwerbstätigkeit	(2 Zellen)

Streuung der Faktoren: 0,202 und 4,802

Effektivität: 46,53%

6.3 Gewichtung der intergrierten LEO 2018-Stichprobe

Wie in Kapitel 2.5.2 dargestellt, wurde die Bruttostichprobe der LEO 2018-Zusatzerhebung auf Basis von wiederbefragungsbereiten Personen mit *niedrigem* Bildungshintergrund gezogen. Die Verteilung der Alterskohorten wurde dabei proportional zu Verteilung in der Grundgesamtheit angelegt. In der Basiserhebung dagegen ist eine Beschränkung der Zielpersonen auf eine bestimmte Gruppe nach Bildungshintergrund nicht vorgesehen. Um diese designbedingten Disproportionalitäten auszugleichen wird eine weitere Designgewichtung zur Anpassung der Bildungsstrukturen durchgeführt. Die personenbezogenen Gewichte der Anpassungen beider Einzelstichproben (s. Kapitel 6.2) gehen als Ausgangsgewicht ein.⁵¹ Um in einer weiteren Strukturgewichtung ausgleichende Randanpassungen zu nutzen, geht die Designgewichtung in den Folgeschritt zur nochmaligen Anpassung der Gesamtstrukturen auf Personenebene ein. Dabei wurden die sieben Schritte zur Anpassung der Ränder aus Personengewichtung der Zusatzerhebung und danach die neun Schritte zur Anpassung der Ränder aus Personengewichtung der Basiserhebung hier nacheinander durchgeführt. Eingefügt sind die Schritte zur *Proportionalisierung der Bildungsstrukturen* (1. Schritt) sowie zur Strukturangepassung der schulischen Bildung (11. Schritt).

Strukturanpassung auf Personenebene: Gesamtstichprobe LEO 2018

(einschließlich der Designgewichtung zum Ausgleich der disproportionalen Bildungsstrukturen; s. 1. Punkt und einer Bildungsanpassung, s. Punkt 11)

1.	Proportionalisierung der Unteren Bildung	(3 Zellen)
2.	Untere Bildung x Nationalität	(4 Zellen)
3.	Untere Bildung x Bundesland	(17 Zellen)
4.	Untere Bildung x Geschlecht	(4 Zellen)
5.	Untere Bildung x Alter	(11 Zellen)
6.	Untere Bildung x Haushaltsgröße	(6 Zellen)
7.	Untere Bildung x Erwerbsumfang - Vollzeit/Teilzeit	(5 Zellen)
8.	Untere Bildung x Erwerbstätigkeit	(4 Zellen)
9.	West/Ost x Schulbildung	(9 Zellen)
10.	Altersgruppen x Schulbildung	(21 Zellen)
11.	Schulbildung	(5 Zellen)
12.	Bundesland (4 Gruppen) x Altersgruppen	(36 Zellen)
13.	Bundesland (4 Gruppen) x Geschlecht	(8 Zellen)
14.	Bundesland	(16 Zellen)
15.	West/Ost x Geschlecht x Altersgruppen	(36 Zellen)
16.	West/Ost x Deutsch/nicht Deutsch	(5 Zellen)
17.	Erwerbsumfang - Vollzeit/Teilzeit	(3 Zellen)
18.	West/Ost x Erwerbstätigkeit	(4 Zellen)

Streuung der Faktoren: 0,201 und 4,891

Effektivität: 51.34%

Zur Gewichtung auf Personenebene sind dem LEO-Datensatz zwei Gewichtungsfaktoren zugespielt, die eine gewichtete Berechnung der Basisstichprobe allein (Auswahl mithilfe der Variablen **typ**; Gewichtungsvariable: **pgew**) und der Gesamtstichprobe einschließlich Zusatzstichprobe (Gewichtungsvariable: **pgewges**) zulassen.

⁵¹ Eine alleinige Designgewichtung würde darauf aufbauend eine Effektivität von 51,93% erzeugen.

Insgesamt und getrennt nach Stichproben sind in Tabelle 36 die Auszählungen der Variablen *höchster Schulabschluss* (**schulab**) dargestellt. In der ungewichteten Basiserhebung findet sich im Vergleich zur Sollstruktur laut Mikrozensus (2016; 28%) mit 24 Prozent ein etwas zu niedriger Anteil von Personen mit niedrigem oder keinem Schulabschluss. In der Zusatzerhebung liegt der entsprechende Wert gemäß der Zielgruppenvorgabe mit 99 Prozent deutlich darüber und in der Gesamtstichprobe liegt er um einen Prozentpunkt (29%) über dem gewünschten Anteilswert. In der Basiserhebung zeigt sich demnach, die aus anderen Erhebungen bekannte Bildungsselektivität (s, Kapitel 5.1.2). Personen mit höherem Bildungsniveau sind in der Regel eher dazu bereit, sich an Umfragen zu beteiligen. Dies spiegelt sich hier in dem niedrigeren Anteilswert von Personen mit niedrigerem schulischem Bildungsniveau in den ungewichteten LEO-Daten wider. In der gewichteten Basiserhebung ebenso wie in der gewichteten Gesamtstichprobe liegt jeweils der Anteilswert von Personen mit niedrigerem schulischem Bildungsabschluss (jeweils 28%) so hoch wie im Mikrozensus.

Tabelle 36: Nutzung der Personengewichtung getrennt nach Stichproben am Beispiel der Variablen „höchster Schulabschluss“ (SCHULAB)

<i>ungewichtet</i>		integrierte LEO-			Sollstruktur lt. MZ 2016	
höchster Schulabschluss (SCHULAB)	Value	Nettostichprobe	Basisstichprobe	Zusatzstichprobe	(in %)	
0 kein Abschluss	366	29,3%	266	100	99,4%	3,82
1 niedrig	1.738		1.330			408
2 mittel	2.176		2.176	0		33,44
3 hoch	2.756		2.756	0		37,08
4 derzeit Schulbesuch	143		140	3		1,44
9 keine Angabe	13		13	0		0,44
Gesamt	7.192		6.681	511		100,00
<i>SPSS-Syntax:</i>		<i>weigh off.</i> <i>fre schulab.</i>	<i>weigh off.</i> <i>cro schulab by typ.</i>			

<i>gewichtet</i>		(integriert)	(einzeln)	Sollstruktur lt. MZ 2016		
höchster Schulabschluss (SCHULAB)	Value	integrierte LEO- Nettostichprobe	Basisstichprobe	(in %)		
0 kein Abschluss	355	27,7%	311	28,2%	3,82	
1 niedrig	1.637		1.571		23,78	
2 mittel	2.410		2.232		33,44	
3 hoch	2.675		2.471		37,08	
4 derzeit Schulbesuch	104		85		1,44	
9 keine Angabe	11		11		0,44	
Gesamt	7.192		6.681		100,00	
<i>SPSS-Syntax:</i>		<i>weigh by pgewges.</i> <i>fre schulab.</i>	<i>weigh by pgew.</i> <i>fre schulab.</i>			

6.4 Erstellung der Gewichtungs- und Hochrechnungsfaktoren

Für eine Reihe von Eckwerten sind nicht nur relative, sondern auch auf die Grundgesamtheit projizierte Verteilungen von Interesse. Diese Hochrechnungsfaktoren lassen sich auf einfache Weise aus den Personengewichtungsfaktoren der einzelnen Stichproben ableiten. Um sie zu ermitteln, werden alle Gewichtungsfaktoren mit dem Verhältnis aus der Anzahl an Personen in der Grundgesamtheit und der Fallzahl pro Sample multipliziert. Die Grundgesamtheit der Basisstichprobe (18- bis 64-Jährige) umfasst laut amtlicher Bevölkerungsfortschreibung rd. 51,6 Mio.⁵² Personen, die Grundgesamtheit der Zusatzerhebung (18- bis 64-Jährige mit höchstens einem Hauptschulabschluss) liegt laut Mikrozensus 2016⁵³ bei rund 14,9 Mio. Die Hochrechnungsfaktoren sind im Datensatz in den Variablen **phochges** für die Gesamtstichprobe und **phoch** für die Basisstichprobe abgelegt.

⁵² Der aus dem Mikrozensus ermittelte Wert (Stand 2016) liegt mit 51,3 Mio. 18- bis 64-Jährigen etwas niedriger.

⁵³ Diese Information wurde aus dem Mikrozensus entnommen, weil sie in der Bevölkerungsfortschreibung nicht gegeben ist.

7 Literatur

- AAPOR (2016): *Standard Definitions*. Verfügbar unter: http://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf.
- Bilger, F. (2011): *AlphaPanel: Repräsentative Befragung von Teilnehmenden in Alphabetisierungskursen*. In: Birte Egloff und Anke Grotlüschen (Hrsg.): *Forschen im Feld der Alphabetisierung – Grundbildung*. Waxmann: Münster/New York/München/Berlin.
- Bilger, F., Jäckle, R., Rosenblatt, B. v. & Strauß, A. (2012): *Studiendesign, Durchführung und Methodik der leo. – Level-One Studie*. In: Anke Grotlüschen & Wibke Riekmann (Hg.): *Funktionaler Analphabetismus in Deutschland. Ergebnisse der ersten leo. – Level-One Studie*. Münster: Waxmann, S. 77-105.
- Bilger, F., Strauß, A., Gensicke, M., Baier, E. & Backes, A. (2017): *Datenerhebung für die LEO.-Grundbildungsstudie (LEO.2018). Pretestergebnisse*. München: Kantar Public (projektinterner Chartbericht).
- Bilger, F., Thümmel, K. & Strauß, A. (2018): *Weiterbildungsverhalten in Deutschland. AES 2016. Materialband 4. Codierung nach der ISCED-Fields*. Kantar Public: München.
- Bilger, F. & Strauß, A. (2017): *Beteiligung an non-formaler Weiterbildung*. In: F. Bilger, F. Behringer, H. Kuper und J. Schrader (Hg.): *Weiterbildungsverhalten in Deutschland 2016. Ergebnisse des Adult Education Survey (AES)*, Bielefeld: W. Bertelsmann, S. 25-55. Verfügbar unter: <https://www.die-bonn.de/doks/2017-weiterbildungsforschung-01.pdf>.
- Deming, W. Edward & Stephan, Frederick (1940): *On a Least Squares Adjustment of a Sampled Frequency Table When the Expected Marginal Totals are Known*. In: *Annals of Mathematics Statistics*, Vol. 11, No. 4, S. 427-444.
- Eurostat (2016): *Classification of Learning Activities (CLA) – Manual*. Luxembourg: Publications Office of the European Union. Verfügbar unter: <http://ec.europa.eu/eurostat/de/web/products-manuals-andguidelines/-/KS-GQ-15-011>.
- Fleiss, J. L. & Cohen, J. (1973). *The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability*. In: *Educational and Psychological Measurement*, 33, 613-619.
- Grotlüschen, Anke, Riekmann, Wibke & Buddeberg, Klaus (2012a): *leo. – Level-One Studie: Methodische Herausforderungen*. In: Anke Grotlüschen & Wibke Riekmann (Hg.): *Funktionaler Analphabetismus in Deutschland. Ergebnisse der ersten leo. – Level-One Studie*. Münster: Waxmann, S. 54-75.
- Grotlüschen, Anke, Riekmann, Wibke & Buddeberg, Klaus (2012b): *Hauptergebnisse der leo. – Level-One Studie*. In: Anke Grotlüschen & Wibke Riekmann (Hg.): *Funktionaler Analphabetismus in Deutschland. Ergebnisse der ersten leo. – Level-One Studie*. Münster: Waxmann, S. 13-53.

- Hartig, Johannes & Riekman, Wibke (2012): *Bestimmung der Level-Grenzen in der leo. – Level-One Studie*. In: Anke Grotlüschen & Wibke Riekman (Hg.): *Funktionaler Analphabetismus in Deutschland. Ergebnisse der ersten leo. – Level-One Studie*. Münster: Waxmann, S. 106-121.
- Mislevy, R., Beaton, A. Kaplan, B. & Sheehan, K. (1992). *Estimating Population Characteristics from Sparse Matrix Samples of Item Responses*. In: *Journal of Educational Measurement*, 133–161, 1992.
- OECD (2009): *PISA Data Analysis Manual SPSS®, 2nd Edition*. Paris, 2009: OECD. Verfügbar unter: <http://browse.oecdbookshop.org/oecd/pdfs/browseit/9809031E.PDF>.
- Ohne Autor (2017): *Anlage zum Grundsatzpapier zur Nationalen Dekade für Alphabetisierung und Grundbildung*. Stand: 08.03.2017, S. 8 (Verfügbar unter: https://www.alphadekade.de/files/01_Arbeitsprogramm_Nationale_Dekade_Alphabetisierung.pdf).
- Post, J. & Schütz, G. (2008). *BSW/AES 2007. Erhebung zum Weiterbildungsverhalten. Materialband 5. Die Codierung offener Nennungen zu Bildungs- und Weiterbildungsaktivitäten nach der ISCED-Fields-Klassifikation*. München: TNS Infratest Sozialforschung.
- Rosenblatt, Bernhard von & Bilger, Frauke (2011): *Erwachsene in Alphabetisierungskursen der Volkshochschulen. Ergebnisse einer repräsentativen Befragung (AlphaPanel)*. Bericht im Rahmen des Forschungsvorhabens „Verbleibsstudie zur biographischen Entwicklung ehemaliger Teilnehmer/-innen an Alphabetisierungskursen“ unter Leitung von Prof. Dr. Rainer Lehmann, Humboldt-Universität Berlin im Forschungsschwerpunkt „Alphabetisierung und Grundbildung“ des Bundesministeriums für Bildung und Forschung. Herausgegeben vom Deutschen Volkshochschul-Verband. Bonn.
- Robitzsch, A., Kiefer, T., & Wu, M. (2018). TAM: Test analysis modules. R package version 2.13-15. <https://CRAN.R-project.org/package=TAM>
- Rubin, D. B. (1987): *Multiple Imputation for Nonresponse in Surveys*. New York: J. Wiley & Sons.
- Statistisches Bundesamt (2008): *Klassifikation der Wirtschaftszweige mit Erläuterungen*. Wiesbaden (Verfügbar unter: https://www.destatis.de/DE/Publikationen/Verzeichnis/Klassifikation/WZ08_3100100089004.pdf;jsessionid=D4D7EC064C9A9BBD164D4966D649DB19.Internet-Live1?__blob=publicationFile).
- Statistisches Bundesamt (2016). *Handbuch der Ausbildungsfelder*. Wiesbaden: Destatis.
- Tamassia, E. et al. (2012). *Scoring Reliability Studies*. In: A. Zabal et al. (Hrsg), *Technical Report of the Survey of Adult Skills* (2nd Edition) (S.12-6). Münster: Waxmann.
- United Nations Educational, Scientific and Cultural Organization (UNESCO), UNESCO Institute for Statistics (1997, 2006). *International Standard Classification of Education – ISCED 1997*.
- United Nations Educational, Scientific and Cultural Organization (UNESCO), UNESCO Institute for Statistics (2012). *International Standard Classification of Education. ISCED 2011*. Montreal, Quebec. Verfügbar unter <http://www.uis.unesco.org/Education/Documents/isced-2011-en.pdf>.
- United Nations Educational, Scientific and Cultural Organization (UNESCO), UNESCO Institute for Statistics (2014). *ISCED Fields of Education and Training 2013 (ISCED-F 2013)*. Montreal. Verfügbar unter: <http://uis.unesco.org/en/topic/international-standard-classification-education-isced>.
- Wauschkuhn, U. (1982): *Anpassung von Stichproben und n-dimensionalen Tabellen an Randbedingungen*, München/Wien.

