

Kolja Briedis | Almuth Lietz | Kai Mühleck | Uwe Ruß | Percy Scheller | Isabelle Schiffer-Fiedler | Ulrike Schwabe | Moritz Seifert | Carola Teichmann | Susanne de Vogel | Robert Birkelbach | Ute Hoffstätter | Tina Oestreich | Henrike Steins | Anne Weber

# Nacaps 2018

Daten- und Methodenbericht zur Datenpaketversion  
3.0.0 der National Academics Panel Study 2018  
(1.-6. Befragungswelle)



Dieses Werk steht unter der Creative Commons Namensnennung Nicht kommerziell-Share Alike 4.0 International (CC BY-NC-SA 4.0) Deed

<https://creativecommons.org/licenses/by-nc-sa/4.0/deed.de>

#### **Bibliographische Angaben**

Briedis, K., Lietz, A., Mühleck, K., Ruß, U., Scheller, P., Schiffer-Fiedler, I., Schwabe, U., Seifert, M., Teichmann, C., de Vogel, S., Birkelbach, R., Hoffstätter, U., Oestreich, T., Steins, H. & Weber, A. (2025). Nacaps 2018. Daten- und Methodenbericht zur Datenpaketversion 3.0.0 der National Academics Panel Study 2018 (1.-6. Befragungswelle). Hannover: Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (DZHW). <https://doi.org/10.21249/DZHW:nac2018-dmr-de:3.0.0>

Zugehöriges Datenpaket: <https://doi.org/10.21249/DZHW:nac2018:3.0.0>

#### **Impressum**

##### **Herausgegeben von**

Deutsches Zentrum für Hochschul- und  
Wissenschaftsforschung GmbH (DZHW)  
Lange Laube 12 | 30159 Hannover | [www.dzhw.eu](http://www.dzhw.eu)  
Postfach 2920 | 30029 Hannover  
Tel.: +49 511 450670-0 | Fax: +49 511 450670-960

##### **Geschäftsführung**

Dr. Marcus Beiner  
Regina Oelfke

##### **Vorsitzender des Aufsichtsrats**

Ministerialdirigent Peter Greisler

##### **Registergericht**

Amtsgericht Hannover | HRB 210251  
Umsatzsteuer-Identifikationsnummer:  
DE291239300

# Inhaltsverzeichnis

Tabellen-/Abbildungsverzeichnis	II
1 Inhalt und Anlage der Studie	3
2 Erhebungsinstrument	5
2.1 Pretests .....	5
2.2 Inhalte der Erhebungsinstrumente .....	6
2.2.1 Erstbefragung .....	6
2.2.2 Folgebefragungen.....	7
3 Grundgesamtheit und Stichprobenverfahren	11
4 Durchführung der Erhebung	12
4.1 Erstbefragung.....	12
4.2 Folgebefragungen .....	14
5 Rücklauf	16
5.1 Erstbefragung.....	16
5.2 Folgebefragungen .....	17
6 Datenaufbereitung	19
6.1 Datenübertragung.....	19
6.2 Datenprüfung und -bereinigung .....	19
6.3 Generierung von Variablen .....	20
6.4 Datenstruktur und Datenformat.....	21
6.5 Vergabe von Variablennamen und Variablen-/Wertelabels.....	21
6.6 Codierung fehlender Werte .....	22
7 Repräsentativität der Daten	25
8 Anonymisierung	31
9 Literatur	33
10 Anhang	35
10.1 Dokumentation der Herkunft sekundär genutzter Fragen .....	35
10.2 Referenzliste: Destatis-Studienbereiche zu projekteigener Fächergruppeneinteilung .....	36

## Tabellen-/Abbildungsverzeichnis

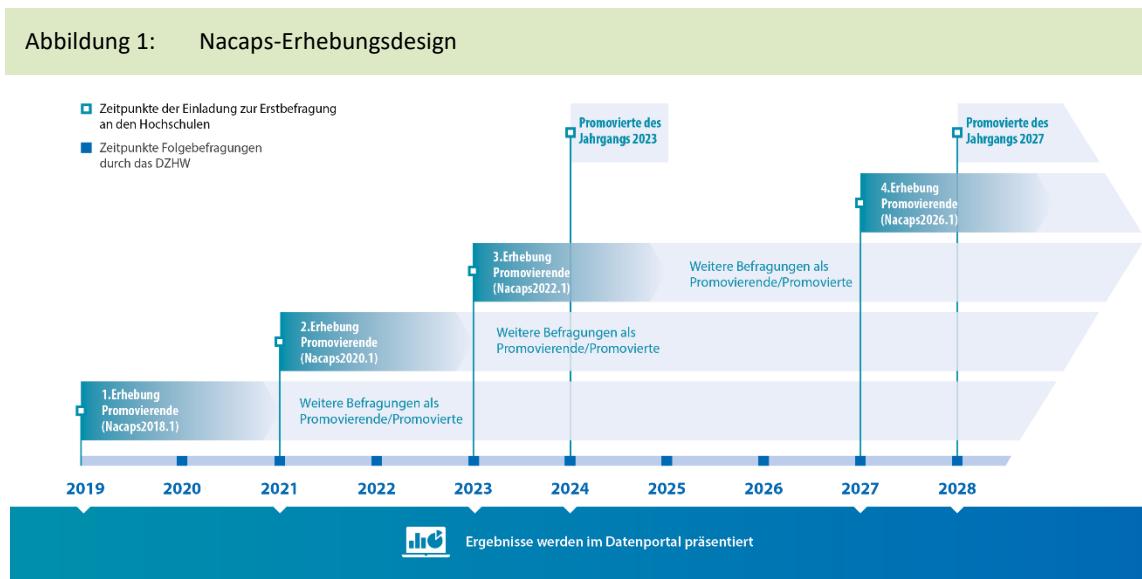
Abbildung 1:	Nacaps-Erhebungsdesign.....	3
Abbildung 2:	Jährliche Befragungen von Nacaps 2018 und der Verlauf der Covid19-Pandemie .....	4
Tabelle 1:	Geschlecht und Fachbereich (DFG) der interviewten Promovierenden .....	5
Tabelle 2:	Versandtage der Erhebungsunterlagen der Erstbefragung .....	12
Tabelle 3:	Versandtage der Erhebungsunterlagen der Folgewellen.....	14
Tabelle 4:	Beteiligung der Hochschulen .....	16
Tabelle 5:	Brutto- und Nettorücklaufquote der ersten Nacaps-Befragungswelle .....	16
Abbildung 3:	Relativer Anteil beendeter Fragebögen (in %) über den Verlauf der Feldzeit .....	17
Tabelle 6:	Brutto- und Nettorücklaufquote der Nacaps-Folgebefragungswellen (nach AAPOR 2016)	18
Tabelle 7:	Themengebiete und Kürzel für Variablenamen in Nacaps 2018.....	22
Tabelle 8:	Systematik des FDZ-DZHW für fehlende Werte.....	24
Tabelle 9:	Verteilung der Promovierenden nach Destatis 2019 und in Nacaps 2018, 1. Welle	
	25	
Tabelle 10:	Verteilung von männlichen und weiblichen Promovierenden nach Destatis 2019 und in Nacaps 2018, 1. Welle .....	26
Tabelle 11:	Logistische Regression der Teilnahmewahrscheinlichkeit an Welle 2, 3 und 4.....	28
Tabelle 12:	Logistische Regression der Teilnahmewahrscheinlichkeit an Welle 5 und 6 .....	29
Tabelle 13:	Zugangswege im FDZ-DZHW.....	31
Tabelle 14:	Zuordnung der Destatis-Studienbereiche zu projekteigener Fächergruppeneinteilung.....	36

