

Methodenbericht

lidA Welle 3



Autoren: Jacob Steinwede, Jonathan Ruiz Marcos, Martin Kleudgen

infas Institut für angewandte
Sozialwissenschaft GmbH

Friedrich-Wilhelm-Straße 18
D-53113 Bonn
Tel. +49 (0)228/38 22-0
Fax +49 (0)228/31 00 71
info@infas.de
www.infas.de

Bericht an

Bergische Universität Wuppertal
Fakultät für Maschinenbau und Sicherheitstechnik
Lehrstuhl für Arbeitswissenschaft
D-42097 Wuppertal

Vorgelegt von

infas Institut für angewandte
Sozialwissenschaft GmbH
Friedrich-Wilhelm-Straße 18
53113 Bonn

Kontakt

Dr. Jacob Steinwede
Bereich Sozialforschung

Tel. +49 (0)228/38 22-408
Fax +49 (0)228/310071
E-Mail j.steinwede@infas.de

Autoren

Dr. Jacob Steinwede
Jonathan Ruiz Marcos
Martin Kleudgen

© infas Institut für angewandte Sozialwissenschaft GmbH
Der Inhalt dieses Berichts darf ganz oder teilweise
nur mit unserer schriftlichen Genehmigung veröffentlicht,
vervielfältigt, gedruckt oder in Informations- und
Dokumentationssystemen (information storage and
retrieval systems) gespeichert, verarbeitet oder ausgegeben
werden.

infas ist zertifiziert
nach ISO 20252 für die Markt-, Mei-
nungs- und Sozialforschung

ISO 20252



infas ist Mitglied im
Arbeitskreis Deutscher Markt- und
Sozialforschungsinstitute e.V. (ADM)
und ESOMAR



ESOMAR
| member

Vorbemerkung

„lidA – leben in der Arbeit“ ist eine Kohortenstudie in Deutschland, die das Zusammenwirken von Arbeit, Alter, Gesundheit und Erwerbsteilhabe seit 2009 untersucht. Dazu werden zwei Arbeitnehmerkohorten der Geburtsjahrgänge 1959 und 1965 im Abstand von ca. 3 Jahren wiederholt befragt. Bei Einwilligung der Befragten werden Daten aus anderen Quellen (z. B. Registerdaten der Bundesagentur für Arbeit) zugespielt.

Die dritte Welle der lidA-Studie wird von einem Konsortium bestehend aus Forschern der Universitäten Wuppertal und Duisburg-Essen, dem Universitätsklinikum Düsseldorf sowie dem infas Institut für angewandte Sozialwissenschaft GmbH durchgeführt. Die dritte Studienwelle wird gefördert von der Deutschen Gesetzlichen Unfallversicherung (DGUV), der Deutschen Rentenversicherung Bund (DRV Bund), der BARMER und dem Ministerium für Arbeit, Gesundheit und Soziales des Landes Nordrhein-Westfalen.

Die Befragungsteilnehmerinnen und -teilnehmer des lidA-Panels wurden nun zum dritten Mal interviewt. Die erste Erhebungswelle zum Aufbau des Panels wurde von März bis November 2011 durchgeführt. Die Felderhebung der zweiten Welle erfolgte zwischen Mitte Januar und Mitte April 2014 und die dritte Welle wurde von April bis Juni 2018 durchgeführt.

Der vorliegende Bericht beschreibt das methodische Vorgehen einschließlich der Panelpflege und die Feldergebnisse der dritten Erhebungswelle.

infas Institut für angewandte Sozialwissenschaft GmbH

Inhaltsverzeichnis

1	Das Erhebungsinstrument	7
1.1	Gliederung des Fragebogens	7
1.2	Schriftlicher Selbstausfüller zum gegenwärtigen Lebensgefühl	8
1.3	Erlaubnis zur Zuspiegelung von IEB-Daten und Rentenversicherungsdaten	9
2	Panelpflege und Einsatzstichprobe	10
3	Parameter für die Felddurchführung	12
3.1	Persönliches Anschreiben und Incentive	12
3.2	Identifizierung der richtigen Zielperson	12
3.3	Umgang mit verzogenen Personen	12
3.4	Nutzung von Preload-Informationen	12
3.5	Übersicht über die Erhebungsmaterialien	13
4	Feldergebnisse	14
4.1	Dokumentation der Kontakte im Feld	14
4.2	Feldbegleitendes Tracking	14
4.3	Realisierte Interviews	14
4.4	Interviewereinsatz und Kontakthäufigkeit	17
4.5	Interviewdauer	18
4.6	Rücklauf beim Selbstausfüller zum gegenwärtige Lebensgefühl	19
4.7	Zuspielungsbereitschaft für Registerdaten und Rentenversicherungsdaten	19
5	Datenprüfung und -lieferung	22
5.1	Interviewerkontrolle	22
5.2	Berufscodierung nach ISCO08 und KldB2010	23
5.3	Datenlieferungen	23
6	Selektivitätsanalyse und Gewichtung	25
6.1	Ausfallanalyse	25
6.2	Gewichtung	27

Abbildungsverzeichnis

Abbildung 1	Gliederung des CAPI-Erhebungsprogramms lidA Welle 3	8
-------------	---	---

Tabellenverzeichnis

Tabelle 1	Ergebnis der lidA-Panelpflege Herbst/Winter 2017/2018	11
Tabelle 2	Studienmaterial	13
Tabelle 3	Stichprobenausschöpfung lidA-Welle 3	16
Tabelle 4	Interviews pro Interviewerin und Interviewer	17
Tabelle 5	Kontakthäufigkeit	18
Tabelle 6	Interviewdauer modulweise	18
Tabelle 7	Vorliegen eines Selbstausfüllers	19
Tabelle 8	Einwilligungen zur Zuspiegelung externer Daten an Befragungsdaten	20
Tabelle 9	Ergebnis der Interviewerkontrolle	22
Tabelle 10	Ergebnis der Datenprüfung und Auslieferung gültig realisierter Fälle	22
Tabelle 11	Schriftliche Einwilligungsformulare BA-Registerdaten	24
Tabelle 12	Schriftliche Einwilligungsformulare Rentenversicherungsdaten	24
Tabelle 13	Multiplres logistisches Regressionsmodell zur Teilnahme von Wiederholungsbefragten an Welle 3	26
Tabelle 14	Vergleich von ungewichteten und gewichteten Merkmalen	28

1 Das Erhebungsinstrument

1.1 Gliederung des Fragebogens

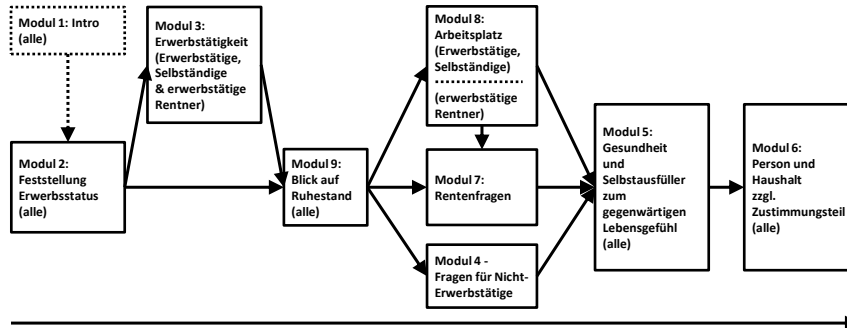
Die lidA-Erhebung der Welle 3 wurde, wie beide vorhergehende Befragungswellen, erneut als persönlich-mündliche Befragung mit Computerunterstützung (Computer Assisted Personal Interview, CAPI) durchgeführt. Mit dem Einsatz eines CAPI-Fragebogens wurde den Befragten auch ein Listenheft vorgelegt.

Der Fragebogen aus der zweiten lidA-Welle wurde für Welle 3 überarbeitet und um besondere inhaltliche Schwerpunkte ergänzt. Erneut abgefragt wurden Merkmale der Arbeitstätigkeit und des Arbeitsverhältnisses, persönliche Lebensumstände und Aspekte von Gesundheit. Formulierungen und Antwortformate wurden weitgehend identisch beibehalten, um eine Vergleichbarkeit über die Wellen hinweg zu ermöglichen. Neu kamen Fragen zu digitalisierter Arbeit/Flexibilisierung, Tätigkeits-/Berufswechseln, Maßnahmen der Arbeitsgestaltung, Inanspruchnahme von Rehabilitationsleistungen und betrieblicher Gesundheitsförderung hinzu. Dazu wurden teils bewährte Formulierungen aus vergleichbaren Untersuchungen eingesetzt, teils validierte Skalen eingesetzt oder gezielt Items durch die Studienleitung an der Bergischen Universität Wuppertal neu entwickelt. Die selbst entwickelten Items wurden vor ihrem Einsatz in der Studie in separaten Pretests - qualitative Interviews und einer quantitativen Online-Befragung, jeweils mit älteren Erwerbstätigen - durch die Studienleitung auf ihre Eignung überprüft und optimiert.