Anhang:

A1 Befragungsunterlagen

A1-1 LEO 2018-Listenheft

Listenheft

Sprache erkennen

- A Hauptschule/Volksschule
- B Realschule
- C Gymnasium
- D Polytechnische Oberschule (POS)
- E Erweiterte Oberschule (EOS)
- F Berufliches, Wirtschafts- oder technisches Gymnasium
- G Übergreifende Schulform, z.B. Sekundarschule oder Gesamtschule
- H Förderschule, Sonderschule (Sonderpädagogische Förderung)
- I Abendhauptschule, Abendrealschule
- J Abendgymnasium, Kolleg
- K Fachoberschule (FOS)
- L Berufsoberschule (BOS), technische Oberschule
- M Volkshochschule mit dem Ziel, einen Schulabschluss zu machen
- N Schule im Ausland besucht
- O eine andere, bisher nicht genannte Schule, und zwar: (*bitte angeben*) 

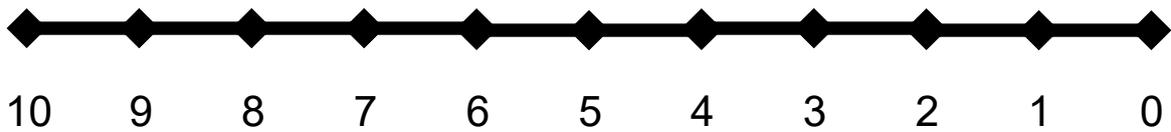
- A berufsvorbereitende Maßnahme, z.B. BVJ
- B Anlernausbildung oder berufliches Praktikum als mindestens sechsmonatige praktische Ausbildung im Betrieb
- C Berufsgrundbildungsjahr (BGJ)
- D 1-jährige Berufsfachschule, die berufliche Grundbildung vermittelt
- E betriebliche Lehre mit Berufsschule
- F Vorbereitungsdienst für Beamte des mittleren Dienstes
- G 2- bis 3-jährige Berufsfachschule, die einen Berufsabschluss vermittelt
- H 1-jährige Schule des Gesundheitswesens
- I 2- oder 3-jährige Schule des Gesundheitswesens
- J Ausbildungsstätte/Schule für Erzieher/-innen
- K Fachschule der DDR
- L Fachakademie (Bayern)
- M Fortbildung zum Meister
- N Fortbildung zum Techniker, Fachwirt/Fachkaufleute
- O Fachschule oder Fachakademie
- P Studium an einer Berufsakademie
- Q Studium an einer Verwaltungsfachhochschule
- R Studium an einer Fachhochschule
- S Studium an einer Universität oder entsprechenden Hochschule
- T Promotionsstudium
- U Andere, und zwar: ____ (*bitte angeben*) ____ 

- A Volks- oder Hauptschulabschluss (DDR: Polytechnischer Oberschule (POS) in Klasse 8 oder 9 beendet)
 - B Mittlere Reife (DDR: Polytechnischer Oberschule (POS) in Klasse 10 beendet)
 - C Abitur, Fachhochschulreife (DDR: Erweiterte Oberschule (EOS))
 - D Sonstigen Schulabschluss
- Keinen Schulabschluss

- A Ich verstehe und spreche nur einzelne Worte und Redewendungen.
- B Ich verstehe die gebräuchlichen Ausdrücke und kann mich in alltäglichen Situationen verständigen.
- C Ich kann das meiste verstehen, mich an Unterhaltungen aktiv beteiligen und einfache Texte verfassen.
- D Ich kann anspruchsvolle Texte lesen und schreiben. Ich beherrsche die Sprache nahezu vollständig.

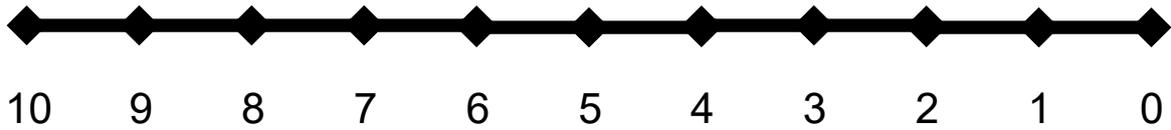
Ich kann auch
komplizierte Texte
problemlos lesen und
schreiben

Ich kann auf
Deutsch nicht
lesen und
schreiben



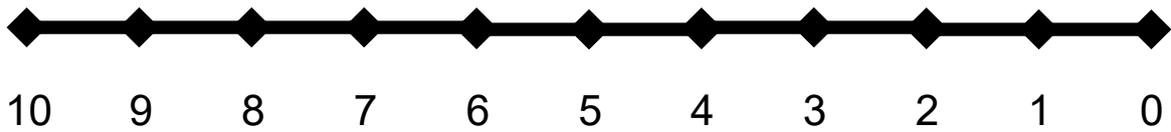
Ich bin voll und ganz
mit meinem Leben
zufrieden

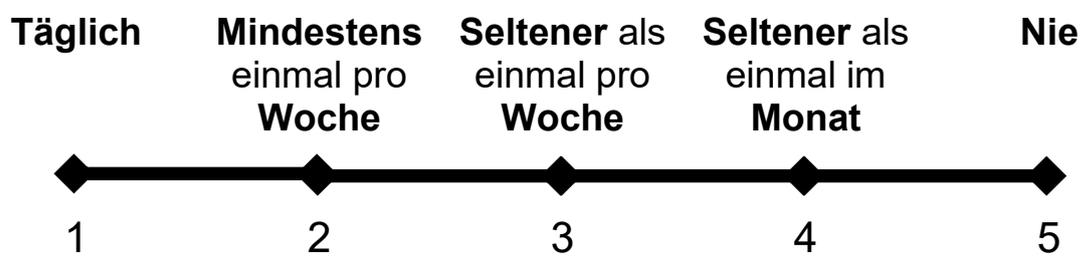
Ich bin völlig
unzufrieden

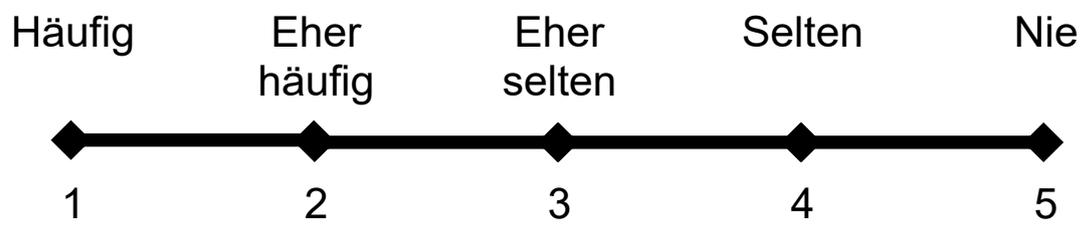


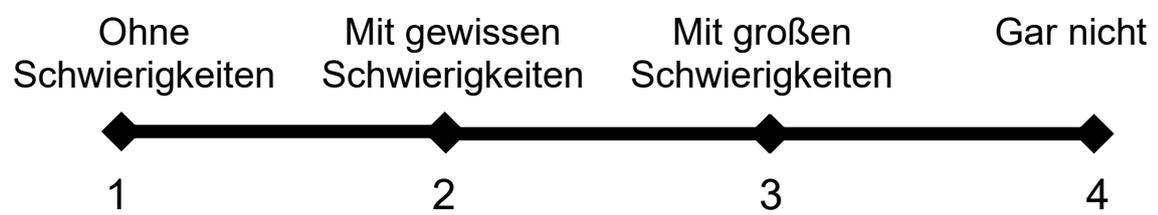
Ich fühle mich
zugehörig

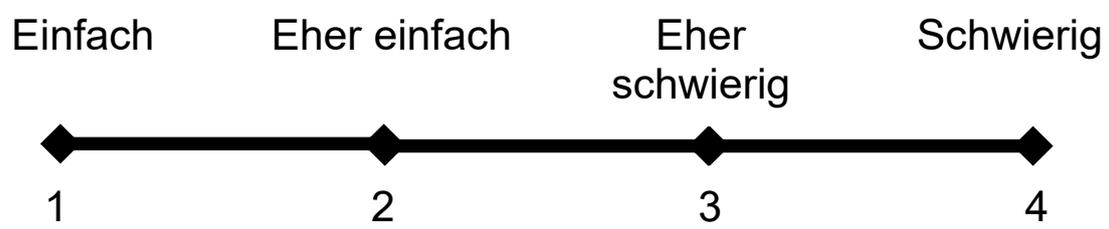
Ich fühle mich
ausgeschlossen











Ich bin
sehr
zufrieden

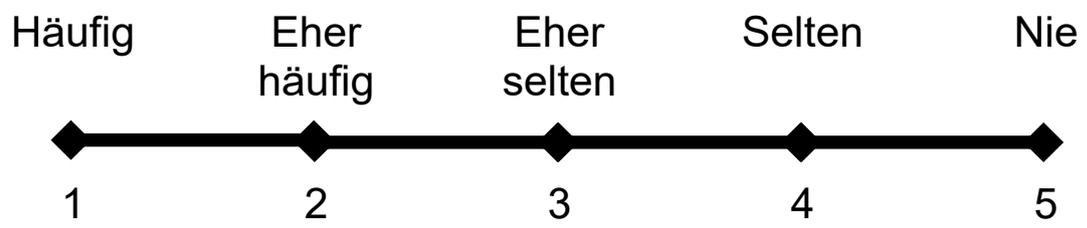
Ich bin
sehr
unzufrieden

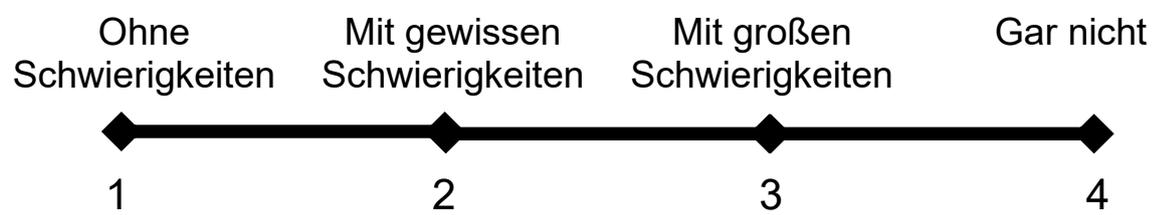


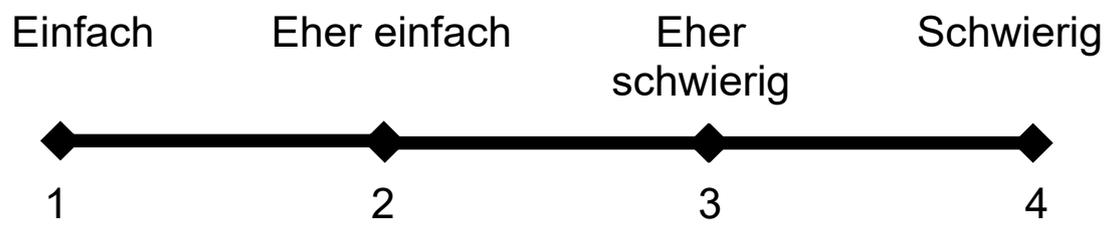
10 9 8 7 6 5 4 3 2 1 0

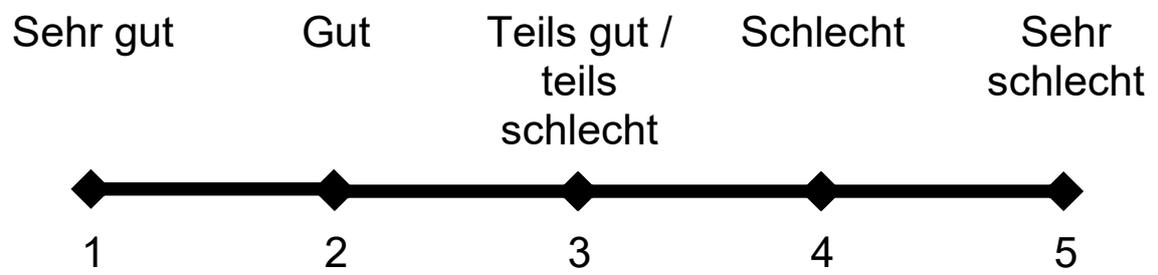
- A Voll erwerbstätig
- B In Teilzeit erwerbstätig
- C Arbeitslos
- D In Elternzeit
- E Auszubildende/r in betrieblicher Lehre
- G Schüler/in, Student/in, in einem Praktikum oder in sonstiger Ausbildung
- H Im Freiwilligen sozialen oder ökologischen Jahr, Bundesfreiwilligendienst
- I Rentner/in, Pensionär/in
- J Erwerbsunfähig
- K Hausfrau/Hausmann
- L Sonstiges