# 1 Inhalt und Anlage der Studie

Nacaps 2018 ist Teil der Studienreihe *National Academics Panel Study (Nacaps)*, einer Längsschnittstudie, die erstmalig systematisch anhand standardisierter Online-Befragungen von Promovierenden und Promovierten Daten zu Karriere- und Lebensverläufen von Hochqualifizierten in Deutschland erhebt – für Forschung, Hochschulen und Wissenschaftspolitik (Wegner & Briedis, 2020).<sup>1</sup> Nacaps wird vom Deutschen Zentrum für Hochschul- und Wissenschaftsforschung (DZHW) durchgeführt und wurde bis einschließlich 2024 vom Bundesministerium für Bildung und Forschung (BMBF) gefördert.

Die Nacaps-Studienreihe weist ein Multi-Kohorten-Panel-Design auf, bei dem im Abstand von zwei Jahren Promovierendenkohorten, d. h. Personen, die zu einem bestimmten Stichtag an einer deutschen Hochschule zur Promotion registriert waren, kontaktiert und in der Folge jährlich befragt werden (vgl. Abbildung 1). Bislang gibt es Promovierendenkohorten mit den Stichtagen im Dezember 2018, 2020 und 2022, die jeweils im Folgejahr zum ersten Mal befragt wurden. Zur Auffrischung und partiellen Ergänzung wurden im Jahr 2024 zusätzlich Promovierte des Jahres 2023 einmalig befragt. Geplant ist zudem eine neue Promovierendenkohorte, die im Jahr 2027 zum ersten Mal befragt werden soll. Im Jahr 2028 sollen ergänzend Promovierte des Jahrgangs 2027 in das Panel integriert werden.

Abbildung 1: Nacaps-Erhebungsdesign

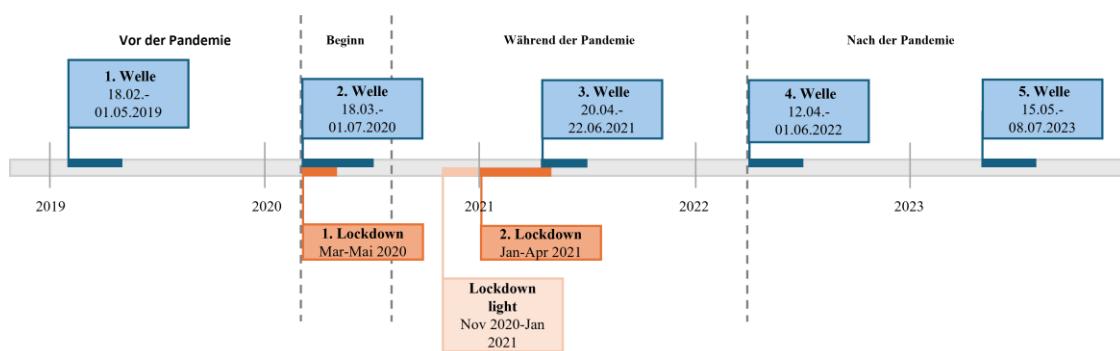


Nacaps 2018 bildet als erste Promovierendenkohorte den Auftakt der Studienreihe. Die Grundgesamtheit beinhaltet alle Personen, die zum Stichtag 1. Dezember 2018 an einer deutschen Hochschule zur Promotion registriert waren. Die erste Befragungswelle fand von Februar bis Mai 2019 statt. Die Folgebefragungswellen 2-6 erfolgten im jährlichen Abstand jeweils im Frühjahr.

<sup>1</sup> Aktuelle Informationen zu Nacaps können über die Website des Projekts (<https://www.nacaps.de>) abgerufen werden.

Damit ist Nacaps 2018 dadurch gekennzeichnet, dass die Datenerhebung teilweise während der Covid19-Pandemie stattgefunden hat. In Abbildung 2 (adaptiert nach Rußmann, Netz & Schwabe, 2025) sind die jährlichen Befragungen von Nacaps 2018 (oberer Teil; nur die Wellen 1-5) und der Verlauf der Covid19-Pandemie (unterer Teil) dargestellt. Die Initialbefragung (Welle 1) hat im Frühjahr 2019 vor der Covid19-Pandemie stattgefunden. Die beiden darauffolgenden Befragungen (Welle 2 und 3) fanden im Frühjahr 2020 und 2021 während der Covid19-Pandemie statt, wobei die zweite Welle fast vollständig mit dem ersten Lockdown zusammenfiel. Ab der vierten Welle fand die Datenerhebung nach der Covid19-Pandemie statt. Die Daten von Nacaps 2018 sind daher für die Untersuchung der Auswirkungen der Covid19-Pandemie mittels eines Vorher-Nachher-Designs geeignet.