Der Fragebogen für die dritte lidA-Welle wurde für eine durchschnittliche Erhebungsdauer von 75 Minuten konzipiert. Gegenüber den Vorgängerbefragungen wurde der Ablauf geändert. Der Fragebogen für Welle 3 gliedert sich in 9 Module (vgl. Abbildung 1). Nach Feststellung von Erwerbstätigkeit und Rentenbezug wurden alle Zielpersonen fünf Gruppen zugeordnet:

- abhängig Erwerbstätige,
- Selbständige,
- Rentnerinnen und Rentner, die noch erwerbstätig sind,
- Rentnerinnen und Rentner ohne Erwerb,
- Erwerbslose, die noch nicht in Rente sind.

Anhand der Gruppenzuordnung erfolgte dann die weitere Steuerung des Fragebogens. Erwerbstätige erhielten Fragen zu derzeitigen oder früheren Erwerbstätigkeiten und zu Merkmalen der Arbeit. Nicht-Erwerbstätige erhielten ein kurzes eigenes Modul. Neu in Welle 3 waren die Module „Blick auf den bevorstehenden Ruhestand“ und „Fragen zur Rente“.

Abbildung 1 Gliederung des CAPI-Erhebungsprogramms lidA Welle 3


Wie in den vorhergehenden Wellen wurden die Befragten gebeten, einen „Selbstaussfüller zum gegenwärtigen Lebensgefühl“ auszufüllen. Abschließend wurden Angaben zur Person und zum Haushaltskontext erhoben.

Abgeschlossen wurde die Erhebung in Welle 3 mit der Abfrage und Dokumentation von Einwilligungen der Befragten in einem sogenannten Zustimmungsblock. Schon in der ersten und zweiten Welle der lidA-Befragung wurde um die Erlaubnis zur Zuspiegung von externen Daten an die lidA-Befragungsdaten gebeten. Unabhängig von früheren Zustimmungen wurden in Welle 3 alle Zielpersonen um ihr Einverständnis gebeten,

- dass ergänzende Daten der Bundesagentur für Arbeit (IEB-Daten) hinzugespielt werden dürfen und
- dass Daten der Deutschen Rentenversicherung Bund hinzugespielt werden dürfen.

1.2 Schriftlicher Selbstaussfüller zum gegenwärtigen Lebensgefühl

Auch im Rahmen der dritten Erhebungswelle von lidA wurden Fragen zum gegenwärtigen Lebensgefühl gestellt. Aus Gründen der Vertraulichkeit wurde bei diesen Fragen ein schriftlicher Selbstaussfüller eingesetzt, damit die Zielperson die entsprechenden Fragen ganz für sich beantworten konnte. Im Interview wurde vom Befragungsprogramm darauf hingewiesen, wann der Selbstaussfüllerfragebogen übergeben werden sollte. Für eine genaue Zuordnung des Selbstaussfüllerfragebogens zu einer Zielperson wurde von den Interviewerinnen und Interviewern auf dem schriftlichen Selbstaussfüller die laufende Nummer der Zielperson eingetragen. Nach Übergabe des Selbstaussfüllerfragebogens und eines Briefumschlags wurde der Bogen im Beisein der Interviewerin oder des Interviewers, aber ohne Ansicht der Antworten, von den befragten Personen ausgefüllt und anschließend in den Briefumschlag gesteckt, dieser wurde verschlossen.

Für eine genaue Zuordnung und Dokumentation des Selbstaussfüllerfragebogens wurde nach dem Erhalt des verschlossenen Umschlags durch die Interviewerin oder den Interviewer die laufende Nummer der Zielperson und die Interviewernummer auf dem Umschlag eingetragen und im Fragebogenprogramm vermerkt, dass der Briefumschlag direkt übergeben wurde. Die Zielperson hatte jedoch auch die Mög-

lichkeit, den Selbstausfüllerfragebogen mit einem beigefügten Rückumschlag selbst zurückzusenden, falls sie dies wünschte.

1.3 Erlaubnis zur Zuspiegelung von IEB-Daten und Rentenversicherungsdaten

Die lidA-Studie ist seit Beginn so konzipiert, dass zur Auswertung von im Interview erhobenen Befragungsdaten auch Informationen aus anderen, bereits bestehenden Datenbeständen (zu Arbeitsleben und Gesundheit) einbezogen werden. Neu in der dritten Welle war nun die ergänzende Zuspiegelung von Daten der Deutschen Rentenversicherung Bund. Um Veränderungen der beruflichen Anforderungen und der Gesundheit im Lauf des Erwerbslebens über einen längeren Zeitraum hinweg beurteilen zu können, sollen diese Daten der Deutschen Rentenversicherung Bund wichtige Informationen beibringen. Dabei geht es auch um Leistungen, die die Rentenversicherung oft schon in der Erwerbsphase erbringt, wie zum Beispiel medizinische oder berufliche Rehabilitationsmaßnahmen.

Datenschutzrechtlich wird für die Zuspiegelung dieser Daten an die Befragungsdaten das Einverständnis der Zielperson vorausgesetzt. Dabei werden im Rahmen der lidA-Studie stets schriftliche Einverständniserklärungen bei den Befragten eingeholt. Form und Inhalt dieser Einwilligungserklärungen werden jeweils durch die beteiligten Datenschützer des IAB – und in der dritten Welle nun auch der Deutschen Rentenversicherung Bund – vorgegeben. Der Zustimmungsblock mit zwei schriftlichen Bereitschaftserklärungen stellt dabei eine hohe Anforderung an die Befragungspersonen dar. Eine besondere Herausforderung war, dass auf der Einverständniserklärung für die Deutsche Rentenversicherung Bund die Sozialversicherungsnummer eingetragen werden sollte. Die Erfassung der Sozialversicherungsnummer bei Survey-Befragten kann aus methodisch-sozialwissenschaftlicher Sicht keineswegs als optimaler Weg angesehen werden. Viele Befragte wissen – auch auf Basis einer Erläuterung – nicht genau, um was es sich bei dieser Nummer handelt. Ebenso wissen viele Befragte nicht, wo sie eine solche Nummer aufbewahrt haben oder finden können. Auch hier wirken erläuternde Hilfestellungen, die im „Zustimmungsbock“ durch die Interviewer und Interviewerinnen gegeben werden, nur bedingt weiter. Denn eine Herausgabe der eigenen Sozialversicherungsnummer gilt aus datenschutzrechtlicher Sicht als bedenklich und sensibel. Und ein solches Argument kommt bei den Befragten auch gerade dann zum Tragen, wenn vorher erläutert wurde, auf welchen Dokumenten eine Sozialversicherungsnummer in der Regel auch zu finden ist.

Dass bei der lidA-Welle 3 für die dort eingesetzte Einverständniserklärung zur Zuspiegelung von Daten der Deutschen Rentenversicherung Bund die Sozialversicherungsnummer vollständig erfasst werden sollte, war eine datenschutzrechtliche Vorgabe der Deutschen Rentenversicherung Bund. Daher musste dies in der Erhebung so umgesetzt werden.

2 Panelpflege und Einsatzstichprobe

In der zweiten Welle der lidA-Haupterhebung waren im Januar 2014 **n=5.605** Adressen panelbereiter Personen der ersten lidA-Welle im Feld eingesetzt worden.

Am Ende der zweiten Welle wurden n=4.249 Interviews (knapp 76 Prozent der Bruttostichprobe) realisiert. Dabei hatten 805 Personen – oder 14,7 Prozent der Bruttostichprobe – in der zweiten Welle ihre Teilnahme an der lidA-Befragung verweigert. Diese Fälle stellten temporäre Ausfälle dar.

n=297 Zielpersonen oder gut 5 Prozent der gesamten Bruttostichprobe hatten indes im Verlauf der zweiten Welle die Panelbereitschaft zurückgezogen und gegenüber infas eine Adresslöschung verfügt. Diese 297 Fälle fielen nach der zweiten lidA-Welle aus dem Panelbestand heraus. Entsprechend den Regeln des Datenschutzes wurden die Adressen dieser Zielpersonen nach Feldende unwiderruflich bei infas gelöscht. Und weitere **n=16** Personen mussten zum Ende der Welle 2 als verstorben gelistet werden. Es verblieb somit als Panelbestand ein Umfang von **n=5.292 Fällen**.

2015 mussten bei infas sämtliche Daten der lidA-Studie aus vertraglichen Gründen unwiderruflich gelöscht werden. infas trug aber dafür Sorge, dass zum 21.10.2015 ein Archiv mit 3 Dateien (Panelstatus für 5.292 Fälle, Stichprobeninformationen, Zusatzmerkmale) an das Forschungsdatenzentrum des Instituts für Arbeitsmarkt- und Berufsforschung (IAB) – als der datenschutzrechtlich verantwortlichen Stelle – übergeben wurde. Die Weiterführung des lidA-Projekts war zu diesem Zeitpunkt eine offene Frage. Die kontrollierte und detailliert dokumentierte Übergabe zwischen infas und IAB sollte indes einen möglichen Fortgang des Projekts zu einem späteren Zeitpunkt sichern helfen. In Folge dieser Löschung konnte dann die unter methodischen Gesichtspunkten notwendige Panelpflege nicht im Jahr 2016 stattfinden.

Nach einem erneuten Antrag auf Übermittlung von Sozialdaten nach § 75 SGB X (auf Basis einer Weiterfinanzierung des lidA-Projekts) wurde der lidA-Datenbestand im Sommer 2017 vom IAB vertragsbasiert wieder an infas zurückgegeben. Im Zeitraum der gut eineinhalb Jahre zwischen Herbst 2015 und Sommer 2017 hatte sich, wie zu diesem Zeitpunkt bekannt, ein Verlust von **n=7 Fällen** aus dem Panelbestand aufgrund von Todesfällen ergeben. infas erhielt somit 2017 einen Panelbestand von **n=5.285 Fällen** vom IAB zurück.