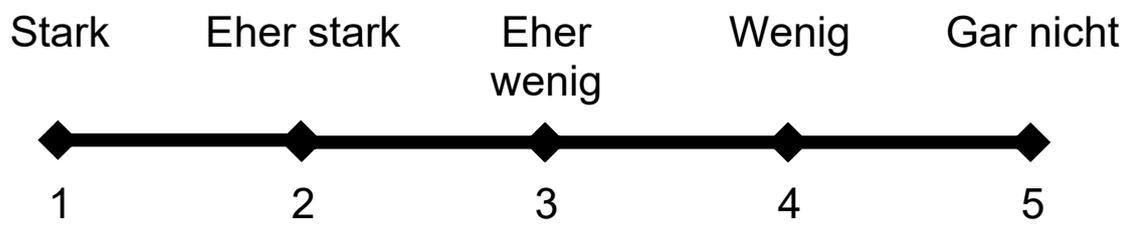
- A Keine Ausbildung erforderlich
- B Einen beruflichen Ausbildungsabschluss
- C Einen Fortbildungsabschluss, z.B. zum Meister,
Techniker oder
sonstigen Fachschulabschluss
- D Einen Hochschulabschluss (Fachhochschule oder
Universität)

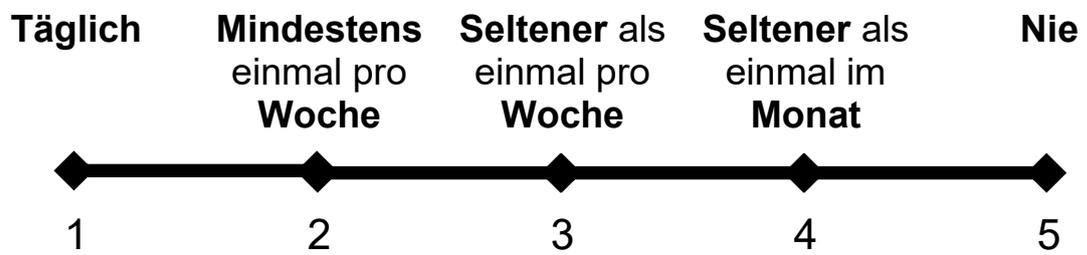


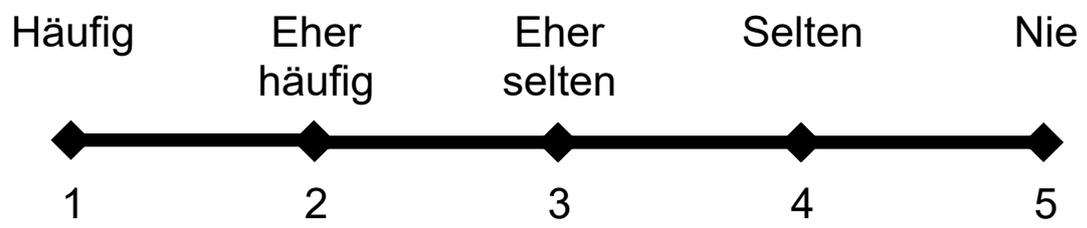


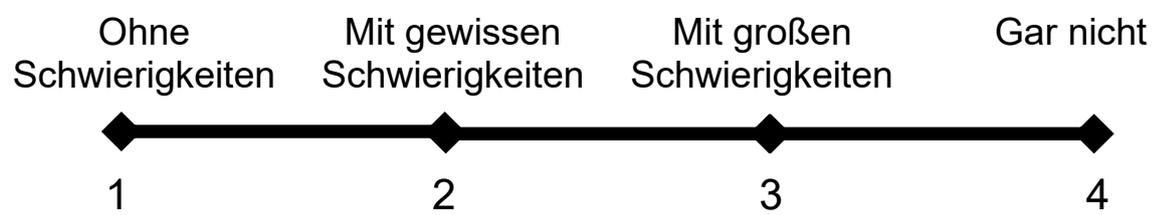


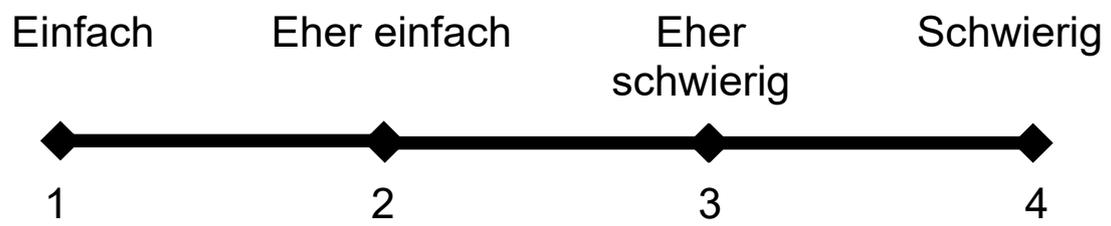








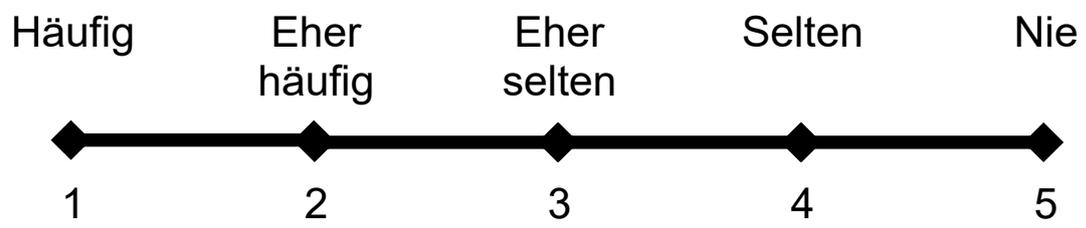




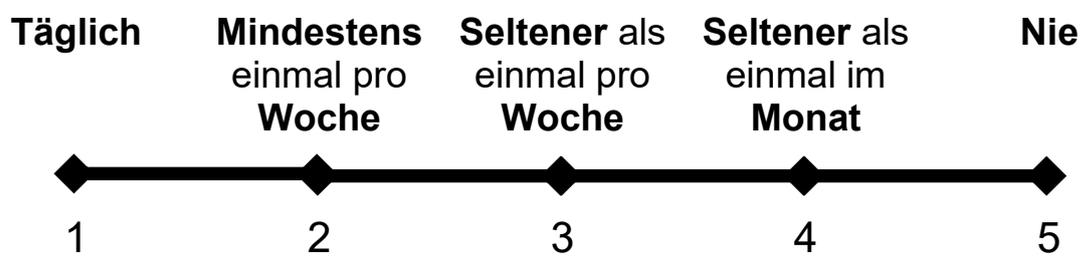
Ja, und zwar ...

- A Legasthenie (auch Lese-Rechtschreibschwäche oder Dyslexie)
- B Sprachfehler
- C Augenprobleme oder eine erhebliche Sehschwäche
- D Schwerhörigkeit
- E eine Schädigung durch einen Unfall
- F eine andere gesundheitliche Beeinträchtigung

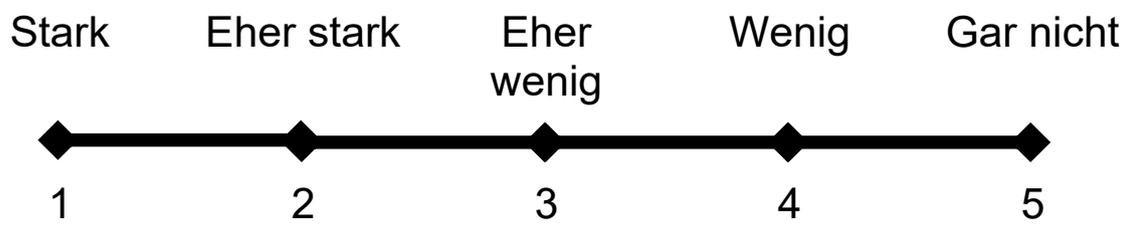
Nein, nichts davon

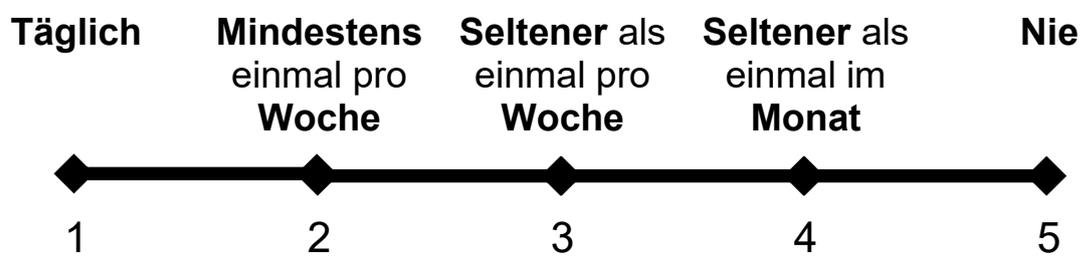


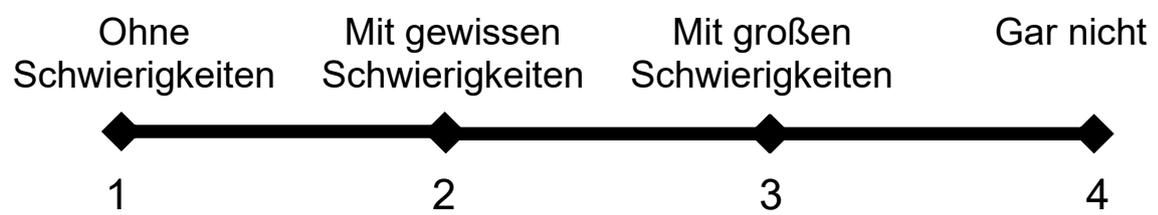
- (1) Ich selbst
- (2) Partnerin / Partner
- (3) Ein anderes Haushaltsmitglied
- (4) Eine andere Person, und zwar: ____(*bitte angeben*)____
- (5) Trifft nicht zu

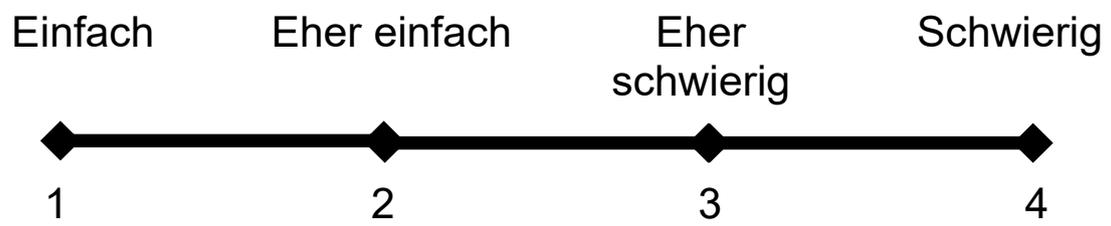


- (1) Ja, mehrmals am Tag
- (2) Ja, einmal am Tag
- (3) Ja, mehrmals in der Woche
- (4) Ja, einmal in der Woche
- (5) Ja, seltener als einmal in der Woche
- (6) Nein, ich lese meinem Kind nicht vor

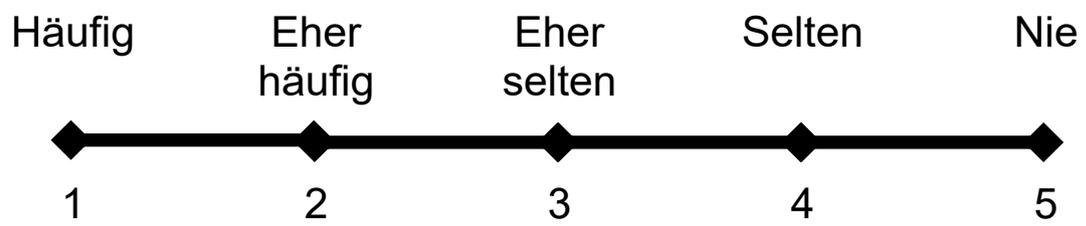


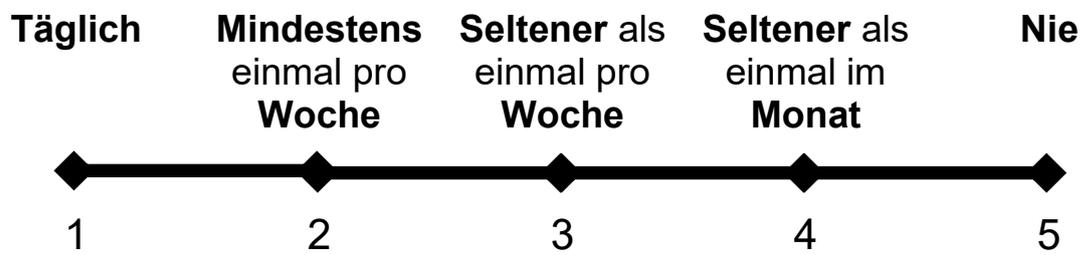






- A von Sportvereinen
 - B aus der Schule, dem Kindergarten, oder von der Eltern- oder Schülervertretung
 - C von Theater-, Musik-, Tanzgruppen oder anderen kulturellen Vereinigungen und Fördervereinen
 - D von Selbsthilfegruppen, Gruppen der Nachbarschaftshilfe, Wohlfahrtsverbänden oder anderen Hilfsorganisationen
 - E von Parteien oder politischen Interessenvertretungen
 - F von einer beruflichen Interessenvertretung, z.B. Gewerkschaft, dem Betriebsrat, einem Berufsverband oder einer Arbeitsloseninitiative
 - G von Kirchengemeinden, religiösen Vereinigungen oder Organisationen
- Nichts davon