Abbildung 2: Jährliche Befragungen von Nacaps 2018 und der Verlauf der Covid19-Pandemie



Quelle: Rußmann, Netz & Schwabe (2025), S. 650

Neben Grunddaten für die Bildungsberichterstattung enthalten die Daten detaillierte Informationen über die Lern- und Entwicklungsbedingungen während der Promotionsphase sowie über die Karriere- und Lebensverläufe von Promovierten und Promotionsabbrecher\*innen im Anschluss an die Promotionsphase. Darüber hinaus beinhalten die Daten eine Reihe von Persönlichkeitsmerkmalen (u. a. Big Five, Selbstwirksamkeit, interne/externe Kontrollüberzeugungen), Angaben zur Gesundheitssituation sowie sozio- und bildungsbiografische Hintergrundinformationen. Ergänzt wird das Befragungsprogramm durch wellenspezifische Schwerpunktthemen (z. B. Wissenschaftskommunikation, Covid19-Pandemie). Dadurch ergibt sich ein großes, bisher nicht vorhandenes Analysepotenzial für die Hochschul- und Wissenschaftsforschung.

## 2 Erhebungsinstrument

In allen Befragungswellen von Nacaps 2018 wurde als Erhebungsinstrument ein standardisierter Online-Fragebogen in deutscher und englischer Sprache eingesetzt. Die Programmierung des Fragebogens und die Durchführung der Befragung erfolgten mit der DZHW-Onlinebefragungssoftware Zofar.

Kapitel 2.1 beschreibt die vor der Erstbefragung zur Prüfung und Optimierung des Erhebungsinstruments durchgeführten kognitiven und quantitativen Pretests. Kapitel 2.2 stellt die zentralen Inhalte der Erhebungsinstrumente vor.

### 2.1 Pretests

**[Ziel und Verfahren]** Das Erhebungsinstrument wurde vor der Erstbefragung zunächst durch einen kognitiven und anschließend durch einen quantitativen Pretest geprüft. Ziel des kognitiven Pretests war es zu validieren, ob das Erhebungsinstrument für die Zielgruppe der Promovierenden geeignet war, ob sich neu entwickelte Items für unterschiedliche Promovierendengruppen gleichermaßen eigneten und ob diese verständlich waren. Ziel des quantitativen Pretests war es, das Erhebungsinstrument im Feld unter ähnlichen Bedingungen wie in der Haupterhebung zu testen. Einerseits sollten die Validität und Reliabilität neu entwickelter Instrumente getestet werden. Andererseits sollte die Dauer der Befragung, der Aufbau des Fragebogens sowie die generelle Durchführbarkeit der Online-Befragung und die Wirksamkeit der Incentivierung geprüft werden.

**[Kognitiver Pretest]** Vom 6. bis 15. Dezember 2017 wurden 17 kognitive Interviews mit Promovierenden an den Standorten Hannover und Berlin durch vier Nacaps-Projektmitarbeiter\*innen *face-to-face* durchgeführt. Es wurden 90 Minuten pro Interview einkalkuliert, wobei die tatsächliche Interviewlänge zwischen 72 Minuten und 110 Minuten lag. Als Aufwandsentschädigung erhielten die Interviewten 50 Euro. Abhängig von den konkret zu testenden Fragen und dem jeweiligen Erkenntnisinteresse kamen Befragungstechniken, wie *Probing* (z. B. Information Retrieval Probing, Comprehension Probing), *Confidence Rating* und *General Probing* zum Einsatz (Lenzner, Neuert & Otto, 2015; Prüfer & Rexroth, 2000; Weichbold, 2019). Dabei wurden Personen aus unterschiedlichen Fächergruppen und sowohl Frauen als auch Männer interviewt (Tabelle 1). Es wurden verschiedene Fragen und Itembatterien zu den Themen *Weg zum Promotionsthema*, *Promotionstypen*, *Soziale Netzwerke und Unterstützung*, *wissenschaftliche Tätigkeit* und *Betreuerverhältnis* getestet. In zwei verschiedenen Versionen wurde jeweils die Reihenfolge der Themen variiert.

Tabelle 1: Geschlecht und Fachbereich (DFG) der interviewten Promovierenden

Fachbereich	Geschlecht		Gesamt
	weiblich	männlich	
Geistes- und Sozialwissenschaften	4	1	5
Lebenswissenschaften	3	0	3
Naturwissenschaften	3	2	5
Ingenieurwissenschaften	0	4	4
Gesamt	10	7	17

**[Quantitativer Pretest]** Für den quantitativen Pretest wurden 10.089 Namen und E-Mail-Adressen von Promovierenden und wissenschaftlichen Mitarbeiter\*innen ohne Doktortitel auf frei zugänglichen Webseiten von 29 promotionsberechtigten deutschen Hochschulen von Januar bis März 2018 recherchiert. Die Programmierung des Fragebogens und die Durchführung der Befragung wurden unter Verwendung der DZHW-Onlinebefragungssoftware Zofar durchgeführt. Die Feldphase begann am 9. Mai und endete am 31. Mai 2018. In diesem Zeitraum wurden zwei Reminder (am 17. und 23. Mai 2018) versandt. Von den 10.089 angeschriebenen Personen haben 1.692 vollständig und 3.109 teilweise an der Befragung teilgenommen. Somit betrug die Rücklaufquote 16,7 bzw. 30,8 Prozent.

## 2.2 Inhalte der Erhebungsinstrumente

Die Inhalte der Erhebungsinstrumente werden zunächst für die Erstbefragung (vgl. Kapitel 2.2.1) und anschließend für die fünf Folgebefragungen (vgl. Kapitel 2.2.2) dargestellt.

Darauf hingewiesen sei, dass einige Fragen der Erhebungsinstrumente aus anderen Instrumenten – teils unmodifiziert, teils modifiziert – sekundär genutzt wurden, vgl. dazu Anhang 10.1.

### 2.2.1 Erstbefragung

Zu Beginn des Fragebogens der ersten Befragungswelle wurden die Befragten nach ihrem derzeitigen Promotionsstatus<sup>2</sup> gefragt. Je nach Promotionsstatus wurden im weiteren Verlauf der Befragung verschiedene Fragen angezeigt.

Der Fragebogen war in sechs thematische Blöcke gegliedert. Es wurden *Angaben zur Promotion (A)*, das *Promotionsthema und die Promotionsbedingungen (B)*, der *Arbeitsalltag in der Promotionsphase (C)*, die *persönliche Lebenssituation und die Karriereziele (D)*, die *Soziodemografie und die vorangegangen Bildungsetappen (E)* und zuletzt Aspekte wie *Panelbereitschaft, Newsletter, Verlosung und Abfrage der Kontaktdaten (F)* erfasst.