Diese 5.285 Adressen wurden unmittelbar im Herbst 2017 in einer Panelpflege eingesetzt. Kombiniert wurde die Sendung mit einer Adressänderungskarte und einem portofreien Rückumschlag. Den Zielpersonen wurde Dank für die bisherige Teilnahme an der lidA-Studie mitgeteilt, die kommende Befragungswelle 2018 vor angekündigt und ferner die Bitte ausgesprochen, dass – falls sich in der Zwischenzeit Änderungen bei der Adresse ergeben haben sollten – diese infas mitzuteilen. Adressänderungen konnten auch über eine kostenfreie Rufnummer mitgeteilt oder an einen studieneigenen E-Mail-Account gesendet werden.

Tabelle 1 Ergebnis der lidA-Panelpflege Herbst/Winter 2017/2018

Rücklaufcodes	Panelbereit		Verstorben/ Auslandsfälle		Rückzug Panelbereitschaft		Gesamt	
	N	%	n	%	n	%	n	%
Keine Rückmeldung	4.878	92,3	-	-	-	-	4.878	92,3
Grundsätzliche Verweigerung	-	-	-	-	39	0,7	39	0,7
Verstorben	-	-	16	0,3	-	-	16	0,3
Unbekannt verzogen	79	1,5	-	-	-	-	79	1,5
Ins Ausland verzogen	-	-	3	0,1	-	-	3	0,1
Mitteilung Adressänderung	221	4,2	-	-	-	-	221	4,2
Aktive Bestätigung der Adresse	49	1,0	-	-	-	-	49	1,0
Gesamt	5.227	99,0	19	0,3	39	0,7	5.285	100,0

Quelle: infas-Sample-Management-System, Stand: 22.03.2018

Im Rahmen der Panelpflege gingen Rückmeldungen von n=407 Zielpersonen ein. Davon teilten n=221 Panelteilnehmerinnen und -teilnehmer eine neue Adresse mit. Weitere n=49 Personen bestätigten ihre Adresse aktiv. n=39 Personen zogen ihre weitere Bereitschaft zur Teilnahme an der lidA-Studie zurück. 16 weitere Personen waren verstorben, 3 Personen ins Ausland verzogen. Da die **n=39** grundlegend verweigernden Personen sowie die **n=19** verstorbenen Personen und die Auslandsfälle abgezogen werden mussten, verblieben im Nachgang der Panelpflege **n=5.227** Panelfälle im Bestand.

Im Rahmen des **Pretests** für die Haupterhebung der lida-Welle 3 mussten **n=100** dieser Adressen bereits vor der Hauptstudie eingesetzt werden, da nur so der beauftragte und methodisch gebotene Test mit Preload-Informationen überhaupt möglich war.¹ Auf dieser Basis gingen schließlich **n=5.127** Adressen **panelbereiter Zielpersonen** in die Einsatzstichprobe der Haupterhebung der lidA-Welle 3 ein.

¹ Im Pretest wurden dabei n=38 Interviews realisiert, die als vollwertige Fälle später mitgezählt werden können, da zwischen Pretest und Haupterhebung der Welle 3 kaum Änderungen am Fragebogen vorgenommen wurden. Es wurden weitestgehend die gleichen Fragen beantwortet wie im Interview der Haupterhebung der Welle 3.

3 Parameter für die Felddurchführung

3.1 Persönliches Anschreiben und Incentive

Alle Zielpersonen der Einsatzstichprobe wurden eine Woche vor dem Feldstart durch ein persönliches Anschreiben über die Studie, über die erneute persönliche Kontaktaufnahme durch einen Interviewer oder eine Interviewerin sowie über die Einhaltung des Datenschutzes informiert. Im Anschreiben wurde auch ein Incentive angekündigt. Gegenüber der zweiten Befragungswelle war das Incentive nun erhöht worden, da die Befragungsdauer höher ausfiel. Für ein vollständig abgeschlossenes Interview erhielten die Zielpersonen in der dritten Welle zusammen mit einem Dankeschreiben 20 Euro in bar per Post zugesandt.

3.2 Identifizierung der richtigen Zielperson

Möglichst alle Panelteilnehmerinnen und -teilnehmer aus der ersten Welle sollten auch in der dritten Welle wieder befragt werden. Dazu war es erforderlich, dass vor Ort überprüft wurde, dass das Interview mit der korrekten Zielperson (i.e. der panelbereiten Person) durchgeführt wird. Für die genaue Identifikation der Zielperson wurden der Name, der Vorname, die Adresse und das Geburtsdatum abgeglichen. Zur Unterstützung waren die aus der Stichprobe vorliegenden Informationen zu Geschlecht und Geburtsjahr der Zielperson im Programm hinterlegt.

3.3 Umgang mit verzogenen Personen

Sofern Zielpersonen an der Kontaktadresse vorübergehend oder dauerhaft nicht anzutreffen waren, sollten die Interviewerinnen und Interviewer, so möglich, Informationen zur neuen Adresse im elektronischen Kontaktprotokoll erfassen. Sofern die Zielperson noch in derselben Gemeinde wohnhaft war, wurde die Zielperson vom selben Interviewer oder derselben Interviewerin unter der neuen Adresse kontaktiert. Wenn die Zielperson weiter weg verzogen war, wurde die Adresse an eine dort eingesetzte Interviewerin oder Interviewer übergeben.

3.4 Nutzung von Preload-Informationen

Im Befragungsprogramm sollten Informationen aus den vorhergehenden Befragungswellen genutzt werden, um den Verlauf des Interviews steuern oder verkürzen zu helfen. Beispielsweise war für Zielpersonen aus der zweiten Welle der Ausbildungsabschluss bekannt oder die Art der Erwerbstätigkeit oder auch die berufliche Stellung zum Zeitpunkt des letzten Interviews im Jahr 2011 oder 2014. Solche Informationen wurden nun in der dritten Welle bei der Formulierung der Fragen genutzt. Auf entsprechend hinterlegte Informationen wurde im CAPI-Programm fallweise gesteuert „still“ zurückgegriffen.

Beispielhaft sei dies an der Frage nach der beruflichen Bildung erläutert. Dazu wurde den Befragten der im letzten Interview genannte höchste Ausbildungsabschluss genannt. Anschließend wurde gefragt, ob seit dem letzten Interview ein weiterer höherer Ausbildungsabschluss erworben worden sei. Sollte dies der Fall sein, konnten in

einer Folgefrage die aktuellen Angaben erfasst werden. Auch war möglich, dass Befragte die Angaben aus der ersten oder zweiten Befragungswelle korrigierten.

Bei diesen sogenannten Preload-Informationen wurden individuelle Zeitbezüge verwendet. Bei Fragen nach Veränderungen wurde dabei nicht ein allgemein gültiger Zeitpunkt gewählt, sondern das Datum des letzten Interviews eingeblendet. Grundsätzlich erleichtert die Nutzung von Preload-Informationen die Befragung und verkürzt sie an entsprechenden Stellen auch. Viele Fragen müssen nicht erneut gestellt werden, da die befragte Person eine vorliegende unveränderte Information einfach bestätigen kann.

3.5 Übersicht über die Erhebungsmaterialien

Für die Durchführung eines lidA-Interviews war in der dritten Welle ein umfangreiches Set an Erhebungsmaterialien erforderlich.

Tabelle 2 Studienmaterial

Für jede Zielperson

1 Selbstausfüllerfragebogen

1 Rückumschlag für den Selbstausfüller

1 Teilnehmerinformation

1 Einverständniserklärung zur Zuspiegelung ausgewählter Registerdaten

1 Einverständniserklärung zur Zuspiegelung ausgewählter Rentenversicherungsdaten

Für die Befragung

1 Adressliste

1 Listenheft

1 Rücklaufcode-Karte

1 Interviewerhandbuch

Per E-Mail/USB-Stick

Befragungsprogramm für das Interview inkl. Testversion

4 Feldergebnisse

4.1 Dokumentation der Kontakte im Feld

Für eine genaue Dokumentation des Bearbeitungsstands der Kontakte führten die Interviewerinnen und Interviewer ein elektronisches Kontaktprotokoll auf ihrem Laptop. Der aktuelle Bearbeitungsstand für jede sich in Arbeit befindende Adresse wurde im Kontaktprotokoll mithilfe einer umfassenden Schlüsselliste dokumentiert.

Während der Feldzeit waren die Interviewerinnen und Interviewer angewiesen, die Kontaktprotokolldateien mit den aktuellen Feldständen für alle bearbeiteten Adressen und Datensätze realisierter Interviews regelmäßig zu übermitteln. Die Feldbearbeitungsstände wurden zentral erfasst und konnten für jeden Sample Point zeitnah für steuernde Maßnahmen und den Interviewereinsatz in den Stichprobenpoints genutzt werden.

Über den Feldstand wurde die Universität Wuppertal während des laufenden Feldes unterrichtet. Neben dem Bearbeitungsstand wurden mit den Reportings auch Informationen zu den im Interview mündlich erteilten Einwilligungen zur Zuspiegelung von Register- bzw. Rentenversicherungsdaten sowie zur Interviewdauer übermittelt.