- A als Asylbewerber/in oder Flüchtling
- B als Aussiedler/in oder Spätaussiedler/in
- C als Familienangehörige/r
- D als Student/in oder Studienbewerber/in, für eine Ausbildung oder ein Praktikum
- E als Arbeitnehmer/in
- F aus einem anderen Grund, und zwar: *__(bitte angeben)__* 

<i>Veranstaltungsarten in den letzten 12 Monaten</i>	
A. Kurse und Lehrgänge in der Arbeits- oder Freizeit	<p>Beispiele für berufliche Weiterbildung (auch im Betrieb)</p> <ul style="list-style-type: none"> · Schweißerlehrgang · Medizinische Fortbildung · Herstellerschulung für Maschinen · EDV-Lehrgang · Kurs Business-Englisch · Managementkurs <p>Beispiele für nicht-berufliche Weiterbildung (auch in der Freizeit)</p> <ul style="list-style-type: none"> · Erste Hilfe, Rettungsdienste · Computerkurs · Sprachkurs · Trainer-Lehrgang im Sport · Gesundheitsbildung · Politische Bildung
B. Kurzzeitige Bildungsveranstaltungen: Vorträge, Schulungen, Seminare, Workshops	<p>Beispiele für berufliche Weiterbildung (auch im Betrieb)</p> <ul style="list-style-type: none"> · Sicherheitsunterweisung · Schulung für eine neue Maschine · Vortrag über ein neues Gesetz · Halbtagesseminar in der IHK · Workshop Qualitätsmanagement · Führungskräfte training <p>Beispiele für nicht-berufliche Weiterbildung (auch in der Freizeit)</p> <ul style="list-style-type: none"> · Vortrag über Kultur / andere Länder · Seminar / Vortrag zur Kindererziehung · Naturkundliche Führung · Gesundheitsbildung · Politische Bildung
C. Schulungen / Training am Arbeitsplatz durch Vorgesetzte, Kollegen, Trainer, Teletutoren	<p>Beispiele</p> <ul style="list-style-type: none"> · Einarbeitung · Qualifizierung am Arbeitsplatz · Coaching · Job-Rotation · Traineeprogramme · Austauschprogramme mit anderen Firmen
D. Privatunterricht in der Freizeit	<p>Beispiele</p> <ul style="list-style-type: none"> · Fahrschule · individuelle Trainerstunden · Musikunterricht · Nachhilfestunden

- A Um meine berufliche Tätigkeit besser ausüben zu können
- B Um meine beruflichen Chancen zu verbessern
- C Um meinen Arbeitsplatz zu sichern
- D Um meine Aussichten auf einen Arbeitsplatz oder eine neue Stelle zu verbessern
- E Um mich selbstständig zu machen
- F Wegen organisatorischer bzw. technischer Veränderungen bei der Arbeit
- G Weil ich zur Teilnahme verpflichtet war
- H Um Kenntnisse/Fähigkeiten zu erwerben, die ich im Alltag nutzen kann
- I Um mein Wissen/meine Fähigkeiten zu einem Thema zu erweitern, das mich interessiert
- J Um ein Zertifikat/einen Prüfungsabschluss zu erwerben
- K Um Leute kennen zu lernen und Spaß zu haben
- L Um zu erfahren, wie ich gesünder leben kann
- M Um eine ehrenamtliche Tätigkeit besser ausführen zu können
- N Andere Gründe

A1-2 Empfehlungsschreiben der Auftraggebenden



FAKULTÄT
FÜR ERZIEHUNGSWISSENSCHAFT

Sehr geehrte Damen und Herren,

das Befragungsinstitut Kantar aus München führt im Auftrag der Universität Hamburg und im Auftrag des Bundesministeriums für Bildung und Forschung (BMBF) eine Haushaltsbefragung durch.

Es geht dabei um Sprache in unterschiedlichen Situationen des Alltags.

Hierzu bitten wir Sie um Ihre Mitwirkung.

Diese Befragung dient nur zu Forschungszwecken, sie dient nicht der Werbung oder Marktforschung.

Für Ihre Unterstützung danken



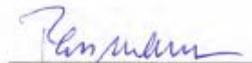
Universität Hamburg

Dr. Klaus Buddeberg
T. 040-42838-9464



Bundesministerium für
Bildung und Forschung

MR Peter Munk (Referatsleiter)



KANTAR PUBLIC
Matthias Passmann
T. 089-5600-2146



KANTAR PUBLIC

A1-3 Ankündigungsschreiben für die Zusatzerhebung

KANTAR PUBLIC

Landsberger Straße 284
80687 München
Germany
T +49 (0)89 5600 1275
F +49 (0)89 5600 1441
www.kantarpublic.com
www.ins-infratest.com/sofo

Anrede Vorname Nachname
Straße Hausnummer
PLZ Ort

25. Februar 2018

the 315107950

ID: ID-ZP zur Verständigung mit Hotline

Befragung zum Thema: „Sprache erkennen“

Sehr geehrte Damen, sehr geehrte Herren,

der digitale Wandel ist in vollem Gang. Die technologischen Entwicklungen sind rasant und verändern die Art, wie wir kommunizieren, arbeiten, lernen und leben. Für die aktive und selbstbestimmte Teilhabe an der digitalen Welt müssen die Bürgerinnen und Bürger mit den neuen Entwicklungen unter anderem durch stetes (Dazu-)Lernen Schritt halten. Daher hat das Bundesministerium für Bildung und Forschung (BMBF) zusammen mit der Universität Hamburg das Forschungsinstitut Kantar Public mit dieser Erhebung beauftragt. Die Studie zielt darauf ab, bundesweit repräsentativ zu ermitteln, wie Erwachsene im Alter von 18 bis 64 Jahren mit Sprache in unterschiedlichen Situationen des Alltags umgehen.

Warum gerade Sie?

Im Rahmen einer vorangegangenen Befragung haben Sie uns freundlicherweise Ihre Zustimmung erteilt, Sie zu einem späteren Zeitpunkt zu einem anderen Thema erneut zu kontaktieren. Ihre Teilnahme an der aktuellen Befragung ist wiederum selbstverständlich freiwillig. Für die Qualität und Aussagekraft dieser wissenschaftlichen Studie sind Ihre Angaben jedoch sehr wichtig. Wir bedanken uns für Ihre Teilnahme an einem vollständigen Interview, falls Sie zu unserer Zielgruppe gehören, mit 10 Euro in bar.

Wann werden Sie befragt?

In den nächsten Tagen oder Wochen wird ein/e Mitarbeiter/in des Kantar-Interviewerstabes mit Ihnen Kontakt aufnehmen, um ein Interview zu der oben genannten Studie durchzuführen.

Was passiert mit Ihren persönlichen Daten?

Wir garantieren Ihnen, dass all Ihre Angaben streng vertraulich behandelt werden. Ihre Antworten im Interview gehen ausschließlich in anonymisierter Form in die Analysen und Ergebnisdarstellungen ein. Weitere Details zu unseren Maßnahmen zum Schutz Ihrer persönlichen Daten können Sie der beiliegenden Datenschutzerklärung entnehmen oder später bei dem/der Interviewer/in erfragen.

Haben Sie weitere Fragen?

Sollten Sie Fragen haben, erreichen Sie uns unter der Rufnummer (089) 5600 2146.

Mit freundlichen Grüßen


Frauke Bilger


Alexandra Strauß


Matthias Passmann

Kantar Deutschland GmbH
Sitz / Registered in: München, AG München, HRB 113021
Umsatzsteuer-ID-Nr. / VAT-Number: DE 813390549

Geschäftsführer / Managing Director: Dr. Stefan Glawatsch, Winfried Hagenhoff, Jens Krüger,
Doris Lohaus, Frank Paule, Hartmut Scheffler, Dr. Nico A. Siegel, Robert A. Wieland
Aufsichtsrat / Supervisory Board: Bruno Käbber (Vorsitzender / Chairman)

Informationen für die Studienteilnehmenden – Fragen und Antworten

Was ist Zweck der BMBF-Erhebung?

Die Erhebung verfolgt mehrere Ziele. Die Daten werden in Deutschland als bildungspolitisches Steuerungsinstrument verwendet und dienen auch dem Ausbau der Bildungsinfrastruktur für Erwachsene. Darüber hinaus werden die anonymisierten Daten der fachwissenschaftlichen Öffentlichkeit zur Verfügung gestellt.

Wer steht hinter der Studie?

Die Erhebung wurde vom Bundesministerium für Bildung und Forschung (BMBF) und der Universität Hamburg beauftragt. Die Federführung der BMBF-Erhebung liegt bei der Universität Hamburg. Kantar Public unterstützt die Universität Hamburg beratend und führt die Befragung durch. Die Ergebnisse werden voraussichtlich im ersten Quartal 2019 auf der Homepage des BMBF veröffentlicht.

Wie läuft die Befragung ab?

Professionell ausgebildete Interviewer/innen des Forschungsinstituts Kantar führen die computergestützte Befragung durch. Ein/e Mitarbeiter/in aus dem Kantar-Interviewerstab wird Sie persönlich kontaktieren, um die Befragung durchzuführen oder mit Ihnen einen Termin dafür zu vereinbaren. Das Interview kann bei Ihnen zu Hause oder an einem neutralen Ort (z. B. in einem Café) stattfinden. Mithilfe eines vorbereiteten computergestützten Fragebogens wird der/die Interviewer/in Ihnen Fragen zu verschiedenen Themenbereichen stellen und Ihre Antworten direkt in den Computer eingeben. Wenn Sie Fragen nicht beantworten möchten, können Sie diese jederzeit überspringen. Ihre Teilnahme ist selbstverständlich freiwillig, d. h., durch eine Nicht-Beantwortung einzelner Fragen entstehen Ihnen keine Nachteile. Gleiches gilt für den Fall Ihrer Nicht-Teilnahme.

Was sind die Inhalte der Befragung?

Wir befragen Sie unter anderem zu den Bereichen Familie, Erwerbsleben, Bildung, (Nicht-)Nutzung digitaler Medien, Mobilität, Weiterbildung und Fremdsprachenkenntnisse. Wir interessieren uns z. B. dafür, ob Sie im Rahmen Ihrer Erwerbssituation überhaupt Bildungsmöglichkeiten haben. Und auch dafür, ob und inwiefern Sie digitale Medien nutzen. Wenn Sie keine digitalen Medien nutzen, interessieren wir uns sehr für Ihre Gründe.

Wie werden meine Daten vor Missbrauch geschützt?

Ihre Angaben werden OHNE Namen und Adresse (anonymisiert) und nur mit einer Codeziffer gespeichert. Ihre Adresse und Ihr Name werden stets getrennt von den Befragungsdaten aufbewahrt, so dass es nicht möglich ist, aus den Angaben im Interview Rückschlüsse auf Ihre Person zu ziehen. Die Daten werden ausschließlich für fachwissenschaftliche und bildungspolitische Zwecke genutzt. Weiteres zum Schutz Ihrer Angaben finden Sie in der beiliegenden Datenschutzerklärung.

Wenn Sie weitere Fragen haben, können Sie sich gerne an uns wenden. Die entsprechenden Kontaktinformationen finden Sie auf der Vorderseite (unten).

KANTAR PUBLIC

Erklärung zum Datenschutz und zur Vertraulichkeit Ihrer Angaben bei persönlichen Interviews

Kantar Public, ein Geschäftsbereich der Kantar Deutschland GmbH, ist Mitglied im Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e. V. (ADM) und arbeitet nach den Vorschriften des Bundesdatenschutzgesetzes (BDSG) und allen anderen datenschutzrechtlichen Bestimmungen.

Die Ergebnisse dieser Befragung werden ausschließlich in anonymisierter Form dargestellt. Das bedeutet: Aus den Ergebnissen kann niemand erkennen, von welcher Person oder Institution die Angaben gemacht worden sind.

Das gilt auch bei einer möglichen Folge- oder Wiederholungsbefragung, wo es wichtig ist, nach einer bestimmten Zeit noch einmal ein Interview mit derselben Person bzw. Institution durchzuführen, dabei auf Angaben aus dem Erstinterview zurückzugreifen und die statistische Auswertung so vorzunehmen, dass die Angaben aus mehreren Befragungen durch eine Codennummer, also ohne Namen und Adresse, miteinander verknüpft werden.

Kantar Public gibt auch keine Einzeldaten weiter, die Ihre Person oder Institution erkennen lassen könnten.

Falls die um Teilnahme gebetene Person noch nicht 18 Jahre alt und zurzeit kein Erwachsener anwesend ist: Bitte zeigen Sie dieses Merkblatt auch Ihren Eltern mit der Bitte, es billigend zur Kenntnis zu nehmen.

Für die Einhaltung der Datenschutzbestimmungen bei Kantar Public sind Dr. Nico A. Siegel und Frank Paule, Geschäftsführer der Kantar Deutschland GmbH, verantwortlich.

**Anschrift der
Kantar Deutschland GmbH**

Landsberger Straße 284
80687 München
t +49 89 5600 - 0
f +49 89 5600 - 1313

**Im Internet finden Sie
Informationen über uns unter:**

www.tns-infratest.com
www.tns-global.com

**Fragen zum Datenschutz beantwortet
der Datenschutzbeauftragte:**

Ass. iur. David Ohlenroth
t +49 89 5600 - 1178
f +49 89 5600 - 1730
e datenschutz@tns-infratest.com

**Nachfolgend zeigen wir Ihnen den Weg Ihrer Daten von der
Erhebung bis zur anonymen Ergebnisdarstellung!**

Was geschieht mit Ihren Angaben?