**[Angaben zur Promotion]** Der Themenblock *Angaben zur Promotion* umfasste Fragen zur Motivation bei Promotionsbeginn, zum inhaltlichen Beginn und zum Zeitpunkt der Registrierung an der Hochschule, zum Namen und Standort der Hochschule, zu Gründen für die Hochschulwahl, zu Kooperationen mit Hochschulen im Ausland oder Fachhochschulen sowie zum Promotionsfach. Sofern die Promotion bereits abgeschlossen wurde, wurde ferner nach dem Datum des Abschlusses gefragt; sofern die Promotion abgebrochen wurde, wurden das Datum und die Gründe des Abbruchs erhoben.

**[Promotionsthema und Promotionsbedingungen]** Der Themenblock *Promotionsthema und Promotionsbedingungen* enthielt Fragen zu kooperativen Promotionsformen in Forschungsprojekten, zum Promotionsthema, zum subjektiv wahrgenommenen Promotionsfortschritt, zu strukturierten Promotionsprogrammen, zu Stipendien, zu Kursen und Lehrveranstaltungen (inkl. der Zufriedenheit damit), zu Kooperationen mit externen Organisationen, zur Form der Promotion (Monographie vs. kumulative Promotion), zu Abbruchintentionen, zum Vorhandensein einer Promotions- bzw. Betreuungsvereinbarung sowie zur Anzahl an Betreuer\*innen. Ferner umfasste er Fragen zum/zur Betreuer\*in der Promotion (Affiliation, Geschlecht, formaler Betreuungsstatus, Position, Austauschhäufigkeit, Betreuungsstil, Zufriedenheit sowie Sympathieeinschätzung), zur guten wissenschaftlichen Praxis, zu der Finanzierungsquelle sowie spezifische Fragen für Promovierende, die an einer Hochschule oder Forschungseinrichtung arbeiten.

**[Arbeitsalltag in der Promotionsphase]** Der Themenblock *Arbeitsalltag in der Promotionsphase* setzte sich zusammen aus Fragen zur Beschäftigung (wie Befristung, Stellenumfang), zu den

---

<sup>2</sup> Mit den vier Antwortmöglichkeiten: „Ich promoviere“, „Ich habe das Promotionsverfahren abgeschlossen“, „Ich habe mein Promotionsvorhaben unterbrochen“ und „Ich habe mein Promotionsvorhaben abgebrochen“.

Qualifizierungszielen sowie zum thematischen Bezug zum Promotionsprojekt. Weiterhin wurde nach konkreten wissenschaftlichen Tätigkeiten, der täglichen Motivation zur Arbeit an der Promotion, den jüngsten kooperativen Projekten mit anderen Forscher\*innen oder Partner\*innen außerhalb der Wissenschaft, nach der Anzahl unterrichteter Semesterwochenstunden und nach Auslandsaufenthalten im Rahmen des Promotionsvorhabens gefragt.

**[Persönliche Lebenssituation und Karriereziele]** Der Themenblock *Persönliche Lebenssituation und Karriereziele* erfasste die allgemeine Lebenszufriedenheit, enthielt Fragen zur Partner- und Elternschaft, zur Zufriedenheit mit der Vereinbarkeit von Arbeits- und Privatleben, zum allgemeinen Gesundheitszustand, zum körperlichen und seelischen Befinden sowie zum Vorhandensein einer Behinderung. Weiterhin wurden als psychologische Konstrukte die Big Five (Fünf-Faktoren-Modell), die allgemeine Risikobereitschaft, die allgemeinen Kontrollüberzeugungen und die allgemeinen Selbstwirksamkeitserwartungen erfasst. Ferner wurden Instrumente entwickelt und eingesetzt, die die Wichtigkeit von bestimmten Stelleneigenschaften, die Attraktivität von verschiedenen Stellen, die Intention zum Verbleib im Wissenschaftsbereich, die Sektorenpräferenz, das Anstreben einer Professur, den Wunsch nach dauerhafter Beschäftigung in der Wissenschaft und die spezifischen Selbstwirksamkeitserwartungen mit Blick auf eine Tätigkeit innerhalb oder außerhalb der Wissenschaft erfassen.

**[Soziodemografie und vorangegangene Bildungsetappen]** Der Themenblock *Soziodemografie und vorangegangene Bildungsetappen* enthielt zum einen Fragen nach Geschlecht, Geburtsjahr und -ort sowie Staatsangehörigkeit(en) der Befragten, zum anderen nach Schulabschluss, Ausbildungsabschluss und beruflicher Stellung sowie Bildungsaspirationen der Eltern. Außerdem wurden der Ort und die Note der Studienberechtigung und Angaben zum promotionsrelevanten Studium erhoben.

**[Panelbereitschaft, Newsletter, Verlosung, Abfrage der Kontaktdaten]** Die Befragung schloss mit Fragen zur Panelbereitschaft, zur Kontaktierung zum Informieren über Nacaps-Ergebnisse und zur Teilnahme an der Verlosung inkl. Abfrage der Kontaktdaten und bot die Möglichkeit zur Angabe von Kommentaren.

## 2.2.2 Folgebefragungen

In den Folgebefragungen waren bestimmte Kernthemen in jeder Befragungswelle enthalten, auch wenn die genauen Fragen pro Kernthema teilweise variieren. Zum Teil wurden dabei auch Angaben aus vorangegangenen Wellen in die aktuelle Befragung eingespielt („preload“), um daran anknüpfen zu können. Darüber hinaus gab es spezifische Themen, die nur in einzelnen Wellen vorkamen. Die genauen Inhalte der Kernthemen und der wellenspezifischen Themen werden im Folgenden genauer erläutert.