4.2 Feldbegleitendes Tracking

Für Zielpersonen, die aufgrund fehlerhafter Adressen nicht erreicht werden konnten, wurden während der Feldzeit Adressrecherchen durchgeführt. Diese sogenannten Trackingmaßnahmen wurden nacheinander eingeleitet, bis eine neue Adressinformation gefunden wurde. Zunächst erfolgte eine Recherche in der Addressfactory der Deutschen Post AG. Sofern auf diesem Weg keine Adressaktualisierung möglich war, wurde anschließend die Anfrage an das jeweilige Einwohnermeldeamt gesendet, zu dem die letzte bekannte Anschrift gehörte.

Während der Feldzeit führten zudem auch die Interviewerinnen und Interviewer vor Ort selbstständig Adressrecherchen durch. Darüber hinaus kontaktierten einige Zielpersonen selbst die infas-Hotline oder nutzten die studienspezifische E-Mail-Adresse, um Adressänderungen oder neue Telefonnummern anzugeben.

Sobald mithilfe der Recherchemaßnahmen oder aufgrund einer Rückmeldung über die infas-Hotline neue Telefonnummern oder andere Adressinformationen in Erfahrung gebracht werden konnten, wurden auch diese sukzessiv an das Feld übergeben.

Insgesamt wurden mit diesen Verfahren 99 Adressen im Feldverlauf getrackt. Davon war in 2 Fällen keine Adresse ermittelbar, in 34 Fällen wurde die bestehende Adresse bestätigt und in 63 Fällen konnte eine neue Adresse eingeholt werden.

4.3 Realisierte Interviews

In der dritten lidA-Haupterhebung wurden innerhalb der 17-wöchigen Feldzeit $n=5.127$ Adressen panelbereiter Zielpersonen intensiv im Feld bearbeitet. Zwischen dem 11.04.2018 und dem 31.07.2018 wurden $n=3.551$ Interviews realisiert. Dies

sind annähernd 70 Prozent der Bruttostichprobe. Aufgrund der nach Feldende durchgeführten detaillierten Datenprüfung sowie der Ergebnisse der Interviewerkontrolle mussten lediglich n=3 realisierte Interviews als nicht auswertbar klassifiziert werden. Somit besteht das Feldergebnis in **n=3.548** auswertbaren Interviews.

Damit liegt für die dritte Welle der lidA-Studie ein gutes Feldergebnis vor. Im Zusammenhang mit dieser Bewertung sind zwei das Feldergebnis beeinflussende Faktoren zu benennen. Zum einen ist in Rechnung zu stellen, dass in Folge der Löschung aller lidA-Daten im Herbst 2015 die Panelpflege erst viel später als eigentlich vorgesehen hatte stattfinden können. Dies trug insgesamt zu einer Schwächung der Panelbindung bei. Zum anderen hat die während der Feldzeit mit großem Medienecho in Kraft getretene EU-Datenschutzgrundverordnung das Feldergebnis negativ beeinflusst. Es ist nicht allein der Anteil aller durch „Verweigerung“ ausgelösten Ausfälle, der in Welle 3 höher ausfällt als in Welle 2. In besonderem Maß fällt in Welle 3 der Anteil jener Personen ins Auge, die im Frühsommer 2018 eine grundsätzliche Ablehnung aus Datenschutzgründen aussprachen, dies auch verbunden mit dem Rückzug der Panelbereitschaft und der sofortigen Aufforderung, die Adresse vollständig zu löschen. Der Anteil der Personen, die genau solche Ablehnungen aussprachen, lag in der dritten lidA-Welle mit 8,8 Prozent deutlich höher als in Welle 2 (gut 5 Prozent).

Tabelle 3 Stichprobenausschöpfung lidA-Welle 3

Response Code	Letzter Kontakt/Feldstatus (AAPOR-Gruppierung)	Bruttostichprobe		Bereinigte Einsatzstichprobe		
		abs.	%	%	abs.	
Nicht Zielgruppe (NE)						
14	Unbewohnt, Gebäude verfallen, abgerissen (NE)	5	0,10	-	135	
15	ZP verstorben (NE)	27	0,53	-		
33	ZP ins Ausland verzogen (NE)	8	0,16	-		
35	ZP wohnt da nicht mehr/neue Anschrift unbekannt (NE)	95	1,85	-		
Systematische Ausfälle (NR)						
1	Nicht erreicht, kein Kontakt (NR-NC)	73	1,42	1,46	452	
43	Nicht befragbar/dauerhaft krank oder behindert (NR-NC)	11	0,21	0,22		
80	Keine Verständigung möglich/ spricht nicht ausreichend Deutsch (NR-O)	8	0,16	0,16		
6	Vager Termin, in Feldzeit nicht realisiert (NR-O)	170	3,32	3,41		
7	Definitiver Termin, in Feldzeit nicht realisiert (NR-O)	16	0,31	0,32		
42	In Feldzeit nicht zu erreichen (NR-O)	174	3,39	3,49		
9	ZP verweigert: keine Zeit/dauert zu lange (NR-R)	134	2,61	2,68		
11	ZP verweigert krank (NR-R)	27	0,53	0,54		
40	KP verweigert neue Anschrift (NR-R)	5	0,10	0,10		
52	Zugang zu ZP verhindert/Teilnahme untersagt/ lt. Auskunft nicht bereit (NR-R)	34	0,66	0,68		
53	KP verweigert jegliche Auskunft (NR-R)	10	0,20	0,20		
54	ZP verweigert: kein Interesse /Thema (NR-R)	45	0,88	0,90		
59	ZP verweigert: Datenschutzgründe / zu persönlich (NR-R)	3	0,06	0,06		
62	ZP verweigert: sonstige Gründe (NR-R)	44	0,86	0,88		
71	ZP verweigert: nicht in dieser Welle (temporärer Ausfall)	247	4,82	4,95		
8	Adresse löschen/ZP verweigert grundsätzlich/ Panelbereitschaft zurückgezogen (NR-R)	437	8,52	8,75		
47	Abbruch im Fragebogen	3	0,06	0,06		989
Interview						
18	Interview realisiert, davon	3.551	69,26	71,13		3.551
	<i>Gültig realisiert</i>	<i>3.548</i>	<i>69,20</i>	<i>71,07</i>		
	<i>Nicht auswertbar</i>	<i>3</i>	<i>0,06</i>	<i>0,06</i>		
Summe						
		5.127	100,00	100,00	4.992	

AAPOR-Gruppen: NC: Non-Contact; NE: not eligible; NR: eligible, nonresponse; R=Refusal; O=Other; ZP: Zielperson;
 KP: Kontaktperson

Im Übrigen stellt sich unter den systematischen Ausfällen der Welle 3 ein breites Spektrum von Ausfallgründen dar, das nach Umfang und Struktur im Vergleich zu anderen Paneluntersuchungen ein ähnliches Muster aufweist. Der Anteil nicht erreichter Personen belief sich zum Ende der Feldzeit der Welle 3 auf lediglich rund 1,5 Prozent (entspricht 73 Personen).

4.4 Interviewereinsatz und Kontakthäufigkeit

In der dritten Erhebungswelle wurden insgesamt 206 Interviewerinnen und Interviewer eingesetzt. Im Durchschnitt realisierte jede interviewende Person 17,2 Interviews. Die Anzahl der realisierten Interviews pro Interviewerin und Interviewer variiert in Abhängigkeit von der Anzahl der zu bearbeitenden Adressen in den jeweiligen Points und der Teilnahmebereitschaft der Zielperson. Rund 70 Prozent aller Interviewerinnen und Interviewer realisierten weniger als 20 Interviews. Nur wenige Interviewerinnen und Interviewer wurden aus feldstrategischen Gründen in mehreren Points eingesetzt und realisierten so eine höhere Zahl an Interviews.

Tabelle 4 Interviews pro Interviewerin und Interviewer

	Interviewerinnen und Interviewer		Interviews	
	<i>n</i>	%	<i>n</i>	%
1-5 Interviews	15	7,3	48	1,4
6-10 Interviews	46	22,3	396	11,2
11-20 Interviews	85	41,3	1.289	36,3
21-30 Interviews	38	18,4	933	26,3
31-40 Interviews	13	6,3	451	12,7
41-50 Interviews	6	2,9	263	7,4
51-60 Interviews	2	1,0	109	3,1
61-70 Interviews	1	0,5	62	1,7
Gesamt	206	100,0	3.551	100,0

Basis: im Feld realisierte Interviews

Bis zur Realisierung eines Interviews waren im Feld durchschnittlich 3,8 Kontakte erforderlich. Bei der Auswertung der Kontakthäufigkeit zeigt sich eine besonders hohe Streuung. Während viele Interviews bereits beim ersten Kontakt realisiert werden konnten, waren in anderen Fällen bis zu 47 Kontakte für die Durchführung notwendig. Die nicht realisierten Fälle weisen eine noch höhere durchschnittliche Kontakthäufigkeit auf. Dies zeigt die intensive Feldbearbeitung. Relevant ist hier auch der Blick auf zwei Gruppen: die „nicht erreichten“ Zielpersonen sowie jene Personen, die während der Feldverlaufs immer nur vage einer Interviewteilnahme zusagten. Für beide Segmente liegen die durchschnittlichen Kontakthäufigkeiten mit 5,0 und 6,3 jeweils noch höher, was die beharrliche Bearbeitung der Interviewerinnen und Interviewer dokumentiert.