1. Ihre Antworten zu den Fragen werden vom Interviewer oder von Ihnen selbst in die Antwortfelder eingegeben bzw. eingetragen, z.B. so:

Welchen Schulabschluss haben Sie?	Volksschule	<input type="checkbox"/>
	Mittlere Reife	<input type="checkbox"/>
	Abitur	<input checked="" type="checkbox"/>

2. Bei der Kantar Deutschland GmbH wird der Fragenteil durchgesehen und von der Adresse getrennt; eventuelle Unklarheiten in den Antworten einzelner Befragter werden vorher telefonisch geklärt. Daten und Adresse erhalten eine Code-Nummer. Wer danach Ihre Antworten sieht, weiß also nicht, von wem sie gegeben wurden. Die Adresse verbleibt bei der Kantar Deutschland GmbH, jedoch nur bis zum Abschluss der Gesamtuntersuchung. Sie dient nur für Interviewerkontrollen (z.B. durch einen Telefonanruf oder Zusendung einer Postkarte mit der Bitte, die Durchführung des Interviews zu bestätigen) und dazu, Sie gegebenenfalls später mit der Bitte um ein neues Interview noch einmal aufzusuchen, anzuschreiben oder anzurufen.
3. Bei Interviews mit einem Papierfragebogen werden Ihre Angaben in Zahlen umgesetzt und ohne Ihren Namen und ohne Ihre Adresse (also anonymisiert) auf einen Datenträger gebracht. Bei PC-/Laptop-Interviews, wo die Fragentexte auf einem Bildschirm erscheinen, geschieht das bereits während des Interviews.
4. Anschließend werden die Interviewdaten (ohne Namen und Adresse) von einem Computer ausgewertet. Der Computer zählt z.B. alle Antworten nach dem Schulabschluss und errechnet die Prozentergebnisse.

5. Das Gesamtergebnis und die Ergebnisse von Teilgruppen (z.B. Arbeiter und Angestellte) werden in Tabellen ausgegeben.

Schulabschluss	Gesamt	Arbeiter	Angestellte
Volksschule	44%	50%	39%
Mittlere Reife	34%	41%	29%
Abitur	22%	9%	32%

6. Auch bei Folge- oder Wiederholungsbefragungen werden Ihr Name und Ihre Anschrift stets von den Daten des Fragenteils getrennt. Bei der Auswertung vergleicht der Computer – während er rechnet – pro Person, aber er tut das über die Codenummer (also niemals über Namen!), und gibt dann die Ergebnisse genauso anonymisiert aus wie bei der Einmalbefragung.
7. In jedem Fall gilt: Ihre Teilnahme ist **freiwillig** und Sie haben jederzeit das Recht, zu widersprechen. Es ist selbstverständlich, dass die Kantar Deutschland GmbH alle Vorschriften des Datenschutzes einhält.

Sie können absolut sicher sein, dass die Kantar Deutschland GmbH

- Ihren Namen und Ihre Anschrift oder die Ihrer Institution nach Abschluss der Gesamtuntersuchung nicht wieder mit den Interviewdaten zusammenführt, so dass niemand erfährt, welche Antworten Sie gegeben haben,
- Ihren Namen und Ihre Anschrift nicht an Dritte weitergibt,
- keine Einzelheiten an Dritte weitergibt, die eine Identifizierung Ihrer Person oder Institution zulassen.

**Wir danken Ihnen für Ihr Mitwirken und Ihr Vertrauen
in unsere Arbeit!**

A1-5 Extrablatt für die Interviewenden

Projekt Nr. 107950 „Sprache erkennen“

WICHTIG für die Befragung in Testteil 1 und 2 („Rätselheft“ und Hefte „Joschi“, „Helga“ und „Leschek“)

Die Befragten müssen einige Aufgaben lösen.
Sie als Interviewer/in lesen die Aufgaben vor, die ZP muss die Fragen durch Eintragungen im „Rätselheft“ oder den anderen Heften beantworten.
Das heißt: Sie werden hier zum „Testleitenden“.

WAS BEDEUTET DAS?

1. Um die Testsituationen für alle Befragten gleich zu gestalten, beachten Sie bitte die Hinweise im Fragenprogramm genau.
2. Die Zielperson soll die Aufgaben selbständig und ohne Ihre Mithilfe oder die Mithilfe anderer Personen lösen! Ansonsten würden die Ergebnisse der Studie verfälscht.
3. Wenn die Aufgaben des „Rätselhefts“ bearbeitet sind, müssen Sie die Antworten im Heft als richtig oder falsch bewerten. Dazu bekommen Sie klare Anleitungen im CAPI. Wenn Sie unsicher sind, weil Sie z.B. die Schrift nicht lesen können, geben Sie bitte „falsch“ an. Erst über diese Eingabe wird im CAPI-Programm ermittelt, ob die ZP sich für Testteil 2 qualifiziert und noch ein weiteres Heft „Joschi“, „Helga“ oder „Leschek“ erhält. Ihre Bewertung dient nur der CAPI-Auswahl. Eine genaue Auswertung der PAPI-Testhefte erfolgt durch Kantar Public.

Noch einige Tipps:

- Zu Beginn des Rätselheftes klären Sie bitte, ob die Zielperson zum Lesen eine Brille/Lesehilfe benötigt. Wenn ja, geben Sie ihr Zeit, diese zu holen und zu benutzen!
- ALLE Personen, die an Projekt **107950 „Sprache erkennen“** teilnehmen, bekommen ein Incentive in Höhe von 10 Euro. Nutzen Sie das Geld bitte, um die Personen zur Teilnahme an der Studie und insbesondere zur Beantwortung der Testteile zu bewegen.
- Bei der Bewertung der Testaufgaben (Fragen T7) ist Ihre Aufmerksamkeit am stärksten gefordert bei zwei lustigen Texten: T7-6a: „Meine Oma...“ und T7-6b: „Großer gemischter Salat...“. Bitte prüfen Sie genau, ob ZP die Texte richtig geschrieben hat – Ihr Bildschirm zeigt die richtige Schreibweise!
- Am Schluss vergewissern Sie sich, dass Sie auf dem Titelblatt der Hefte die nötigen Eintragungen vorgenommen haben: Welle, Listen- und laufende Nr. und **Ihre Interviewer- bzw. Abrechnungs-Nr. Hefte** ohne diese Angaben (LESBAR und KORREKT!) können wir leider nicht vergüten.

A2 Standartexte zur Vercodung

A2-1 Vercodung der Beruf nach ISCO 08

A2-1.1 Die Vercodung der Angaben zur beruflichen Tätigkeit nach der International Standard Classification of Occupations 2008 (ISCO 08)

Zur Ordnung von Berufsinformationen und, darauf aufbauend, zur Beschreibung von Struktur und Verteilung beruflicher Tätigkeiten sind in Deutschland derzeit nebeneinander zwei aktuelle Klassifikationssysteme vorhanden. Das erste ist die nationale *Klassifikation der Berufe 2010* (KldB 2010), beim anderen handelt es sich um die Internationale Standardklassifikation der Berufe (*International Standard Classification of Occupations*) des *International Labour Office* (ILO) aus dem Jahr 2008 (ISCO 08). Im Folgenden wird die Vorgehensweise der Vercodung des ISCO 08 erläutert.

Kantar Public ordnet den beruflichen Tätigkeiten den Code der Berufsgattung zu. Beim ISCO 08 handelt es sich dabei um einen vierstelligen Code. Das Kantar-Verfahren basiert auf einem umfangreichen *Dictionary* und der nachgelagerten *manuellen* direkten Vercodung der sogenannten *Leftovers*.

Grundlage der automatischen Zuordnung der Codes sind die elektronisch verfügbaren Verzeichnisse des ISCO 08⁵⁴. Hierbei wurden Unterschiede zu deutschen Berufsausbildungen (z.B. Pflegeberufe, die in Deutschland derzeit keine akademische Ausbildung erfordern) berücksichtigt. Diese Verzeichnisse wurden so überarbeitet, dass sie eindeutige Begriffe enthielten, denen die vierstelligen Codes zugeordnet sind. Dieses *Dictionary* wurde auf Basis der Ergebnisse der manuellen Vercodung der *Leftovers* weiterentwickelt, indem die *Leftovers*, die sich bei der manuellen Bearbeitung als eindeutig zuordenbar erwiesen hatten, ebenfalls in das *Dictionary* aufgenommen wurden. Die nach der Anwendung der Automatik verbleibenden offenen Angaben ohne Code werden anschließend manuell vercodet. Bei nicht eindeutig zuzuordnenden Berufen werden im Rahmen der manuellen Vercodung unter Verwendung zusätzlicher Angaben wie *Stellung im Beruf* oder *erforderliche Ausbildung* die Codes zugewiesen. Die bei der manuellen Vercodung angewendeten Regeln sind eindeutig formuliert und dokumentiert (s. nächster Abschnitt).

A2-1.2 Allgemeine Regeln bei der Vercodung der Berufe

Grundsätzlich wird die Angabe vercodet, die als berufliche Tätigkeit genannt wird (in LEO 2018 in der Frage **F010a** bzw. der Stringvariablen **BERUF** erfasst). Bei mehreren Angaben wird die erste vercodet, es sei denn, die zusätzliche/n Angabe/n erläutern die erste näher. Liegen weitere Angaben zu anderen Merkmalen (z.B. Branche (in LEO 2018: **F018**), Stellung im Beruf (in LEO 2018: **F011**, **F012**, **F013** und **F016**), Beamtenlaufbahnstufe (in LEO 2018: **F014**), für den Beruf erforderliche Ausbildung (in LEO 2018: **F010c**) vor, werden diese – sofern sinnvoll und schlüssig – zur genaueren manuellen Vercodung herangezogen. Bei zusammengesetzten Begriffen, die so nicht vercodet werden können, wird, wenn möglich, der übergeordnete Begriff vercodet (Beispiel: Hellermonteur als Monteur).

Typischerweise enthalten die offenen Angaben teilweise Nennungen, die so allgemein sind, dass die Zuordnung von Codes schwierig oder nicht eindeutig möglich ist. Ein hoher Vercodungserfolg impliziert, dass in

⁵⁴ ISCO 08 – gemeinsame deutschsprachige Titel und Erläuterungen auf Basis der englischsprachigen Version von Juni 2009, Stand der Bearbeitung 2010; erstellt von „STATISTIK AUSTRIA – Die Informationsmanager“.

Zweifelsfällen durchaus in gewisser Weise als willkürlich zu betrachtende Entscheidungen getroffen werden. Dies erscheint dann allerdings wenig problematisch, wenn die Entscheidungsregeln dokumentiert und diese Fälle entsprechend in einer Variable **isco08_prob** gekennzeichnet werden (s.u.).

Die von Kantar angewendeten allgemeinen Regeln sind folgende.

- **Regel 1:** Verkodet wird grundsätzlich der angegebene Begriff, wenn er in der ISCO08 aufgeführt ist.
- **Regel 2:** Bei mehreren Nennungen in einer Klartextangabe wird die erste vercodet, es sei denn, die weiteren Angaben dienen einer näheren Spezifizierung. Bei Doppelangaben, bei denen die eine die andere näher spezifiziert, wird die jeweils spezifischere Angabe vercodet. Bei zusammengesetzten Begriffen, die so nicht vercodet werden können, wird, wenn möglich, der übergeordnete Begriff vercodet (z.B. Hellermonteur als Monteur).
- **Regel 3:** Wenn eine Vercodung auf der Basis der offenen Angabe allein nicht möglich ist, werden zusätzliche Informationen – soweit vorhanden – hinzugezogen; diese eventuell zudem in geeigneter Kombination.
- **Regel 4:** Ist die offene Angabe so unspezifisch, dass sie mehreren Kennziffern zugeordnet werden kann (z.B. Händler – es ist nicht klar ob Einzel- oder Großhändler), wird sie derjenigen zugeordnet, für die angenommen werden kann, dass sie empirisch am häufigsten vorkommt. Diese Fälle werden in der Variable **isco08_prob** gekennzeichnet (s.u.).
- **Regel 5:** Wenn die Angabe zur beruflichen Tätigkeit keinen Aufschluss über die für die Tätigkeit erforderliche Qualifikation gibt und keine zusätzlichen Informationen hierzu vorliegen (s. auch Regel 3), wird der Code der Tätigkeit zugewiesen, für die die geringste Qualifikation notwendig ist. Diese Fälle werden in der Variable **isco08_prob** (s.u.) gekennzeichnet (z.B.: Betriebswirt aufgrund von Ausbildung, Weiterbildung oder Studium? Grafikdesigner über Ausbildung an Berufsfachschule, Fachschule oder Hochschule?).

Problematische Angaben, bei denen auf die Regeln 4 oder 5 zurückgegriffen wird, werden in einer Zusatzvariablen **isco08_prob** als solche gekennzeichnet. Dadurch besteht die Möglichkeit, Auswertungen mit oder ohne diese Fälle vorzunehmen und Sensitivitätstests durchzuführen. Um darüber hinaus eine Hilfestellung zur Beantwortung der Frage zu geben, inwiefern ein solches Vorgehen inhaltlich notwendig ist, werden im Folgenden die konkreten Sachverhalte, bei denen die Entscheidungsregeln angewendet werden, detailliert dokumentiert.

Während in früheren Versionen des ISCO88 für bestimmte Berufe Allgemeincodes vorgesehen waren (z.B. Angestellter, Ingenieur, Selbstständiger usw.), gibt es diese beim ISCO08 nicht mehr. Kantar Public codiert solche Angaben mit deutlich erkennbaren Sondercodes im Minusbereich (-10, -11 usw.), die in Kapitel A2.1.3 erläutert werden.

Die Angabe „Bachelor“ wird auf dem Level eines Technikers vercodet. Wenn für eine erweiterte Angabe ein dem Technikerniveau entsprechender Code vorhanden ist (z.B. Bachelor Ingenieurwissenschaften Maschinenbau), wird dieser vergeben. Die Angabe „Master“ wird auf dem Level eines Ingenieurs vercodet. Wenn für eine erweiterte Angabe ein dem Ingenieurniveau entsprechender Code vorhanden ist (z. B. Master Ingenieurwissenschaften Maschinenbau), wird dieser vergeben. Ist die Vergabe eines konkreten Codes für Bachelor oder Master nicht möglich, werden die Sondercodes -31 für Bachelor und -32 für Master vergeben (s. Kapitel A2.1.3).