### Kernthemen

**[Allgemeine Angaben zur Promotion]** In allen fünf Befragungswellen wurde einleitend nach dem aktuellen Promotionsstatus (promovierend, unterbrochen, Dissertationsschrift bzw. alle relevanten Arbeiten zur Begutachtung eingereicht (ab dritter Welle), abgeschlossen, abgebrochen) gefragt. Sofern die Promotion im Anschluss an die vorausgehende Befragung abgeschlossen wurde, wurde ferner nach dem Datum des Abschlusses, nach Änderungen des Promotionsthemas, der Dissertationsart (Monographie vs. kumulative Promotion) und -veröffentlichung sowie nach der Promotionsnote gefragt; sofern die Promotion abgebrochen/unterbrochen wurde, wurden in der zweiten Welle die Gründe des Abbruchs / der Unterbrechung erhoben sowie in allen fünf Wellen das Datum des Abbruchs erfasst und erfragt, ob die Hochschule über den Abbruch / die Unterbrechung informiert ist. Für Befragte, die ihre Promotion im Anschluss an die Vorwelle abgeschlossen oder abgebrochen haben, wurden zudem (nochmals) die Promotionshochschule und das (angestrebte) Promotionsfach

erhoben. Ab der dritten Welle wurden Personen, die seit der letzten Befragung promoviert haben, nach dem Datum der Urkundenverleihung und nach Promotionspreisen gefragt. Zudem wurden Informationen zu den Gutachter\*innen (Anzahl, Institution, formaler Status der Gutachter\*innen an Fachhochschulen, Betreuungsstatus) sowie der selbsteingeschätzten Reputation der erstbegutachtenden Person und der Institution erhoben.

**[Promotionsthema/-bedingungen]** Alle fünf Befragungswellen enthielten Fragen zum subjektiv wahrgenommenen Promotionsfortschritt, zu Erfolgserwartung und Abbruchintentionen (inkl. Gründen), zur Anzahl der Betreuer\*innen, zur Hauptbetreuungsperson (Wechsel, Affiliation, Geschlecht, formaler Betreuungsstatus, Position, Austauschhäufigkeit, Betreuungsstil, Zufriedenheit sowie Sympathieeinschätzung) und zum Besuch von Kursen (in Welle 2 inkl. Zufriedenheit). In der zweiten Welle liefen zudem ein aus mehreren Fragen bestehendes Instrument zur Messung der Lernumwelt in der Promotionsphase sowie Fragen zur Unterstützung bezüglich guter wissenschaftlicher Praxis und zur ideellen Förderung durch Förderprogramme. Ergänzend wurden in der dritten und vierten Welle Fragen zu Kooperationen mit Partner\*innen innerhalb und außerhalb der Wissenschaft (in Welle 3 inkl. Kooperationen mit Unternehmen und externen Organisationen) sowie zu Lehrverpflichtungen gestellt; ab der vierten Welle gab es zudem Fragen zur Mitgliedschaft in einem Promotionsprogramm und in der vierten Welle vertiefende Fragen zur Art des Promotionsprogramms und Details der Mitgliedschaft.

**[Wissenschaftliche Aktivitäten während der Promotion]** In der zweiten Welle wurde nach den Publikationsmedien in der Promotionsphase gefragt. Ab der dritten Welle wurden die wissenschaftlichen Aktivitäten während der Promotion detaillierter erfasst mit Fragen zu Anzahl und Format wissenschaftlicher Publikationen und Beiträgen in Publikums- und Online-Medien, Anzahl von Konferenzbesuchen und -beiträgen, Gutachten sowie Anträgen auf Drittmittel und Patente (gestellt/bewilligt).

**[Mobilitätserfahrungen]** Ab der dritten Welle wurden die promotionsbedingten oder zu anderen wissenschaftlichen Zwecken absolvierten Auslandaufenthalte (Aufenthaltsart, Land, Datum des Beginns, Dauer) seit Beginn der Promotion anhand eines Tableaus erhoben. Erfragt wurden außerdem die Mobilitätspläne (Aufenthaltsart, Land). In den Wellen 3 und 4 wurden diese Fragen ausschließlich den Personen gestellt, die in der jeweiligen Welle angegeben haben, dass sie ihre Promotion abgeschlossen oder abgebrochen haben. Somit wurden in diesen beiden Wellen für diese Statusgruppen einmalig die Auslandaufenthalte seit Promotionsbeginn erfasst. Ab Welle 5 wurden dann alle Befragten nach ihren Auslandaufenthalten und Auslandsplänen gefragt, mit Ausnahme der Personen, die bereits in den Vorwellen ihre Promotion abgebrochen haben. Ursprünglich sollten die Personen mit laufenden, unterbrochenen oder in den Vorwellen abgeschlossenen Promotionen einen anderen Fragetext erhalten und die Auslandaufenthalte seit der letzten Befragung angeben. Leider kam es zu einem Übertragungsfehler, sodass alle Personen den gleichen Fragetext erhalten haben und (in einigen Fällen erneut) nach ihren Auslandaufenthalten seit Promotionsbeginn gefragt wurden, unabhängig davon, ob sie in der Vorwelle bereits abgeschlossen haben oder eine laufende oder unterbrochene Promotion haben. Dies ist sowohl in Welle 5 als auch in Welle 6 der Fall gewesen.

**[Finanzierung und Erwerbstätigkeit]** In allen fünf Befragungswellen wurden zunächst Fragen zur Finanzierung des Lebensunterhalts Promovierender und Unterbrecher\*innen erhoben. Dazu gehören die Finanzierungsquellen, die monatlichen Netto-Einkünfte und die selbsteingeschätzte Finanzierungssicherheit. Zusätzliche Fragen erhielten Stipendiat\*innen, Erwerbstätige an Universitäten und Elterngeldempfänger\*innen. Bei Promovierten und Abbrecher\*innen wurde einleitend der Erwerbsstatus erhoben. Bei Nicht-Erwerbstätigen wurden sodann die Finanzierungsquellen und die Netto-Einkünfte erfragt. Bei allen Erwerbstätigen (Promovierende, Unterbrecher\*innen, Promovierte und Abbrecher\*innen) folgten detailliertere Fragen zu u. a. Beschäftigungszeitraum, Befristung, Stellenumfang, der Anzahl der Verträge, monatlichem Brutto-Einkommen, Betriebsgröße,

Sektor/Branche, Adäquanz, berufliche Stellung, wissenschaftliche Personalkategorie, Führungsverantwortung und berufliche Zufriedenheit, ab der dritten Befragung ergänzend zum Arbeitsort. Zudem wurden mehrere Fragen zum Thema Wissenschaftsbezug der Erwerbstätigkeit gestellt. Von Promovierenden wurden darüber hinaus Informationen zum Promotionsbezug der Erwerbstätigkeit (z. B. Promotion als arbeitsvertraglich festgelegtes Qualifikationsziel, Arbeitszeitanteile, thematischer Bezug) gesammelt.