Tabelle 5 Kontakthäufigkeit

	n	Minimum	Maximum	Mittelwert	Standardabweichung
Realisierte Interviews**)	3.551	1	47	3,8	2,35
Kein Interview realisiert, aber mindestens 1 Feldkontakt durch interviewende Person, davon:	1.561	1	37	4,2	3,68
– nicht erreichte Zielpersonen	73	1	29	5,0	3,75
– vage Termine	170	1	30	6,3	4,45
Gesamt (mindestens 1 Feldkontakt durch Interviewer*)	5.112	1	47	4,0	2,83

*) Basis: bearbeitete Fälle mit mindestens 1 Feldkontakt durch Interviewerinnen und Interviewer.

**) Berücksichtigt sind die im Feld realisierten Interviews.

4.5 Interviewdauer

Die Interviewdauer in der dritten Erhebungswelle war mit 75 Minuten angesetzt. Die Messung der tatsächlichen durchschnittlichen Dauer von 78 Minuten wies auf einen akzeptablen Korridor für die Erhebung hin.

Tabelle 6 Interviewdauer modulweise

Angaben in Minuten, Basis: Zeitstempel (Sekunden)	n (*)	Minimum (**)	Maximum	Mittelwert	Standardabweichung
1 Intro	3.535	0,0	72,1	1,0	3,29
2 Feststellung Erwerbstätigkeit	3.535	0,0	62,2	1,1	2,63
3 Erwerbstätige	3.292	0,1	63,2	11,8	5,29
9 Blick auf den Ruhestand	3.535	0,0	46,2	8,4	3,93
8 Arbeitsplatz	3.292	0,9	101,7	32,6	10,29
4 Nicht Erwerbstätige	157	1,9	34,7	7,5	3,99
7 Rentner	160	1,6	48,9	13,4	7,47
5 Gesundheit	3.535	1,7	88,0	8,3	3,96
Selbstaufüller	3.535	0,0	36,5	3,0	2,14
6 Personenfragen	3.535	1,8	109,7	8,6	8,04
Zustimmungsblock	3.535	0,0	175,3	4,6	7,03
Interviewerfragen	3.520	0,0	76,8	0,8	2,52
Interviewdauer gesamt ohne Interviewerfragen	3.540	20,4	236,1	77,3	23,39
Interviewdauer gesamt inkl. Interviewerfragen	3.520	20,6	237,1	78,0	23,58

(*) Basis: n=3.548 gültig realisierte Fälle, n=8 gültig realisierte Fälle nicht in Dauerberechnung berücksichtigt, da un-plausible Zeiten wg. Unterbrechung oder negativer Dauern.

(**) Fälle mit unplausibel kurzen Minimalzeiten in einzelnen Modulen ergeben sich durch Zurück- und schnelles Wieder-Vorblättern im Fragenprogramm, wodurch ein Zeitstempel überschrieben oder mit negativer Dauer ausgewiesen wird.

Quelle: Zeitstempel in CAPI-Interviews, eigene Berechnungen

Das Modul zur beruflichen Situation („8 Arbeitsplatz“) war im Interview mit durchschnittlich 32 Minuten der längste Erhebungsblock. An zweiter Stelle folgte mit durchschnittlich 13 Minuten der Befragungsteil zum Thema Rente. Die Klärung der Einwilligungen zur Zuspiegelung externer Datenquellen nahm durchschnittlich knapp 5 Minuten Erhebungszeit ein.

4.6 Rücklauf beim Selbstaufüller zum gegenwärtige Lebensgefühl

Laut Angabe im Interview wurde der Selbstaufüller zur Depressivität mit Fragen zum gegenwärtigen Lebensgefühl an nahezu alle Befragungspersonen übergeben (99,1 Prozent).

Wie bereits in der ersten und zweiten Welle wurde der Fragebogen von der Zielperson in einem verschlossenen Rückumschlag an die Interviewerinnen und Interviewer zurückgegeben. Eine Prüfung auf das tatsächliche Ausfüllen des Selbstaufüllers fand somit erst bei der Erfassung der Daten statt. Auch in der dritten Erhebungswelle bestand die Möglichkeit, dass die Zielperson die Fragen im Nachgang ohne das Beisein der interviewenden Person ausfüllt und den Fragebogen im vorfrankierten Rückumschlag an infas schickt. Nicht in allen Fällen wurde der Fragebogen dann auch tatsächlich per Post an infas zurückgeschickt.

Insgesamt liegen zum Feldende für n=3.117 Fälle ein ausgefüllter Papierfragebogen vor. Bezogen auf die 3.548 gültig realisierten CAPI-Interviews liegt die Quote damit bei 87,9 Prozent. Angesichts des sensiblen Themas und des Verfahrens stellt dies auch in der dritten Erhebungswelle ein gutes Ergebnis dar.

Tabelle 7 Vorliegen eines Selbstaufüllers

	Ja	Nein	Gesamt	Gesamt
	%	%	%	n
Selbstaufüller während des Interviews zur Bearbeitung übergeben/ Umschlag angenommen	99,1	0,9	100,0	3.548
Ausgefüllte Selbstaufüller nach Feldende	87,9	12,1	100,0	3.548

Basis: gültig realisierte Fälle/Quelle: Rücklauferfassungssystem infas

4.7 Zuspiegelungsbereitschaft für Registerdaten und Rentenversicherungsdaten

Im Interview wurden die Zielpersonen um ihr Einverständnis zur Zuspiegelung von Informationen aus bereits bestehenden Datenbeständen gebeten. Dies betraf ausgewählte Registerdaten, die beim Institut für Arbeitsmarkt- und Berufsforschung (IAB) der Bundesagentur für Arbeit in Nürnberg vorliegen, sowie auch Versicherungsdaten der Deutschen Rentenversicherung Bund.

Im Interview gaben knapp 84 Prozent der Befragten eine mündliche Einwilligung zur Zuspiegelung von DRV-Daten an Befragungsdaten. Der Anteil schriftlich unterzeichneter Einwilligungen mit Bezug auf die Rentendaten fiel rund 7 Prozentpunkte geringer aus (knapp 77 Prozent).

Eine mündliche Einwilligung im Interview zur Zuspiegelung der Registerdaten der Bundesagentur für Arbeit gaben knapp 89 Prozent der Befragten, die relevante schriftliche Erlaubnis fiel mit gut 72 Prozent deutlich geringer aus.

Tabelle 8 Einwilligungen zur Zuspiegelung externer Daten an Befragungsdaten

	Einwilligung	Keine Einwilligung	Gesamt	
	%	%	%	<i>n</i>
Mündliche Zustimmung zur Zuspiegelung von DRV-Daten an Befragungsdaten	83,9	16,1	100,0	3.548
Schriftliche Zustimmung zur Zuspiegelung von DRV-Daten an Befragungsdaten	76,7	23,3	100,0	3.548
Mündliche Zustimmung zur Zuspiegelung von BA-Registerdaten an Befragungsdaten	88,9	11,1	100,0	3.548
Schriftliche Zustimmung zur Zuspiegelung von BA-Registerdaten an Befragungsdaten	72,2	27,8	100,0	3.548

Basis: gültig realisierte Fälle/Quelle: Rückläuferfassungssystem infas

Die Erfahrung aus verschiedenen Erhebungen zeigt, dass das Unterschreiben von Formularen im Rahmen eines persönlich-mündlichen Interviews grundsätzlich eine Schwierigkeit darstellt. Dies zeigt sich bereits bei dem einfachen Quittieren des Erhalts von Incentives, die in der Interviewsituation übergeben werden. Grundsätzlich bestehen Vorbehalte, eine Unterschrift zu leisten; und gerade im Rahmen einer durch eine andere Person (die Interviewerinnen und Interviewer) gelenkten Kommunikationssituation ist dies besonders virulent. Dabei bestehen auch Befürchtungen vor Bevormundung, ferner wirkt die Einschätzung der Betroffenen, einen Sachverhalt, den man unterschreiben soll, „jetzt“ in der Situation nicht angemessen genau überprüfen zu können. Die Hürde ist noch deutlich höher, wenn es darum geht, komplexe Formulare zu unterschreiben, mit denen die Zielperson der Zuspiegelung von ihr in der Regel überhaupt nicht bekannten oder nicht nachvollziehbaren Datenbeständen zustimmen soll. Diese Aspekte erklären die Differenz zwischen mündlichen und schriftlichen Zustimmungsraten. In der Feldzeit der lidA-Welle 3 ergab sich dabei noch zusätzlich eine Verschärfung durch das Inkrafttreten der EU-Datenschutzgrundverordnung ab Mai 2018, die eine massive Verunsicherung der befragten Bürgerinnen und Bürger auslöste. Dass vor diesem Hintergrund die erzielten Zustimmungsraten gut ausfielen (sie übertrafen dabei die Zustimmungsraten aus Welle 1), ist durchaus positiv zu bewerten.