Die Variable **isco08_prob** hat damit folgende Ausprägungen:

- isco08_prob:**
- 1: Qualifikationsstufe nicht bekannt, niedrigste vercodet
 - 2: Mehrere Codes möglich, Entscheidung getroffen
 - (3: *Angabe, aber nicht mehr erwerbstätig [emeritiert, Rente, Ehrenamt] bzw. ohne Entgelt und dann mit –1 vercodet; wird nicht immer angewandt*)

A2-1.3 Sondercodes bei nicht eindeutigen Angaben

Trotz der in Kapitel A2.1.2 dokumentierten Entscheidungsregeln beim Vorliegen nicht eindeutiger Angaben ist es nicht möglich, allen offenen Angaben einen Code zuzuordnen. Um die vorliegende Information zur beruflichen Tätigkeit dennoch zu verwerten und um Hinweise auf deren Inhalt zu geben, werden in solchen Fällen, wo möglich, Sondercodes vergeben, deren Bedeutung in Tabelle A2-1 gelistet sind

Tabelle A2-1: ISCO08-Sondercodes

Code	Value Label	Code	Value Label
-10	Aushilfe	-50	Berater, Consultant
-11	Auszubildende, Azubi, Berufsvorbereitung Arbeiter, Facharbeiter, Hilfsarbeiter, Produktionsarbeiter, Werker, Industrieangestellter, Industriearbeiter,	-51	Einsatzleiter, Leiter, leitender Angestellter
-12	Industriefacharbeiter, Teilwerker	-52	Betriebsleiter
-13	Industriemeister, Meister	-53	Kaufmann/-frau
-14	Techniker, technischer Angestellter	-54	Gutachter, Sachverständiger
-15	Ingenieur, technischer Ingenieur	-55	Referent
-16	Praktikant, Trainee, Volontär	-60	sozial, Sozialbereich, Sozialwesen
-17	Freiberufler, Geschäftsmann/-frau, Selbstständiger mithelfender Familienangehöriger (sowohl in als auch	-61	medizinisch, Medizinbereich, medizinischer Bereich
-18	außerhalb der Landwirtschaft)	-62	Therapeut
	Bundesfreiwilligendienst, Freiwilliges soziales Jahr, FSJ,		Hauptschulbesuch, Hauptschulabschluss, Berufseinstiegsjahr (BEJ), Berufsgrundbildungsjahr (BGJ),
-19	Wehrdienst, Zivildienst	-90	Berufsvorbereitungsjahr Realschulbesuch, Realschulabschluss, mittlere Reife, Fachoberschulreife, Berufsgrundschuljahr,
-20	Dienstleister, Dienstleistung	-91	Wirtschaftsschule, Wirtschaftsschulabschluss
-21	Zeitarbeit	-92	Berufsfachschulbesuch, Fachschulreife Fachschulbesuch, Fachschulabschluss (nur wenn eine Fachrichtung nicht erkennbar ist; wenn es sich z. B. um eine soziale, medizinische, pflegerische oder erzieherische
-22	Fachkraft studentische Angestellte/Hilfskraft, Werkstudent,	-93	Fachschule handelt, wird der entsprechende Code vergeben)
-30	wissenschaftliche Hilfskraft	-94	Handelsschulbesuch, Abschluss einer Handelsschule Fachoberschulbesuch (FOS), Fachhochschulreife,
-31	Bachelor	-95	Berufsfachoberschule, Berufsoberschule Gymnasium, Abitur, Hochschulreife, Fachabitur,
-32	Diplom, Magister, Master, 1. Staatsexamen, Staatsexamen	-96	fachbezogener Hochschulzugang
-33	2. Staatsexamen, Assessor, 2. Staatsprüfung, Referendariat Studium, Studium an Hochschule, Universitätsstudium, Fachhochschulabschluss, Fachhochschule, Aufbaustudium,		
-34	Fachdiplom	-1	keine Zuordnung möglich

A2-2 Vercodung der Wirtschaftszweige nach NACE 2008

A2-2.1 Die Vercodung der Angaben zum Wirtschaftszweig

Die Angaben zu den *Wirtschaftszweigen* können zum einen nach der älteren Wirtschaftszweigklassifikation der Europäischen Union (*Nomenclature des statistiques des activités économiques de la Communauté Européenne*: NACE Rev. 1.1) bzw. der deutschen *Klassifikation der Wirtschaftszweige 2003* (WZ 2003) eingeteilt werden, zum anderen nach der neusten *Wirtschaftszweigklassifikation der Europäischen Union (Nomenclature des statistiques des activités économiques de la Communauté Européenne*: NACE Rev. 2) bzw. der deutschen *Klassifikation der Wirtschaftszweige 2008* (WZ 2008). Bei der Wirtschaftszweigklassifikation nach NACE handelt es sich um das System der Europäischen Union, das System WZ stellt das der Bundesrepublik Deutschland dar. Die beiden Klassifikationssysteme stimmen jeweils bis zur Ebene der Klassen überein, die letzte Ebene der Unterklassen stellt ein nationales Spezifikum dar.

Durch das Verfahren von Kantar Public wird der zweistellige numerische Code für die Abteilungen zugewiesen. Grundlage für die manuelle Bearbeitung der nicht automatisch vercodeten Angaben zum Wirtschaftszweig nach WZ 2008 ist die *Klassifikation der Wirtschaftszweige* (Statistisches Bundesamt 2008). Neben den dort vorgegebenen konkreten Codes werden bei sehr allgemeinen Angaben vier zusätzliche Codes für „Industrie“, „Handwerk“, „Dienstleistung“ und „Produzierendes Gewerbe“ vergeben.

Zudem können bei nicht eindeutigen Angaben zum Wirtschaftszweig (LEO 2018: **F018**) weitere Informationen zur Spezifizierung herangezogen werden (z.B. Angabe zu beruflicher Tätigkeit [LEO 2018: **BERUF**], Stellung im Beruf [LEO 2018: **F011**, **F012**, **F013** und **F016**] o.ä.).

A2-2.2 Vercodung der Wirtschaftszweige nach WZ 2008

Bei den Angaben gibt es typischerweise Nennungen, die so allgemein sind, dass die Zuordnung von Codes schwierig oder nicht eindeutig möglich ist. Es ist selbstverständlich, dass ein hoher Vercodungserfolg impliziert, dass in Zweifelsfällen durchaus in gewisser Weise als willkürlich zu betrachtende Entscheidungen getroffen werden. Dies erscheint dann allerdings wenig problematisch, wenn die Entscheidungsregeln dokumentiert werden. Kantar Public hat alle bislang aufgetretenen Problemfälle erfasst und festgehalten, wie entschieden wurde. So entstand eine Sammlung sowohl von allgemeinen als auch speziellen Regeln, die genau festlegen, wie in solchen Fällen zu verfahren ist.

Es wurden folgende allgemeine Regeln angewendet.

- **Regel 1:** Ver codet wird grundsätzlich der angegebene Begriff, wenn er in der *Klassifikation der Wirtschaftszweige* des Statistischen Bundesamtes (2008) aufgeführt ist.
- **Regel 2:** Bei mehreren Nennungen in einer Klartextangabe wird die erste vercodet, es sei denn, die weiteren Angaben dienen einer näheren Spezifizierung. Bei Doppelangaben, bei denen die eine die andere näher spezifiziert, wird die jeweils spezifischere Angabe vercodet.
- **Regel 3:** Wenn eine Ver codung auf der Basis der offenen Angabe allein nicht möglich ist, werden zusätzliche Informationen – soweit vorhanden – hinzugezogen; diese eventuell zudem in geeigneter Kombination.
- **Regel 4:** Ist die offene Angabe so unspezifisch, dass sie mehreren Kennziffern zugeordnet werden kann, wird sie derjenigen zugeordnet, für die angenommen werden kann, dass sie empirisch am häufigsten vorkommt.

- **Regel 5:** Bei Angaben zum Wirtschaftszweig, die nicht erkennen lassen, ob Herstellung oder Handel gemeint ist, wird Herstellung vercodet.

Problematische Angaben, bei denen auf Regel 4 zurückgegriffen wird, werden in der Zusatzvariablen **wz2008_prob** mit Wert „2“ als solche gekennzeichnet. Problematische Angaben, bei denen Regel 5 zur Verwendung kommt, werden in der Zusatzvariablen **wz2008_prob** mit Wert „1“ gekennzeichnet. Dadurch besteht die Möglichkeit, Auswertungen mit diesen Fällen und ohne sie durchzuführen und somit Sensitivitätstests durchzuführen.

Spezielle Regeln

Die Befragten nennen als Branche oftmals den Namen der Firma, nochmals die berufliche Tätigkeit oder eine Bezeichnung, die nicht in das Klassifikationsschema passt bzw. unter mehrere der Wirtschaftszweig-Abteilungen fallen könnte, z.B. wenn der Betrieb eine Mischform darstellt. Grundsätzlich stellt die Vercodung nach NACE bei Letzterem auf den *überwiegenden* Geschäftsbereich ab, der von uns aber oft nicht feststellbar ist.

Vielfach kennen die Befragten ihre Branche nicht, wie Angaben wie Industrie, Handwerk, IG Metall, IG *Phantasienamen, die den Bereich irgendwie umschreiben*, zeigen.

Ebenso sind allgemeine Angaben wie Elektro, Elektriker, Elektrohandwerk, Elektrobranche, Elektrobetrieb usw. (möglich 26 bis 28 oder 43), Metall, Metallbearbeitung (24 oder 25), Metallbereich, Metallindustrie, Metallgewerbe (möglich 24, 25 oder 28, evtl. 26) nicht eindeutig zu vercoden.

Auch Angaben wie z.B. Agrargenossenschaft, Kommune, öffentlicher Dienst oder ein Gewerkschaftsbereich beziehen sich, wie den zugeordneten Berufen zu entnehmen ist, vielfach nicht auf diese Branche, sondern dokumentieren offensichtlich das Zugehörigkeitsgefühl bzw. die Mitgliedschaft der Befragten zu, nicht aber ihr Beschäftigungsverhältnis in dieser Kategorie. Hier einfach die entsprechenden Codes (Kommune = Kommunalverwaltung = 84, Agrargenossenschaft = Interessengemeinschaft = 94, IG Metall oder IG Druck und Papier = Gewerkschaft = Interessenvertretung = 94) zu vergeben, würde die Aussagekraft der Angaben deutlich verfälschen. In solchen Fällen wird – sofern möglich – nach der beruflichen Tätigkeit der entsprechende Wirtschaftszweig ermittelt. Lässt die Angabe zur beruflichen Tätigkeit darauf schließen, dass der Befragte tatsächlich in dem fraglichen Bereich tätig ist oder sein könnte, also z.B. in der Kirchenverwaltung oder als Gewerkschaftssekretär, oder keine Berufsangabe vorliegt, wird der entsprechende Code vergeben.

Bei zu allgemeinen Angaben, wie z.B. Bergbau oder Fensterbau, werden Codes vergeben, die zugegebenermaßen willkürlich und spekulativ sind, aber aufgrund von überwiegendem Vorkommen getroffen werden (z.B. Fenster-/Türen Bau = Holzverarbeitung = 16, da Holz hierfür das überwiegende Material darstellt). Diese Codes werden in **wz2008_prob** mit „2“ gekennzeichnet.

Bei mehreren Angaben (z.B. Metall-Kunststoff) wird in der Regel die erste vercodet, es sei denn, die weiteren Angaben dienen einer näheren Bestimmung der ersten.

Allgemeine Angaben ohne nähere Charakterisierung, die nicht erkennen lassen, ob Herstellung oder Handel gemeint ist, wie z.B. Kraftfahrzeuge/Lebensmittel o.ä., werden dem Bereich zugeordnet, der in der Wertschöpfungskette am weitesten vorne liegt, also beispielsweise dem Bereich Förderung, Erzeugung oder Herstellung und nicht dem Bereich Handel, Vermietung oder Reparatur. Auf der oben angesprochenen Zusatzvariablen **wz2008_prob** wird in diesen Fällen der Wert „1“ gesetzt.

Eine Ausnahme von dieser Regel bildet die Angabe *Computer, Computerbranche* ohne nähere Spezifizierung. Hier wird nicht *Computerherstellung* vercodet, sondern *Erbringung von Dienstleistungen der Informationstechnologie* (= 62), da diesem Wirtschaftszweig im Vergleich zur Herstellung die größere Reichweite zugeordnet werden kann. Auch hier wird auf der Zusatzvariablen **wz2008_prob** der Wert „1“ gesetzt.

Bei Metallbranchen wird die allgemeine Angabe Metall, Metallindustrie o.ä. wie *Verarbeitung*, also mit 25 (hier **wz2008_prob** Wert „2“), vercodet; bei den einzelnen Metallen (wie Aluminium, Stahl o.ä.) werden die analogen Angaben wie *Erzeugung*, also mit 24 (hier **wz2008_prob** Wert „1“), vercodet, da die zusätzlichen Berufsangaben weit überwiegend diese Einordnung nahelegen.

Außerdem werden die vier folgenden Codes für ansonsten nicht zuordenbare Angaben vergeben:

201	Industrie
202	Handwerk
203	Dienstleistungen
204	Produzierendes Gewerbe

A2-3 Vercodung der Themen von Weiterbildungsaktivitäten nach ISCED-Fields

Die *International Classification of Education* (ISCED) ermöglicht eine Klassifikation von Bildungsaktivitäten nach Bildungsniveau (*Level*) und -inhalten (*Fields*). Die ISCED ist eine hierarchische Beschreibung von Bildungsgängen getrennt nach *Bildungsfeldern* (*Fields*), also nach den Inhalten und Themen von Bildungsaktivitäten, für den internationalen Vergleich. Sie wurde in den 1970er Jahre von der UNESCO (ISCED 1976) entwickelt und aufgrund der zunehmenden Bildungsmöglichkeiten im Zeitverlauf angepasst. Im Jahr 1997 wurden sowohl *Levels* als auch *Fields* überarbeitet (United Nations Educational, Scientific and Cultural Organization [UNESCO], UNESCO Institute for Statistics 1997, 2006). Die Klassifikation wurde z.B. für den AES 2010, der Trägerstudie von LEO 2010 (s.o.), angewendet. Im Jahr 2011 wurde die Klassifikation der *Level* neuerlich überarbeitet (ISCED 2011; United Nations Educational, Scientific and Cultural Organization [UNESCO], UNESCO Institute for Statistics 2012) und eine separate Revision *ISCED-Fields-Klassifikation* beschlossen. Die überarbeitete englischsprachige Fassung der (Aus-)Bildungsfelder erschien mit ISCED-F 2013 (United Nations Educational, Scientific and Cultural Organization [UNESCO], UNESCO Institute for Statistics 2014). Die deutsche Fassung wurde im Sommer 2016 durch das statistische Bundesamt⁵⁵ erarbeitet, ging kurze Zeit später in den AES 2016 und nun in LEO 2018 ein.