**[Persönliche Lebenssituation]** Die fünf Befragungswellen enthielten jeweils verschiedene Fragen zur Partner- und Elternschaft (in Welle 2 und 3 inkl. einer Nachfrage zu den Schwierigkeiten bei der Familiplanung), zum allgemeinen Gesundheitszustand sowie zum körperlichen und seelischen Befinden. Ab der dritten Welle wurde zudem der Wohnort erhoben.

**[Einstellungen, Selbsteinschätzung und Ziele]** In allen fünf Befragungswellen wurden die allgemeine Lebenszufriedenheit, die Zufriedenheit mit der Vereinbarkeit von Arbeits- und Privatleben, die allgemeinen Selbstwirksamkeitserwartungen und das TEN/FLEX-Instrument (vgl. Brandstädter & Renner, 1990) erfasst. Ferner wurden Instrumente entwickelt und eingesetzt, die die Wichtigkeit von bestimmten Stelleneigenschaften (nur in Welle 2), die Attraktivität von verschiedenen Stellen, die Intention zum Verbleib im Wissenschaftsbereich, die Sektorenpräferenz, das Anstreben einer Professur, die Promotionsmotivation (nur in Welle 3 und 4), den Wunsch nach dauerhafter Beschäftigung in der Wissenschaft, Selbständigkeitspläne und die spezifischen Selbstwirksamkeitserwartungen mit Blick auf eine Tätigkeit innerhalb oder außerhalb der Wissenschaft erfassen.

**[Konflikte während der Promotionsphase]** Ab Welle 6 wurde ein Themenblock zu erlebten Konflikten während der Promotion erhoben. Hierbei handelte es sich um Fragen der eigens erlebten Konflikte, den involvierten Personen, der Diskriminierung aufgrund von Merkmalen (Ethnizität, Religion, etc.) sowie den Erfahrungen mit sexueller Belästigung und Mobbing.

**[Kontaktdaten/Verlosung]** Alle fünf Befragungswellen schlossen mit Fragen zur Abfrage bzw. Aktualisierung der Kontaktdaten und boten die Möglichkeit zur Angabe von Kommentaren.

### **Wellenspezifische Themen**

**[Politische Partizipation]** In der zweiten Befragungswelle liefen sechs Fragen zum Themenschwerpunkt politische Partizipation. Abgefragt wurde das Interesse der Befragten für Politik, die selbsteingeschätzte Eignung, politisches Engagement zu übernehmen, die wahrgenommene Möglichkeit der politischen Einflussnahme sowie der eigenen Beteiligung am politischen Geschehen, die Einschätzung der Befragten, inwieweit Politiker\*innen ihre Anliegen berücksichtigen und schließlich die eigenen politischen Aktivitäten in den letzten zwölf Monaten.

**[Weiterbildung]** In der zweiten Befragungswelle wurde außerdem anhand von zwei Fragen erhoben, ob die Befragten eine praktische Ausbildungsphase oder eine berufsständisch organisierte Weiterbildung durchlaufen (haben).

**[Großeltern]** Zudem wurden im Rahmen der zweiten Befragungswelle mit vier Fragen der Bildungshintergrund der Großeltern, das Verhältnis und die Kontakthäufigkeit zu den Großeltern erfragt.

**[Panel Commitment]** Der dritten Befragungswelle vorangestellt wurde eine Frage zur Erfassung des Panel Commitments der Teilnehmenden.

**[Wissenschaftskommunikation]** Ein weiterer Schwerpunkt wurde in der dritten Befragungswelle mit fünf Fragen auf das Thema Wissenschaftskommunikation gelegt. Neben der Frage, inwiefern die

Befragten Wissenschaftskommunikation betreiben, wurden die Art der Aktivitäten, die Kommunikationsmedien und die Gründe einer Beteiligung bzw. Nicht-Beteiligung an Wissenschaftskommunikation erhoben.

**[Covid19-Pandemie]** In der vierten Befragungswelle wurde der Themenschwerpunkt auf die Arbeitsbedingungen im Kontext der Covid19-Pandemie gelegt. Abgefragt wurden der Arbeitsumfang, der Arbeitsort, der Einfluss der Covid19-Pandemie auf die Arbeitsbedingungen, die Bewertung der Umgangsweise mit der Pandemie durch die Arbeitgeber\*in und die persönliche Betroffenheit der Befragten.

### 3 Grundgesamtheit und Stichprobenverfahren

**[Grundgesamtheit]** Als wissenschaftlicher Nachwuchs werden in Anlehnung an die Definition in der Berichterstattung des Bundes Hochschulabsolvent\*innen verstanden, die nach dem Studium innerhalb oder außerhalb von Hochschulen oder außeruniversitären Forschungseinrichtungen wissenschaftlich tätig sind oder promovieren, sowie Promovierte, die innerhalb und außerhalb von Hochschulen und außeruniversitären Forschungseinrichtungen einer wissenschaftlichen Tätigkeit nachgehen oder nachgehen wollen. Die genaue Definition, wer als Promovierende\*r gilt, ist lange Zeit nicht einheitlich geregelt gewesen. Nacaps orientiert sich daher an der Definition des Hochschulstatistikgesetzes<sup>3</sup> sowie an der 2017 eingeführten Promovierendenstatistik des Statistischen Bundesamts. Demnach bestand die Grundgesamtheit der Promovierenden bei Nacaps 2018 aus denjenigen Personen, die zum Stichtag 1. Dezember 2018 an einer deutschen Hochschule zur Promotion registriert waren. Nach Angaben des Statistischen Bundesamts belief sich die Anzahl der so definierten Promovierenden in Deutschland im Jahr 2018 auf 173.339 Personen (Statistisches Bundesamt, 2019c).

**[Vollerhebung]** Im ersten Schritt wurden alle 155 promotionsberechtigten Hochschulen (Stand Oktober 2017) zu einer Teilnahme aufgefordert, anschließend wurden in einem zweiten Schritt alle (registrierten) Promovierenden der 53 teilnehmenden Hochschulen zur Befragung eingeladen.<sup>4</sup> Da weder eine Auswahl bzw. Ziehung von promotionsberechtigten Hochschulen noch von Promovierenden erfolgte, handelt es sich bei dem gewählten Verfahren um eine Vollerhebung. Nicht-teilnehmende Hochschulen und Promovierende werden dabei als Non-Response behandelt.