Neben dem Unterschreiben der Formulare stellte insbesondere das von der Deutschen Rentenversicherung geforderte Eintragen der Sozialversicherungsnummer auf der Einwilligung zur Zuspiegelung der Rentenversicherungsdaten eine erheblich hohe Hürde dar. Die Bürgerinnen und Bürger wissen vielfach nicht, wo sie eine solche Nummer aufbewahrt haben. Dazu kamen häufig auch Vorbehalte, die Nummer überhaupt anzugeben. Für den Eintrag waren auf dem Formular zwei getrennte Felder vorgesehen, je nachdem, ob die Zielperson Rente bezog oder nicht. Erwerbstätige Personen sollten ihre vollständige Sozialversicherungsnummer eintragen, Rentner sollten die auf dem Rentenbescheid aufgeführte Versicherungsnummer vollständig angeben.

Im Ergebnis lagen n=2.722 unterschriebene Einverständniserklärungen zur Zuspie- lung der Rentenversicherungsdaten vor. Davon wurde in 974 Fällen (35,8 Prozent) ein Einverständnis zur Zuspie- lung gegeben ohne dass eine Sozialversicherungsnum- mer eingetragen wurde. In 1.748 Fällen (64,2 Prozent) wurde eine Versicherungs- nummer angegeben.

Im Zuge der Datenprüfung wurde bei infas (unter Nutzung eines speziell dafür entwi- ckelten Makros) untersucht, ob die Zeichenfolge der eingetragenen Nummern auf eine korrekte Sozialversicherungsnummer hinwies. Die Prüfung ergab, dass rund 9 Prozent der Angaben nicht dem gültigen Format entsprechen.² Somit verbleiben 1.587 Fälle mit gültiger Zeichenfolge bei den angegebenen Sozialversicherungsnum- mern.

² Die Versicherungsnummer besteht aus 11 Ziffern und einem Buchstaben, also insgesamt aus 12 Zeichen. Der Buchstabe steht immer an der 9. Stelle. Für alle unterschriebenen Einverständniserklärungen mit eingetragener Versicherungs- nummer in einem der vorgesehenen Felder wurde geprüft, ob die eingetragene Zeichenfolge diesem Format entspricht.

5 Datenprüfung und -lieferung

5.1 Interviewerkontrolle

Die Definition eines gültigen realisierten Interviews machte sich in lidA-Welle 3 auch an der Durchführung des Interviewgesprächs mit der richtigen Zielperson fest. Zusätzlich zur laufenden Kontrolle der Interviewerinnen und Interviewer über die Kontaktprotokolldatei wurde wiederum ein schriftliches Kontrollverfahren bei allen befragten Zielpersonen durchgeführt. Alle Befragten erhielten dabei zeitnah nach dem Interviewgespräch zusammen mit dem Dankschreiben einen kurzen Fragebogen (nebst Rückumschlag), in dem um eine Rückmeldung zum Interviewgespräch gebeten wurde. Gefragt wurde, ob sie von einem infas-Interviewer oder Interviewerin im Projekt „leben in der Arbeit“ befragt und ob das Interview persönlich durchgeführt worden war.

Tabelle 9 Ergebnis der Interviewerkontrolle

	<i>n</i>	%
Keine Rückmeldung gegeben	1.348	38,0
Korrekt durchgeführtes Interview (alles ok lt. Interviewerkontrolle)	2.202	62,0
Nicht korrekt durchgeführtes Interview	1	0,0
Gesamt	3.551	100,0

Quelle: Sample-Management-System, lidA-Interviewerkontrolle

3.551 Personen wurden angeschrieben, davon antworteten 2.202 (rd. 62 Prozent). Aus diesen nachträglichen Bestätigungen der Befragten geht hervor, dass die Interviews korrekt durchgeführt worden waren. In dem einen in der Interviewerkontrolle auffällig gemeldeten Fall hatte aufgrund einer Personenverwechslung kein Interview stattgefunden. Daher wurde ein Interview durch die richtige Zielperson dann auch nicht bestätigt.

Tabelle 10 Ergebnis der Datenprüfung und Auslieferung gültig realisierter Fälle

	<i>n</i>	%
Im Feld realisierte Interviews, davon	3.551	100,0
Gültig realisiert	3.548	99,9
Ungültig realisiert bzw. auffällig laut Interviewerkontrolle	3	0,0

Basis: im Feld realisierte Interviews

Zudem ergab die nachgehende Datenprüfung (Ableich von Befragungsdaten und Stichprobeninformationen), dass zusätzlich 2 weitere Fälle aus dem Datensatz zu entfernen waren. Hier handelte es sich um eine Abweichung von der Stichprobe (Befragung einer falschen Person) und im anderen Fall um ein technisches Problem eines Interviewers, das zu einer fehlerhaften Rückmeldung führte.

Zusammengenommen mussten auf Basis der Interviewerkontrolle und der Datenprüfung also 3 Interviews als nicht regelkonform bewertet werden. Der Datensatz umfasst somit 3.548 gültig realisierte Fälle.

5.2 Berufscodierung nach ISCO08 und KldB2010

In der lidA-Welle 3 wurde eine Berufscodierung nach ISCO08 und KldB2010 vorgenommen für alle Fälle, die gegenüber den letzten Erhebungswellen als Berufswechsler verlistet wurden. Insgesamt wurden 492 Nennungen codiert. Grundlage der Codierungen waren die offenen Nennungen zur beruflichen Tätigkeit. Für die Fälle aus dem Pretest wurde die Berufscodierung für insgesamt 6 Fälle nachgeholt und separat ausgeliefert.

Die Codierung verlief in drei Prozessschritten:

Zunächst wurde eine Codierungsunterlage erstellt, die auch Regelungen zur Codevergabe bei nicht eindeutigen Merkmalsausprägungen beinhaltetete.

Die Erstcodierung bestand aus zwei Teilschritten: einer maschinellen Vorcodierung sowie einer manuellen Vercodung. Die Daten wurden in ein elektronisches Codiersystem importiert und dort anhand eines umfassenden Wörterbuches maschinell vorcodiert. Rund 50 Prozent der Fälle konnten auf diese Weise maschinell codiert werden, alle übrigen Fälle wurden manuell codiert. Im Anschluss erfolgte eine kritische Prüfung der Fälle der maschinellen Vorcodierung auf inhaltliche Plausibilität. Als Hilfestellung wurden dabei auch zusätzliche Angaben aus der lidA-Erhebung zur beruflichen Stellung und zur Ausbildung (aus den Vorwellen) zugespielt.

Für alle Nennungen erfolgte schließlich eine blinde Zweitcodierung. Codeabweichungen zwischen Erst- und Zweitcodierung wurden von einem dritten Codierer entschieden. Wenige Problemfälle wurden in Diskussionsrunden besprochen und entschieden. In einer abschließenden Prüfung wurde die Richtigkeit der Codes, die Einhaltung der Codevergaberegeln und die Eingabe der Codes (Zahlendreher) geprüft.

5.3 Datenlieferungen

Nach Abschluss der Prüfungen wurden alle vollständigen Befragungsdatensätze an die Universitäten Wuppertal und Duisburg-Essen sowie an das Universitätsklinikum Düsseldorf übermittelt. Ferner wurden schriftliche Einwilligungserklärungen zur Zuspiegelung der BA-Registerdaten an die Universität Wuppertal übermittelt. Die schriftlichen Einwilligungserklärungen zur Zuspiegelung der Daten der Deutschen Rentenversicherung wurden an die Deutsche Rentenversicherung Bund übergeben. Alle Datenlieferungen und Dokumentübergaben wurden gemäß der für das lidA-Projekt geschlossenen datenschutzrechtlichen Vereinbarung vorgenommen.

CAPI-Befragungsdaten und Selbstaussfüller-Daten: Datensatz mit n=3.548 gültig realisierten Interviews (CAPI-Befragungsdaten). Im Rahmen dieses Befragungsdatensatzes sind die erfassten Informationen aus dem im Feld eingesetzten schriftlichen Selbstaussfüller zum gegenwärtigen Lebensgefühl in einem Umfang von n=3.117 Fällen enthalten.

Einwilligungserklärung BA-Registerdaten: Lieferbedingungen waren, dass ein Registereinwilligungsblatt unterschrieben vorlag, das Formular eindeutig einem Interviewfall zuzuordnen war, keine nachträgliche Verweigerung der Zielperson vorlag und dass das Formular vollständig ausgefüllt war. Ausgeliefert wurden somit 2.561 Formulare.

Tabelle 11 Schriftliche Einwilligungformulare BA-Registerdaten

Materialbeschaffenheit	n	%
Vollständiges Formular	2.561	98,16
Unvollständiges Formular	2	0,08
Leer	20	0,77
Verweigerung auf Formular	26	1,00
Gesamt	2.609	100,0

Quelle: infas-Rücklaufverwaltung

Einwilligungserklärung Rentenversicherungsdaten: Lieferbedingungen waren, dass ein Rentendateneinwilligungsblatt unterschrieben vorlag, das Formular eindeutig zuzuordnen war und keine nachträgliche Verweigerung der Zielperson vorlag. Ausgeliefert wurden 2.722 Formulare, darunter 1.587 Dokumente mit gültiger Zeichenfolge bei angegebenen Sozialversicherungsnummern und 1.135 Dokumente, die – ansonsten vollständig – diese Bedingung nicht erfüllten.³

Tabelle 12 Schriftliche Einwilligungformulare Rentenversicherungsdaten

Materialbeschaffenheit	n	%
Vollständiges Formular und gültige Zeichenfolge bei angegebenen Sozialversicherungsnummern	1.587	55,60
Vollständiges Formular	1.135	39,80
Unvollständiges Formular	64	2,24
Leer	37	1,30
Verweigerung auf Formular	31	1,06
Gesamt	2.854	100,0

Quelle: infas-Rücklaufverwaltung

³ Möglicherweise besteht bei der Deutschen Rentenversicherung eine Verwendungsmöglichkeit, auch ohne dass die Sozialversicherungsnummer angegeben ist.