Im Rahmen von LEO 2018 werden Fachrichtungen bzw. Themenbereiche der *non-formalen Weiterbildungsaktivitäten in den letzten 12 Monaten* offen erfragt (s. Fragen **Wb003**, **Wb005**, **Wb007** und **Wb009**) und – analog zum AES 2016⁵⁶ – mithilfe der ISCED-F klassifiziert. Weil die ISCED-F an sich darauf abstellt (Aus-)Bildungsfelder international zu vergleichen, war die Anwendung auf die Bereiche der non-formalen Weiterbildung (*non-formal education*⁵⁷) im AES 2007 ein Novum. Ähnlich wie die *International Standard Classification of Occupation* (ISCO; s. Kapitel A2-1) findet mit der ISCED eine Vercodung auf einer differenzierten (dreistelligen Ebene, *detailed fields*) statt, die dann wiederum auf eine zwei- (*narrow fields*) bzw. einstellige (*broad fields*) Ebene aggregiert werden kann.

Auf Ebene der detaillierten Felder, also auf dritter Ebene, eignet sich die ISCED-F für die Vercodung von Bildungs- und Lernaktivitäten. Zuordnungsschwierigkeiten ergeben sich allenfalls im Detail und in Einzelfällen. Für einen Überblick sind eher die zweiten oder ersten Aggregationsebenen zu nutzen. Bereits im AES 2007 zeigte sich, dass auf den höher aggregierten Ebenen die ISCED-F keinen inhaltlich sinnvollen Überblick über die *Aktivitäten non-formaler Weiterbildung* gibt (Post & Schütz, 2008). Damals wurde für eine geeignetere Darstellung auf Basis der detaillierten (Aus-)Bildungsfelder eine modifizierte Aggregation erarbeitet, bezeichnet als *ISCED-Fields 5-25*. Sie umfasste fünf breite Felder, innerhalb derer insgesamt 25 engere Felder, bezeichnet als Basiskategorien, unterschieden wurden. Mit dieser Klassifikation ließ sich die Themenstruktur aller in den AES-Erhebungen 2007 bis 2014 erhobenen Bildungs- und Lernaktivitäten inhaltlich sinnvoll und in einem angemessenen Grad der Differenzierung darstellen und analysieren. Eine analoge Darstellungsmöglichkeit wurde im Rahmen der AES-2016-Erhebung auf Basis der ISCED-F 2013 entwickelt und nun auch auf LEO 2018 angewendet. Eine genaue Dokumentation findet sich im vierten Materialband, der parallel zu den AES-2016-Daten im GESIS-Datenarchiv erhältlich ist (Bilger, Thümmel & Strauß, 2018).

⁵⁵ Statistische Ämter Deutschland, Österreich, Schweiz für die deutsche Übersetzung; Statistisches Bundesamt (2016)

⁵⁶ Das in LEO 2018 eingesetzten Instrument entstammt dem deutschen AES 2016.

⁵⁷ Mit einer *formal education*, also einer *formalen oder regulären Bildungsaktivität*, sind laut der *Classification of Learning Activities* (CLA; Eurostat, 2016) alle Aktivitäten gemeint, die wenigstens über sechs Monate andauern und zu einem Abschluss, verortet im *national qualification frame* (NQF), in Deutschland also dem *Deutschen Qualifikationsrahmen* (DQR), führen. Alle anderen Bildungsaktivitäten, die zwar durch eine (Lehr-)Person vorstrukturiert wurden – sei es off- oder online –, aber nicht zu einem solchen Abschluss führen, werden als *non-formal education*, also *non-formale Weiterbildungsaktivitäten* bezeichnet (s. auch Bilger & Strauß 2017).

Um einerseits den Vergleich mit der ISCED 1997 zu ermöglichen, andererseits die inhaltlichen Neuerungen der Ausbildungsfelder in den Blick zu bekommen, wurde die ISCED-F 2013 so wenig wie möglich und so viel wie nötig verändert. Die Klassifikation ist hierarchisch aufgebaut. Die Codierung nach Ausbildungsfeldern erfolgt auf der dritten Ebene, den sogenannten *detaillierten Feldern (detailed fields)* mit einem nunmehr vierstelligen Code (UNESCO Institute for Statistics 2014). Diese detaillierten Felder werden auf der zweiten Ebene zu den sogenannten engen Feldern (*narrow fields*) und der auf ersten Ebene der sogenannten breiten Feldern (*broad fields*) aggregiert. Wie Tabelle A2-2 zeigt, hat sich die Anzahl der Kategorien der Ausbildungsfelder für beide Klassifikationen nur unwesentlich verändert. Dabei sind die ergänzenden Codes (z.B. „nicht näher definiert“), die weiter unten erläutert werden, nicht berücksichtigt.

Tabelle A2-2: ISCED Fields 1997 und 2013 Anzahl der Ausbildungsfelder

Ebene	ISCED-Fields	ISCED 1997	ISCED-F 2013
1. Ebene	broad fields / breite Felder: 1997: 1-Steller 2013: 2-Steller	9	11
2. Ebene	narrow fields / enge Felder: 1997: 2-Steller 2013: 3-Steller	25	29
3. Ebene	detailed fields / detaillierte Felder: 1997: 3-Steller 2013: 4-Steller	79	78

Die wesentlichen Änderungen zwischen den beiden Klassifikationen der Jahre 1997 und 2013 sind in der Zuordnung der detaillierten Informationen zu den breiten Felder (1. Ebene) „Sozialwissenschaften, Wirtschaft und Recht“ (1997) in „Sozialwissenschaften, Journalismus und Informationswesen“ und „Wirtschaft, Verwaltung und Recht“ (2013) sowie „Naturwissenschaften“ (1997) in „Naturwissenschaften, Mathematik und Statistik“ und „Informatik und Kommunikationstechnologie“ (2013) zu finden. Die ISCED-F 2013 weist demnach elf anstelle von neun breiten Feldern auf. Eine weitere wichtige Änderung erfolgte auf der zweiten Ebene, den engen Feldern, mit der Herauslösung des Feldes „Sprachen“ aus dem Feld „Geisteswissenschaften“ (1997). In der ISCED-F 1997 gab es unter den detaillierten Feldern Codes, die zugleich denjenigen der engen Felder (2. Ebene) bzw. der breiten Felder (1. Ebene) entsprachen. Dies ist nun nicht mehr der Fall. Dagegen wurden in der ISCED-F 2013 auf der detaillierten, dritten Ebene allgemeine Codes eingeführt. Für weitere detailliertere Informationen zu den Veränderungen verweisen wir an dieser Stelle auf den o.g. vierten Materialband (Bilger, Thümmel & Strauß 2018).

Vercodung nach ISCED-Fields

Die Vercodung der insgesamt 6.205 offen benannten Themenbereiche oder Fachgebiete der *non-formalen Weiterbildungsaktivitäten* in LEO2018 nach detaillierten Feldern (3. Ebene) der ISCED-F 2013 erfolgte auf Basis der ins Deutsche übertragenen Liste „Umsetzung der Fachrichtungen, Berufe und Studienfächer in die ISCED-Fields“ des Statistischen Bundesamt (s.o.).

Die deutschen Labels und Codes liegen in einem *Dictionary* zur automatischen Vercodung der offenen Angaben vor. In einem ersten Schritt wurden die offenen Nennungen zu den Themenbereichen oder Fachgebieten automatisch vercodet. Die *Leftovers* wurden manuell vercodet. Im Rahmen der Vercodung waren auf allen

drei ISCED-F-Ebenen jeweils knapp zwei Prozent der Themenbereiche oder Fachgebiete der Weiterbildungsaktivitäten nicht klassifizierbar, was im Umkehrschluss bedeutet, dass 98 Prozent der Nennungen nach ISCED-F codiert werden konnten.

Aggregierte Darstellung nach ISCED-F (5-25/26-Klassifikation)

Die ISCED-F dient vor allem der Darstellung von Ausbildungsfeldern. Zur geeigneteren Darstellung der Themenbereiche oder -gebiete der Weiterbildungsaktivitäten auf aggregierter Ebene wurde bereits auf Basis der ISCED-F 1997 für den AES 2007 eine andere Form der Zusammenfassung der engen und breiten Felder (Ebenen 1 und 2) entwickelt. Diese sogenannte *5-25-Klassifikation* wurde auf Basis der detaillierten Felder zu 25 enger definierten Feldern, den sogenannten Basiskategorien, zusammengefasst, die wiederum zu fünf breiten Lernfeldern gebündelt wurden (Post, J. & Schütz, G. 2008).

Mit der erstmaligen Anwendung der ISCED-F 2013 im Rahmen des AES 2016 war zu prüfen, ob die 5-25-Klassifikation weiterhin anwendbar ist. In einem ersten Schritt erfolgte daher ein genauer Vergleich der detaillierten Felder der beiden ISCED-F der Jahre 1997 und 2013. Das war notwendig, um für gegebene Veränderungen in der Klassifikation die entsprechenden Codeziffern zu identifizieren. Alles in allem gelang die Übertragung. Die Mehrheit der detaillierten Felder der ISCED-F 2013 konnten somit inhaltlich den 25 Basiskategorien der Lernfelder der 5-25-Klassifikation problemlos zugeordnet werden, wenngleich eine neue Basiskategorie aufgrund der o.g. Herauslösung der „Sprachen“ aus „Sprachen Literatur, Geschichte, Religion, Philosophie“ eingeführt wurde. Vor diesem Hintergrund wurde auch der Klassifikationsname in *5-26-Klassifikation* geändert (genauer s. nächste Tabelle). Die in ISCED-F 2013 eingeführten Codes für „interdisziplinäre Programme und Qualifikationen“ konnten keiner Basiskategorie eindeutig zugeordnet werden und bekamen deshalb den Code 99 „nicht, oder nur einstellig klassifizierbar“, wobei sich „einstellig klassifizierbar“ hier auf die 5 Lernfelder der 5-26 Klassifikation bezieht. Ebenso wurde mit einigen Codes „nicht näher definiert“ verfahren. Tabelle A2-3, entnommen aus dem o.g. Materialband zeigt, welche dreistelligen Codes welchen Basiskategorien zugeordnet wurden und dies sowohl für ISCED-F 1997 als auch ISCED-F 2013.

Tabelle A2.3: Gegenüberstellung der 5-25-Klassifikation auf Basis der ISCED 1997 und der 5-26-Klassifikation auf Basis der ISCED-F 2013

detaillierte Felder der ISCED-F 1997		Basiskategorien der Lernfelder auf Basis der ISCED 1997		Lernfeldkategorien auf Basis der ISCED(-F) 1997 und 2013		Basiskategorien der Lernfelder auf Basis der ISCED-F 2013		detaillierte Felder der ISCED-F 2013	
Code	Label	Code	Label	Code	Label	Code	Label	Code	Label
10	Grundbildungsgänge	1	1-01: Grundbildungsgänge, Alphabetisierung	1	Sprachen, Kultur, Politik	1	1-01: Grundbildungsgänge, Alphabetisierung	11	Basisprogramme und Qualifikationen
80	Alphabetisierung und Vermittlung von Rechenfähigkeiten					21	Alphabetisierung und Vermittlung von Rechenfähigkeiten		
200	Geisteswiss. und Künste - weites Feld	2	1-04: Kunst, Musik, Medien			2	1-04: Kunst, Musik, Medien	210	Künste nicht näher definiert
210	Künste - breites Feld					211	Audiovisuelle Techniken und Medienproduktion		
211	Bildende Kunst					212	Mode, Innenarchitektur und industrielles Design		
212	Musik und darstellende Kunst					213	Bildende Kunst		
213	Audiovisuelle Techniken und Medienproduktion					214	Musik und darstellende Kunst		
214	Design					215	Kunsthandwerk		
215	Kunstgewerbe								
220	Geisteswissenschaften - breites Feld	3	1-05: Sprachen, Literatur, Geschichte, Religion, Philosophie			3	1-05: Sprachen	230	Sprachen nicht näher definiert
221	Religion und Theologie					231	Spracherwerb		
222	Fremdsprachenphilologie					232	Literatur und Linguistik		
223	Muttersprache					4	1-06: Literatur, Geschichte, Religion, Philosophie	220	Geisteswissenschaften (ohne Sprachen) nicht näher definiert
225	Geschichte					221	Religion und Theologie		
226	Philosophie					222	Geschichte und Archäologie		
						223	Philosophie und Ethik		
310	Sozial- und Verhaltenswiss. - breites Feld	4	1-06: Politik und Gesellschaft, Informationswesen			5	1-07: Politik und Gesellschaft, Informationswesen	311	Volkswirtschaftslehre
311	Psychologie					312	Politikwissenschaft und politische Bildung		
312	Kulturwissenschaften, Sozialwissenschaften					313	Psychologie		
313	Politikwissenschaften					314	Soziologie und Kulturwissenschaften		
314	Volkswirtschaftslehre					321	Journalismus und Berichterstattung		
321	Journalismus und Berichterstattung					322	Bibliothek, Informationswesen, Archiv		