<sup>3</sup> „Als Promovierende gelten Personen, die von einer zur Promotion berechtigten Einrichtung eine schriftliche Bestätigung über die Annahme als Doktorandin oder Doktorand in dieser Einrichtung erhalten haben.“ (§5 (1) HStatG)

<sup>4</sup> Es ist möglich, dass Personen zum Zeitpunkt der Kontaktierung noch an einer der Partnerhochschulen registriert, faktisch aber bereits an eine andere Hochschule gewechselt waren und die Angaben in der Befragung sich schon auf diese neue Hochschule beziehen.

# 4 Durchführung der Erhebung

## 4.1 Erstbefragung

**[Kontaktaufnahme und Adresspflege]** Von den insgesamt 53 teilnehmenden Hochschulen erfolgte an 48 teilnehmenden Hochschulen der Einladungs- und Reminderversand durch die Hochschulen selbst (Adressmittlung) und drei Hochschulen übermittelten Kontaktdaten an das DZHW, das in diesen Fällen den Einladungs- und Reminderversand übernahm (Adressübermittlung). Zwei teilnehmende Hochschulen konnten weder Kontaktdaten an das DZHW übermitteln noch direkt eine Einladung zur Befragung selbst verschicken. Deswegen wurden die Promovierenden an diesen beiden Hochschulen gebeten, sich für die Befragung auf einer separaten Website zu registrieren. Anschließend wurden die so registrierten Promovierenden durch das DZHW zur Befragung eingeladen.

An den teilnehmenden Hochschulen war der Einladungs- und Reminderversand größtenteils an zentralen Stellen wie Graduiertenzentren und Evaluationsbüros angesiedelt, die den Promovierenden der Hochschule in vielen Fällen bereits bekannt waren. Um einen möglichst einheitlichen Einladungs- und Reminderversand zu gewährleisten, wurden denjenigen Hochschulen, die ihre Promovierenden selbst zur Befragung einluden, verschiedene Materialien durch das DZHW zur Verfügung gestellt. Dabei handelte es sich um eine Handreichung und ein Video-Tutorial zum Einladungs- und Reminderversand sowie Serienbriefvorlagen für die Vorankündigung, das Einladungs- und die Reminderanschreiben. Zusätzlich wurden die teilnehmenden Hochschulen zu einem Workshop zur Befragungsdurchführung an das DZHW in Hannover eingeladen.

**[Erhebungsunterlagen]** Die Erhebungsunterlagen umfassten ein Vorankündigungsschreiben, ein Einladungsschreiben und drei Reminder. Mit der Einladungs-E-Mail sowie mit allen nachfolgenden Remindern erhielten die Befragten einen personalisierten Link zum Online-Fragebogen. Die Befragung wurde mit der DZHW-Onlinebefragungssoftware *Zofar* programmiert und durchgeführt. Reminder wurden nur an diejenigen Personen verschickt, die noch nicht vollständig an der Befragung teilgenommen hatten. Informationen zum Datenschutz wurden zu Beginn des Online-Fragebogens bei der Abfrage der Zustimmung zu den Datenschutzbestimmungen bereitgestellt. Alle Anschreiben wurden den beteiligten Hochschulen durch das DZHW zur Verfügung gestellt und durch diese nur unwesentlich verändert, indem die jeweils übliche Anrede und die jeweilige Grußformel der versendenden Stelle der Hochschule eingefügt wurden.

**[Feldphase]** Die gesamte Feldphase lief vom 18. Februar bis 01. Mai 2019. Vorankündigung, Einladungen und Reminder zur Befragungsteilnahme wurden an folgenden Tagen verschickt (vgl. Tabelle 2).

Tabelle 2: Versandtage der Erhebungsunterlagen der Erstbefragung

Versandtage	Erhebungsunterlage
11.02.2019	Vorankündigung
18.02.2019	Einladung / Beginn der Feldphase
05.03.2019	1. Reminder
20.03.2019	2. Reminder
04.04.2019	3. Reminder

An einigen Hochschulen wurde die Einladung wenige Tage später als geplant verschickt. An fünf Hochschulen konnte die Einladungs-E-Mail aus organisatorischen Gründen erst am 05. März 2019 verschickt werden; in einem weiteren Fall wurde die Einladungs-E-Mail erst am 22. März 2019 versendet. Auf Grund dieser verspäteten Einladungen wurde die Feldphase erst am 01. Mai 2019 beendet. Dadurch konnte auch für die Promovierenden dieser Hochschulen eine ausreichend lange Feldphase gewährleistet und ebenfalls bis zu drei Reminder verschickt werden. Einige wenige Hochschulen hatten sich im Vorfeld gegen das Versenden von drei Remindern ausgesprochen und stattdessen nur einen oder zwei Reminder versendet.

**[Rücklaufsteigernde Maßnahmen]** Zur Steigerung des Rücklaufs wurden mehrere Maßnahmen ergriffen. Zum einen wurde mit der Befragungseinladung und den Remindern ein nicht-monetäres promised Incentive in Form einer Verlosung angeboten. Unter allen Teilnehmer\*innen wurden dabei folgende Sachpreise verlost:

- 1 x Apple iPad Pro 11“ mit 256 GB im Wert von 1.049 Euro
- 2 x Microsoft Surface Go mit 64 GB im Wert von je 429 Euro
- 10 x Wunschgutscheine im Wert von je 50 Euro
- 25 x Netflix-Gutscheine im Wert von je 25 Euro

In einem Fall wurde durch eine teilnehmende Hochschule eine zusätzliche eigene Verlosung für ihre eigenen Teilnehmenden angeboten.