6 Selektivitätsanalyse und Gewichtung

Die Gewichtung der dritten Erhebungswelle der lidA-Studie baut im Wesentlichen auf den Gewichtungen der ersten und zweiten Welle auf. Diese sind in den vorherigen Methodenberichten der Haupterhebung „lidA – leben in der Arbeit“ ausführlich dokumentiert. Dort wurden alle grundlegenden Schritte der Designgewichtung und Kalibrierung zur ersten Welle sowie die erstmalige Längsschnittgewichtung zur zweiten Welle beschrieben.

Die Grundgesamtheit der lidA-Erhebung bilden Beschäftigte in den Alterskohorten 1959 und 1965, gemäß Beschäftigten-Historik der Bundesagentur für Arbeit zum Stichtag 31.12.2009. Auf ihrer strukturellen Verteilung zum Stichtag beruhen die Querschnittgewichte der ersten Welle. Auch für die zweite und dritte Welle war die Verteilung der Grundgesamtheit mit diesem Stand maßgeblich und definiert die Zielverteilung für die aktuelle Gewichtung.

Für die Längsschnittgewichtung von der zweiten zur dritten Welle wurden wieder die Bleibewahrscheinlichkeiten für die wiederholte Teilnahme an der aktuellen Erhebung berechnet. Die Bleibewahrscheinlichkeit ist definiert aus der Wahrscheinlichkeit, an der dritten Welle teilzunehmen, gegeben letzte Teilnahme in Welle 2. Mit Berücksichtigung der Bleibewahrscheinlichkeiten können die Querschnittgewichte der ersten Welle um die systematisch entstandenen Ausfälle zur zweiten und dritten Welle adjustiert und mögliche Selektivitätseffekte über die Panelwellen kontrolliert werden.

Für die Bestimmung von Längsschnittgewichten musste ausgehend von den $n=4.244$ gültig realisierten Fällen der zweiten Welle abzüglich $n=79$ Fälle, die im Pretest eingesetzt wurden, ein Ausfallmodell erstellt werden. Insgesamt wurden für den Pretest $n=100$ Fälle eingesetzt, allerdings befanden sich darunter auch $n=21$ Fälle mit letzter Teilnahme in der ersten Welle (temporäre Ausfälle).

Für die Modellierung der Wahrscheinlichkeit, an der dritten Welle teilzunehmen, gegeben letzte Teilnahme in Welle 2 (Wiederholer), verbleiben dann $n=4.165$ panelbereite Fälle. Diese Datenbasis wurde für die Ausfallanalyse verwendet. Sie enthält im Wesentlichen die mit der Stichprobe gelieferten Individualmerkmale für jedes Element der Bruttostichprobe zum Ziehungszeitpunkt ergänzt mit Merkmalen aus den Befragungsdaten der ersten Welle.

6.1 Ausfallanalyse

Die Ausfallanalyse wurde mithilfe eines logistischen Regressionsmodells durchgeführt. Die abhängige Variable, Teilnahme an Welle 3, wurde dichotom mit den Ausprägungen $y=1$ (Teilnahme) und $y=0$ (Nichtteilnahme) codiert. Die Ergebnisse der Ausfallanalyse zeigen ähnlich wie bei der ersten und zweiten Welle eine nahezu selektionsfreie Realisierung der Stichprobe bezogen auf die in der Analyse verwendeten Merkmale. Nur geringfügige, größtenteils nicht signifikante Effekte konnten beobachtet werden, die einen moderaten Einfluss auf die Teilnahmewahrscheinlichkeit haben. In der folgenden Übersicht sind alle Parameter und Ergebnisse der logistischen Regressionsanalyse für die Teilnahme an der dritten Welle dargestellt.

Tabelle 13 Multiples logistisches Regressionsmodell zur Teilnahme von Wiederholungsbefragten an Welle 3

Merkmal/Ausprägung	Referenzkategorie	Odds Ratio	P-Wert
Geschlecht	Männlich		
Weiblich		1,161	0,139
Kohorte	1959		
1965		0,891	0,150
Geringfügige Beschäftigung	Sozialversicherungspflichtig beschäftigt		
Geringfügig beschäftigt		0,990	0,947
Stellung im Beruf und Arbeitszeit	Vollzeit		
Teilzeit		1,104	0,373
Tagesentgelt in Kategorien	Tagesentgelt unter 50 €		
Tagesentgelt 50 bis unter 85 €		1,058	0,652
Tagesentgelt 85 bis unter 120 €		1,125	0,403
Tagesentgelt 120 € und mehr		1,152	0,388
Ausbildung	Volks-, Haupt-, Realschule mit Ausbildung		
Volks-, Haupt-, Realschule ohne Ausbildung		0,756	0,130
Abitur ohne u. mit Ausbildung		1,417	0,005
Fachhochschulabschluss		1,303	0,091
Hochschulabschluss		1,295	0,081
Nationalität	Deutsch		
Nicht deutsch		0,894	0,601
Klassifikation der Berufe nach Blossfeld	Qualifizierte kaufmännische u. Verwaltungsberufe		
Agrarberufe		2,336	0,083
Einfache manuelle Berufe		0,888	0,461
Qualifizierte manuelle Berufe		0,876	0,400
Techniker		1,002	0,993
Ingenieure		0,854	0,473
Einfache Dienste		0,832	0,215
Qualifizierte Dienste		0,929	0,682
Semiprofessionen		1,041	0,768
Professionen		0,593	0,044
Einfache kaufmännische u. Verwaltungsberufe		1,047	0,768
Manager		1,016	0,944
Sonstige		1,279	0,468
Arbeitslosigkeitserfahrung	Nein		
Ja		1,011	0,932
Jahre in Beschäftigung/5		1,051	0,162
Anzahl der Beschäftigungsverhältnisse		0,990	0,669
Arbeitslosigkeit Dauer/100		0,977	0,068

Merkmal/Ausprägung	Referenzkategorie	Odds Ratio	P-Wert
Bundesland (Wohnort zum Zeitpunkt 1. Welle)	Nordrhein-Westfalen		
Schleswig-Holstein		1,666	0,124
Hamburg		0,316	0,002
Niedersachsen		1,183	0,255
Bremen		0,909	0,807
Hessen		0,734	0,035
Rheinland-Pfalz		0,628	0,017
Baden-Württemberg		0,895	0,377
Bayern		1,063	0,634
Saarland		5,977	0,085
Berlin		0,828	0,423
Brandenburg		1,303	0,332
Mecklenburg-Vorpommern		0,736	0,314
Sachsen		1,311	0,153
Sachsen-Anhalt		1,271	0,336
Thüringen		2,213	0,004
Gemeindegröße BIK 1-10 (des Wohnorts zum Zeitpunkt 1. Welle)	500.000 und mehr EW STyp 1 (Kernbereich)		
Unter 2.000 EW		2,395	0,010
2.000 bis u. 5.000 EW		1,213	0,241
5.000 bis u. 20.000 EW		1,224	0,168
20.000 bis u. 50.000 EW		1,210	0,236
50.000 bis u. 100.000 EW STyp 2/3/4 (Randbereich)		1,037	0,903
50.000 bis u. 100.000 EW STyp 1 (Kernbereich)		1,244	0,107
100.000 bis u. 500.000 EW STyp 2/3/4 (Randbereich)		0,994	0,964
100.000 bis u. 500.000 EW STyp 1 (Kernbereich)		1,099	0,530
500.000 und mehr EW STyp 2/3/4 (Randbereich)		2,395	0,010
Pseudo R² (McFadden)	0,0255		
Fallzahl	4.165		

Alle Prädiktoren sind Merkmale der Befragten aus Registerdaten der Bundesagentur für Arbeit zum Stichtag 31.12.2009.

6.2 Gewichtung

Da es sich bei Welle 3 um eine reine Panelstichprobe handelt, wurde für jeden Fall der Bruttostichprobe das bereits berechnete Längsschnittgewicht der Welle 2 als Ausgangsgewicht für die Längsschnittgewichtung der Welle 3 übernommen und, wie nachfolgend beschrieben, angepasst. Für die Adjustierung des Längsschnittgewichts der Welle 2 zum Längsschnittgewicht der Welle 3 wurde das inverse Produkt aus der resultierenden Teilnahmewahrscheinlichkeit des logistischen Regressionsmodells ($P_{\text{Teilnahme}}$) gebildet und mit dem Ausgangsgewicht ($\text{Längsschnittgewicht}_{w2}$) nach folgender Formel multipliziert und anschließend fallzahlnormiert:

$$\text{Längsschnittgewicht}_{w3} = \text{Längsschnittgewicht}_{w2} * (1/(P_{\text{Teilnahme}}))$$

Das neue Längsschnittgewicht passt die Verteilungen aufgrund der geringen Selektivität nur sehr moderat an und unterscheidet sich in seiner Varianz nur unwesentlich vom Ausgangsgewicht der zweiten Welle. In Folge des Gewichts resultieren Verteilungen der dritten Befragungswelle (nur für Wiederholer mit letzter Teilnahme in Welle 2), die nah an den gewichteten Merkmalen der ersten sowie zweiten Welle liegen (siehe dazu den Vergleich von ungewichteten und gewichteten Merkmalen über die Wellen in folgender Tabelle). Für den berechneten Längsschnittgewichtungsfaktor ergeben sich die folgenden Effektivitätswerte: Effektivität E = 75,6 Prozent und Effektive Fallzahl $n' = 2.417$.