Fortsetzung

detaillierte Felder der ISCED-F 1997		Basiskategorien der Lernfelder auf Basis der ISCED 1997		Lernfeldkategorien auf Basis der ISCED(-F) 1997 und 2013		Basiskategorien der Lernfelder auf Basis der ISCED-F 2013		detaillierte Felder der ISCED-F 2013	
Code	Label	Code	Label	Code	Label	Code	Label	Code	Label
322	Bibliothek, Informationswesen, Archiv							388	Interdisziplinäre Programme und Qualifikationen mit dem Schwerpunkt Sozialwiss.sch., Journalismus und Informationswesen
90	Persönlichkeitsentwicklung	5	2-02: Persönlichkeitsentwicklung, Kommunikation, Sozialkompetenz	2	Pädagogik und Sozialkompetenz	6	2-02: Persönlichkeitsentwicklung, Kommunikation, Sozialkompetenz	31	Persönliche Fähigkeiten und Persönlichkeitsentwicklung
140	Lehrerausbildung und Erziehungswiss. - breites Feld	6	2-03: Unterrichten, Ausbilden			7	2-03: Unterrichten, Ausbilden	110	Pädagogik nicht näher definiert
141	Lehrerausbildung					111		Erziehungswissenschaft	
142	Erziehungswissenschaft					112		Ausbildung von Lehrkräften für den vorschulischen Bereich	
						113		Ausbildung von Lehrkräften ohne Fachspezialisierung	
						114		Ausbildung von Lehrkräften mit Fachspezialisierung	
						188		Interdisziplinäre Programme und Qualifikationen mit dem Schwerpunkt Pädagogik	
760	Sozialwesen - breites Feld	7	2-20: Jugend- und Sozialarbeit			8	2-21: Jugend- und Sozialarbeit	922	Kinder- und Jugendarbeit
761	Kinder- und Jugendarbeit					923		Sozialarbeit und Beratung	
762	Sozialarbeit und Beratung								
720	Gesundheitswesen - breites Feld	8	3-18: Medizinische Themen, Pharmazie	3	Gesundheit und Sport	9	3-19: Medizinische Themen, Pharmazie	911	Zahnmedizin
721	Medizin					912		Humanmedizin	
724	Zahnmedizin					914		Medizinische Diagnostik und Behandlungstechnik	
727	Pharmazie					916		Pharmazie	
						917		Traditionelle und alternative Heilmethoden und Therapien	

Fortsetzung

detaillierte Felder der ISCED-F 1997		Basiskategorien der Lernfelder auf Basis der ISCED 1997		Lernfeldkategorien auf Basis der ISCED(-F) 1997 und 2013		Basiskategorien der Lernfelder auf Basis der ISCED-F 2013		detaillierte Felder der ISCED-F 2013	
Code	Label	Code	Label	Code	Label	Code	Label	Code	Label
722	Medizinische Dienste	9	3-19: Gesundheit, Pflege, medizinische Dienstleistung			10	3-20: Gesundheit, Pflege, medizinische Dienstleistung	910	Gesundheit nicht näher definiert
723	Krankenpflege							913	Krankenpflege und Geburtshilfe
								915	Therapie und Rehabilitation
								921	Pflege von alten oder behinderten Personen
813	Sport	10	3-22: Sport			11	3-23: Sport	1014	Sport
340	Wirtschaft und Verwaltung - breites Feld	11	4-07: Handel, Marketing, Rechnungswesen	4	Wirtschaft, Arbeit, Recht	12	4-08: Handel, Marketing, Rechnungswesen	411	Steuer- und Rechnungswesen
341	Handel							412	Finanz-, Bank- und Versicherungswesen
342	Marketing und Werbung							414	Marketing und Werbung
343	Kredit- und Versicherungswesen							415	Sekretariats- und Büroarbeit
344	Steuer- und Rechnungswesen							416	Groß- und Einzelhandel
346	Sekretariats- und Büroarbeit								
345	Management und Verwaltung							12	4-08: Management, Verwaltung
347	Arbeitswelt	13	4-09: Arbeit u. Organisation, Arbeitswelt	14	4-10: Arbeit u. Organisation, Arbeitswelt	417	Arbeitswelt		
380	Recht	14	4-10: Rechtsthemen	15	4-11: Rechtsthemen	421	Recht		
811	Gastgewerbe und Catering	15	4-21: Persönliche Dienstleistungen			16	4-22: Persönliche Dienstleistungen	1011	Hauswirtschaftliche Dienste
812	Reisebüros, Fremdenverkehrsgew. u. Freizeitindustrie							1012	Friseurgewerbe und Schönheitspflege
814	Hauswirtschaftliche Dienste							1013	Gastgewerbe und Catering
815	Friseurgewerbe und Schönheitspflege							1015	Reisebüros, Tourismus und Freizeitindustrie
850	Umweltschutz	16	4-24: Arbeitsschutz, Umweltschutz			17	4-25: Arbeitsschutz	1022	Gesundheit und Sicherheit am Arbeitsplatz

Fortsetzung

detaillierte Felder der ISCED-F 1997		Basiskategorien der Lernfelder auf Basis der ISCED 1997		Lernfeldkategorien auf Basis der ISCED(-F) 1997 und 2013		Basiskategorien der Lernfelder auf Basis der ISCED-F 2013		detaillierte Felder der ISCED-F 2013			
Code	Label	Code	Label	Code	Label	Code	Label	Code	Label		
862	Gesundheit und Sicherheit am Arbeitsplatz										
420	Biowissenschaften	17	5-11: Naturwissenschaftliche Themen	5	5: Natur, Technik, Computer"	18	5-12: Naturwissenschaftliche Themen	511	Biologie		
440	Exakte Naturwissenschaften - breites Feld							512	Biochemie		
441	Physik, Astronomie							521	Umweltwissenschaften		
442	Chemie, Lebensmittelchemie							522	Natürliche Lebensräume und Wildtiere		
443	Geowissenschaften, Geographie							531	Chemie		
								532	Geowissenschaften		
		533	Physik								
460	Mathematik und Statistik	18	5-12: Mathematik, Statistik			19	5-13: Mathematik, Statistik	19	5-13: Mathematik, Statistik	541	Mathematik
461	Mathematik									542	Statistik
462	Statistik										
481	Informatik	19	5-13: Informatik			20	5-14: Informatik	20	5-14: Informatik	612	Datenbanken, Netzwerkdesign und -administration
										613	Software- und Applikationsentwicklung und -analyse
480	Informatik und Computer-Bedienung - breites Feld	20	5-14: Umgang mit dem Computer, Softwarethemen			21	5-25: Umgang mit dem Computer, Softwarethemen	21	5-25: Umgang mit dem Computer, Softwarethemen	611	Computeranwendung
482	Computer-Bedienung										
520	Ingenieurwesen und technische Berufe - breites Feld	21	5-15: Ingenieurwesen, Technik			22	5-16: Ingenieurwesen, Technik	22	5-16: Ingenieurwesen, Technik	711	Chemie und Verfahrenstechnik
521	Maschinenbau und Metallverarbeitung									712	Umweltschutztechnologien
522	Elektrizität und Energie									713	Elektrizität und Energie
523	Elektronik und Automation									714	Elektronik und Automation
524	Chemie und Verfahrenstechnik			715	Maschinenbau und Metallverarbeitung						
525	Kraftfahrzeuge, Schiffe und Flugzeuge			716	Kraftfahrzeuge, Schiffe und Flugzeuge						
540	Verarbeitendes Gewerbe und Bergbau - breites Feld			22	5-16: Verarbeitendes Gewerbe, Bauwesen					23	5-17: Verarbeitendes Gewerbe, Bauwesen
541	Ernährungsgewerbe	722	Werkstoffe (Glas, Papier, Kunststoff und Holz)								
542	Textil, Bekleidung, Schuhe, Leder	723	Textilien (Kleidung, Schuhwerk und Leder)								
543	Holz, Papier, Kunststoff, Glas	724	Bergbau und Gewinnung von Steinen und Erden								

Fortsetzung

detaillierte Felder der ISCED-F 1997		Basiskategorien der Lernfelder auf Basis der ISCED 1997		Lernfeldkategorien auf Basis der ISCED(-F) 1997 und 2013		Basiskategorien der Lernfelder auf Basis der ISCED-F 2013		detaillierte Felder der ISCED-F 2013	
Code	Label	Code	Label	Code	Label	Code	Label	Code	Label
544	Bergbau und Gewinnung von Steinen und Erden							731	Architektur und Städteplanung
580	Architektur und Baugewerbe - breites Feld							732	Baugewerbe, Hoch- und Tiefbau
581	Architektur und Städteplanung								
582	Baugewerbe, Hoch- und Tiefbau								
620	Landwirtsch., Forstwirt. und Fischerei - breites Feld	23	5-17: Landwirtschaft, Tierheilkunde			24	5-18: Landwirtschaft, Tierheilkunde	811	Pflanzenbau und Tierzucht
6S21	Pflanzenbau und Tierzucht							812	Gartenbau
622	Gartenbau							821	Forstwirtschaft
623	Forstwirtschaft							831	Fischereiwirtschaft
624	Fischerei							841	Tiermedizin
640	Tiermedizin							888	Interdisziplinäre Programme und Qualifikationen mit dem Schwerpunkt Land-, Forstwirtschaft, Fischerei und Tiermedizin
840	Verkehrsdienstleistungen	24	5-23: Verkehr, Transport			25	5-24: Verkehr, Transport, Entsorgung	1021	Entsorgungsdienstleistungen
								1041	Verkehrsdienstleistungen
861	Schutz von Eigentum und Personen	25	5-25: Sicherheitsdienstleistungen			26	5-26: Sicherheitsdienstleistungen	1031	Militär und Verteidigung
863	Militär							1032	Schutz von Personen und Eigentum
300	Sozialwiss., Wirtschaft und Recht - weites Feld	99	nicht, oder nur einstellig klassifizierbar	9	nicht, oder nur einstellig klassifizierbar	99	nicht, oder nur einstellig klassifizierbar	288	Interdisziplinäre Programme und Qualifikationen mit dem Schwerpunkt Geisteswissenschaften und Künste
400	Naturwissenschaften - weites Feld							488	Interdisziplinäre Programme und Qualifikationen mit dem Schwerpunkt Wirtschaft, Verwaltung und Recht
500	Ingenieurw., Verarb. Gewerbe u. Baugew. - weites Feld							500	Naturwissenschaften, Mathematik und Statistik nicht näher definiert
800	Dienstleistungen - weites Feld							588	Interdisziplinäre Programme und Qualifikationen mit dem Schwerpunkt Naturwissenschaften, Mathematik und Statistik

Fortsetzung

detaillierte Felder der ISCED-F 1997		Basiskategorien der Lernfelder auf Basis der ISCED 1997		Lernfeldkategorien auf Basis der ISCED(-F) 1997 und 2013		Basiskategorien der Lernfelder auf Basis der ISCED-F 2013		detaillierte Felder der ISCED-F 2013	
Code	Label	Code	Label	Code	Label	Code	Label	Code	Label
999	nicht klassifizierbar							610	Informatik und Kommunikations-technologie nicht näher definiert
								688	Interdisziplinäre Programme und Qualifikationen mit dem Schwerpunkt Informatik und Kommunikationstechnologie
								700	Ingenieurwesen, verarbeitendes Gewerbe und Baugewerbe nicht näher definiert
								788	Interdisziplinäre Programme und Qualifikationen mit dem Schwerpunkt Ingenieurwesen, verarbeitendes und Baugewerbe
								900	Gesundheit und Sozialwesen nicht näher definiert
								988	Interdisziplinäre Programme und Qualifikationen mit dem Schwerpunkt Gesundheit und Sozialwesen
								1088	Interdisziplinäre Programme und Qualifikationen mit dem Schwerpunkt Dienstleistungen

Quelle: Kantar Public 2017

Die Anzahl nicht klassifizierbarer Nennungen in der 5-26-Klassifikation konnte reduziert werden, indem die detaillierten Felder (3. Ebene) „nicht näher definiert“ und „Interdisziplinäre Programme und Qualifikationen“ den fünf Lernfeldern der ISCED-F 2013 wie in Tabelle A2-3 gezeigt, zugeordnet wurden.

Tabelle A2-3: 5-26 Klassifikation Zusätzliche Codes in den 5 Lernfeldern

Lernfeld	ISCED Fields 2013 Codes
1 Sprachen, Kultur, Politik	288 Interdisziplinäre Programme und Qualifikationen mit dem Schwerpunkt Geisteswissenschaften und Künste
4 Wirtschaft, Arbeit, Recht	488 Interdisziplinäre Programme und Qualifikationen mit dem Schwerpunkt Wirtschaft, Verwaltung und Recht
5 Natur, Technik, Computer	500 Naturwissenschaften, Mathematik und Statistik nicht näher definiert
	588 Interdisziplinäre Programme und Qualifikationen mit dem Schwerpunkt Naturwissenschaften, Mathematik und Statistik
	610 Informatik und Kommunikationstechnologie nicht näher definiert
	688 Interdisziplinäre Programme und Qualifikationen mit dem Schwerpunkt Informatik und Kommunikationstechnologie
	788 Interdisziplinäre Programme und Qualifikationen mit dem Schwerpunkt Ingenieurwesen, verarbeitendes und Baugewerbe
9 nicht klassifizierbar	900 Gesundheit und Sozialwesen nicht näher definiert
	988 Interdisziplinäre Programme und Qualifikationen mit dem Schwerpunkt Gesundheit und Sozialwesen
	1088 Interdisziplinäre Programme und Qualifikationen mit dem Schwerpunkt Dienstleistungen

Eine Ergebnisübersicht der non-formalen Weiterbildungsaktivitäten aus LEO 2018, codiert auf Ebene der 5-26-Klassifikation zeigt Tabelle A2-4. Insgesamt sind aufgrund der Interdisziplinären Codes und der gar nicht klassifizierbaren Fälle Aussagen über die Inhalte von 94 Prozent der Weiterbildungsaktivitäten möglich.

Tabelle A2-4: Verteilung der Weiterbildungsaktivitäten nach Basiskategorien der 5-26-Klassifikation

Basis: Weiterbildungsaktivitäten (n=6.180)	Anteilswerte in %
1: Sprachen, Kultur, Politik	14,5
2: Pädagogik und Sozialkompetenz	10,3
3: Gesundheit und Sport	21,5
4: Wirtschaft, Arbeit, Recht	27,0
5: Natur, Technik, Computer	20,2
9: Nicht oder nur einseitig klassifizierbar	6,5
gesamt	100,0