Die Effektivität der Gewichte von 75,6 Prozent ist als sehr gut zu bewerten. Die statistischen Analysen sind mit einer ausreichenden Fallzahl (statistische Power) und einer hinreichenden statistischen Genauigkeit möglich.

Tabelle 14 Vergleich von ungewichteten und gewichteten Merkmalen

Merkmale*	Welle 1	Welle 2	Welle 3 (nur Wiederholer mit letzter Teilnahme in Welle 2)		
	Gewichtet (mit Querschnittgewicht Welle 1)	Gewichtet (mit Längsschnittgewicht Welle 2)	Ungewichtet	Gewichtet (mit Längsschnittgewicht Welle 3)	
	%	%	abs.	%	%
Gesamt	100,0	100,0	3.197	100,0	100,0
Geschlecht					
Männlich	46,6	46,4	1.420	44,4	46,3
Weiblich	53,5	53,6	1.777	55,6	53,7
Kohorte					
1959	45,3	45,2	1.500	46,9	45,2
1965	54,7	54,8	1.697	53,1	54,8
Personengruppe					
Sozialversicherungspflichtig	87,3	87,1	2.821	88,2	87,3
Geringfügig beschäftigt	12,7	12,9	376	11,8	12,7
Stellung im Beruf und Arbeitszeit					
Vollzeit	65,9	65,3	2.092	65,4	65,9
Teilzeit	34,1	34,7	1.105	34,6	34,1
Tagesentgelt in Kategorien					
Tagesentgelt unter 50 €	30,7	31,0	946	29,6	31,0
Tagesentgelt 50 bis unter 85 €	25,0	24,8	750	23,4	24,6
Tagesentgelt 85 bis unter 120 €	21,9	21,7	722	22,6	21,7
Tagesentgelt 120 € und mehr	22,4	22,5	779	24,4	22,7
Ausbildung u. Schulbildung kombiniert					
Volks-, Haupt-, Realschule ohne Ausbildung	5,0	5,2	110	3,4	5,1
Volks-, Haupt-, Realschule mit Ausbildung	61,5	61,2	1.934	60,5	61,4
Abitur ohne u. mit Ausbildung	13,3	13,5	473	14,8	13,6

Merkmale*	Welle 1	Welle 2	Welle 3 (nur Wiederholer mit letzter Teilnahme in Welle 2)		
	Gewichtet (mit Querschnitt- gewicht Welle 1)	Gewichtet (mit Längsschnitt- gewicht Welle 2)	Ungewichtet		Gewichtet (mit Längs- schnittgewicht Welle 3)
	%	%	abs.	%	%
Gesamt	100,0	100,0	3.197	100,0	100,0
Fachhochschulabschluss	8,4	8,4	275	8,6	8,2
Hochschulabschluss	11,8	11,7	405	12,7	11,7
Nationalität					
Deutsch	96,4	96,5	3.116	97,5	96,5
Nicht deutsch	3,6	3,5	81	2,5	3,5
Berufe					
Agrarberufe	1,3	1,3	41	1,3	1,4
Einfache manuelle Berufe	9,5	9,9	275	8,6	10,0
Qualifizierte manuelle Berufe	10,0	9,4	316	9,9	9,7
Techniker	5,7	5,6	185	5,8	5,7
Ingenieure	3,4	3,3	121	3,8	3,3
Einfache Dienste	14,3	14,0	358	11,2	13,6
Qualifizierte Dienste	5,1	5,0	179	5,6	5,1
Semiprofessionen	13,1	13,3	472	14,8	13,0
Professionen	2,0	2,0	65	2,0	2,0
Einfache kaufmänn. u. Verwaltungsberufe	9,3	9,6	304	9,5	9,6
Qualifiz. kaufmänn. u. Verwaltungsberufe	21,7	21,9	720	22,5	21,9
Manager	3,2	3,1	103	3,2	3,2
Sonstiges	1,4	1,4	58	1,8	1,4
Arbeitslosigkeitserfahrung					
Nein	76,6	76,5	2.547	79,7	76,9
Ja	23,4	23,5	650	20,3	23,1
Ost-West-Kennzeichen der Betriebsstätte					
Ost	15,9	15,6	508	15,9	15,2
West	84,1	84,4	2.689	84,1	84,8

Merkmale*	Welle 1	Welle 2	Welle 3 (nur Wiederholer mit letzter Teilnahme in Welle 2)		
	Gewichtet (mit Querschnitt- gewicht Welle 1)	Gewichtet (mit Längsschnitt- gewicht Welle 2)	Ungewichtet		Gewichtet (mit Längs- schnittgewicht Welle 3)
	%	%	abs.	%	%
Gesamt	100,0	100,0	3.197	100,0	100,0

Bundesland (Wohnort zum Zeitpunkt 1. Welle)

Schleswig-Holstein	3,4	3,4	62	1,9	3,4
Hamburg	1,9	1,4	17	0,5	1,4
Niedersachsen	9,8	10,2	365	11,4	10,4
Bremen	0,7	0,8	29	0,9	0,8
Nordrhein-Westfalen	21,9	21,6	664	20,8	22,1
Hessen	7,5	7,6	294	9,2	7,8
Rheinland-Pfalz	5,1	5,3	113	3,5	5,0
Baden-Württemberg	13,4	14,0	487	15,2	14,2
Bayern	15,6	15,0	538	16,8	15,0
Saarland	1,3	1,4	19	0,6	1,0
Berlin	3,4	3,7	83	2,6	3,5
Brandenburg	3,3	3,0	77	2,4	3,1
Mecklenburg-Vorpommern	2,0	1,8	40	1,3	1,4
Sachsen	4,9	4,6	190	5,9	4,5
Sachsen-Anhalt	3,0	3,1	103	3,2	3,1
Thüringen	2,8	3,0	116	3,6	3,1

Gemeindegröße BIK 1-10 (des Wohnorts zum Zeitpunkt 1. Welle)

Unter 2.000 EW	2,2	2,3	9	0,3	2,0
2.000 bis u. 5.000 EW	3,0	2,4	59	1,9	2,5
5.000 bis u. 20.000 EW	8,4	8,1	297	9,3	8,1
20.000 bis u. 50.000 EW	12,2	12,3	435	13,6	12,7
50.000 bis u. 100.000 EW STyp 2/3/4 (Randbereich)	8,5	8,5	346	10,8	8,8
50.000 bis u. 100.000 EW STyp 1 (Kernber.)	2,3	2,3	55	1,7	2,3
100.000 bis u. 500.000 EW STyp 2/3/4 (Randbereich)	16,3	16,9	625	19,6	16,8
100.000 bis u. 500.000 EW STyp 1 (Kernber.)	14,1	14,4	426	13,3	14,0
500.000 und mehr EW STyp 2/3/4 (Randbereich)	10,3	10,5	378	11,8	10,7

Merkmale*	Welle 1	Welle 2	Welle 3 (nur Wiederholer mit letzter Teilnahme in Welle 2)		
	Gewichtet (mit Querschnitt- gewicht Welle 1)	Gewichtet (mit Längsschnitt- gewicht Welle 2)	Ungewichtet		Gewichtet (mit Längs- schnittgewicht Welle 3)
	%	%	abs.	%	%
Gesamt	100,0	100,0	3.197	100,0	100,0
500.000 und mehr EW STyp 1 (Kernbereich)	22,7	22,3	567	17,7	22,1

Die Merkmale der Befragten stammen alle aus Registerdaten der Bundesagentur für Arbeit zum Stichtag 31.12.2009. Teilweise wurden fehlende Werte der hier dargestellten Merkmale zum Zwecke der Vergleichbarkeit und Übersichtlichkeit mit imputierten (Ersatz durch Modal- oder Mittelwert) Werten ausgewertet.

Der ausgelieferte Gewichtungssatz enthält neben dem aktuellen Längsschnittgewicht für Welle 3 erneut alle Vorwellingewichte sowie die reziproken Bleibewahrscheinlichkeiten. Damit können jegliche Längsschnittanalysen einfach durchgeführt werden.

Impressum:

lidA-Methodenreport 4/2020

Autoren:

Jacob Steinwede, Jonathan Ruiz Marcos, Martin Kleudgen
infas Institut für angewandte Sozialwissenschaft

Herausgeber + Redaktion:

Bergische Universität Wuppertal, Lehrstuhl für Arbeitswissenschaft

Nachdruck - aus auszugsweise - nur mit Genehmigung des
Herausgebers gestattet

Bezugsmöglichkeit:

https://arbeit.uni-wuppertal.de/fileadmin/arbeit/Publikationen/Methodenbericht_lidA_W3

Internet: www.lida-studie.de