Zum anderen wurden die teilnehmenden Hochschulen gebeten, bereits vor der Feldphase ihre Promovierenden auf die bevorstehende Nacaps-Befragung aufmerksam zu machen. Zu diesem Zweck wurden den Hochschulen Plakate, Flyer, Beispieltexte für Newseinträge auf Webseiten und Bildmaterialien zur Verfügung gestellt. Zusätzlich wurden die Hochschulen gebeten, eine Woche vor Beginn der Feldphase die Vorankündigungs-E-Mail an ihre Promovierenden zu verschicken. Einerseits sollten damit die Promovierenden auf die anstehende Befragung und das Projekt aufmerksam gemacht werden. Andererseits sollte damit der E-Mail-Versand durch die Hochschulen getestet werden. Neben der direkten Ansprache der Promovierenden wurden schließlich verschiedene begleitende Maßnahmen der Öffentlichkeitsarbeit durchgeführt, die das Interesse der Promovierenden am Projekt steigern sollten. So wurden bereits einige Monate vor Beginn der Feldphase eine Projekt-Homepage (<https://www.nacaps.de/>) und ein Twitter-Account (@Nacaps\_Panel) angelegt, über die insbesondere vor und während der Feldphase zielgruppenrelevante Beiträge geteilt wurden.

Des Weiteren wurden die Einladung und die drei Reminder jeweils an verschiedenen Wochentagen verschickt. So wurde die Einladung am Montag verschickt, der erste Reminder zwei Wochen später folgte dann an einem Dienstag. Der zweite Reminder weitere zwei Wochen später wurde wiederum an einem Mittwoch und der dritte Reminder ebenfalls zwei Wochen später am Donnerstag versendet.

Da die Durchführung des Projekts einerseits und die Erreichbarkeit der Promovierenden andererseits in starkem Maße von der Kooperationsbereitschaft promotionsberechtigter Hochschulen abhängen, wurden auch die promotionsberechtigten Hochschulen mit einer Reihe von Maßnahmen incentiviert. So erhielten teilnehmende Hochschulen Zugang zu einem Datenportal, das hochschulspezifische Auswertungen der eigenen Befragungsdaten ermöglicht. Des Weiteren wurden die teilnehmenden Hochschulen mit einem Willkommenspaket (Nacaps-Kaffebecher, Kühlschranksmagnet, Post-It, Flyer, Poster usw.) begrüßt.

## 4.2 Folgebefragungen

**[Kontaktaufnahme und Adresspflege]** Für die weiteren Befragungen konnten die Einladungsschreiben – im Unterschied zur Erstbefragung – direkt über die Nacaps-E-Mail-Adresse an die Befragten gesendet werden, da die Befragten in der Erstbefragung um Einwilligung zur Kontaktaufnahme (für weitere Befragungen) und die Angabe ihrer Adressen gebeten wurden.

Im Verlauf der Folgewellen wurden Widerruffälle, Wünsche zur Einstellung der Kontaktaufnahme, Adressänderungen und Aktualisierungen durch Abwesenheitsnotizen (im Fall einer neu angegebenen Adresse) oder Nichtzustellbarkeitsnachrichten (ggf. Rückgriff auf alternative Adresse) in die Adressdatenbank eingepflegt. In Welle 3 und 4 fanden zudem auch Kontaktaktionen vor dem Einladungsversand statt, die u. a. der Adresspflege dienten.

Ein Großteil der befragten Personen hat sich nach der Erstbefragung außerdem für den Nacaps-Newsletter angemeldet, der jährlich – meist zum Jahresende – an alle Abonnent\*innen mit Informationen zum Projekt und relevanten Rezensionen, Veröffentlichungen oder Tagungsbeiträgen der Projektmitarbeiter\*innen versandt wurde.

**[Erhebungsunterlagen und Feldphase]** Die Folgebefragungen erfolgten wie die Erstbefragung online mit der DZHW-internen Zofar-Befragungssoftware. Bei allen Folgewellen erfolgte der Versand der Einladungs- und Erinnerungsschreiben (Reminder) per E-Mail mit personalisiertem Link zum Online-Fragebogen.

Der Ablauf der Feldphasen der Folgebefragungen ist datumsgenau in der Tabelle 3 dargestellt. Die zweite Befragungswelle startete Mitte März, ab der dritten Befragungswelle wurde der Feldstart aus organisatorischen Gründen auf Mitte April verlegt. In der Welle 5 erfolgte der Feldstart im Mai.

Tabelle 3: Versandtage der Erhebungsunterlagen der Folgewellen

Daten	2. Welle	3. Welle	4. Welle	5. Welle	6. Welle
Vorankündigung/ Aktionen	-	Weihnachts- karte, posta- lisch am 11.12.2020	Vorankündi- gung Post- karte, posta- lisch am 04.04.2022	-	-
Start	18.03.2020	20.04.2021	12.04.2022	15.05.2023	17.04.2024
1. Reminder	25.03.2020	27.04.2021	20.04.2022	23.05.2023	22.04.2024
2. Reminder	09.04.2020	20.05.2021	05.05.2022	12.06.2023	07.05.2024
3. Reminder	03.06.2020 („Restart“)	16.06.2021	25.05.2022	28.06.2023	22.05.2024
4. Reminder	16.06.2020	-	-	-	-
Ende	01.07.2020	22.06.2021	01.06.2022	08.07.2023	07.06.2024

Durch den Covid19-bedingten Lockdown wurde die Feldphase der zweiten Befragungswelle über vier Monate ausgedehnt und ein „Restart-Anschreiben“ sowie eine zusätzliche Erinnerung verschickt, um dem niedrigen Rücklauf in der Anfangsphase der Covid19-Pandemie entgegenzuwirken.

Vor der dritten Befragung wurde an alle Panelteilnehmer\*innen mit gültiger Postadresse eine Weihnachtskarte per Post versandt. Vor Feldstart der vierten Befragungswelle wurde ein postalisches Briefanschreiben als Vorankündigung an die Panelteilnehmer\*innen mit gültiger Postadresse versandt.

**[Rücklaufsteigernde Maßnahmen]** Eine konstante Maßnahme zur Steigerung des Rücklaufs in den Folgebefragungen ist die Incentivierung durch eine angekündigte Verlosung (technische Sachpreise und/oder Gutscheine). Die Preise wurden in jedem Anschreiben genannt. Zumeist wurde ein großer Sachpreis (wie z. B. ein fair hergestelltes aktuelles Smartphone), zwei kleinere Sachpreise (wie zwei Audiogeräte namhafter Hersteller\*innen) und zehn Gutscheine mit einem Wert von jeweils 50 Euro verlost.

Zudem sollten die zusätzlichen Kontaktaufnahmen (Newsletterversand jährlich, zweite Welle: Feldphasenverlängerung und zusätzliches Erinnerungsschreiben; dritte Welle: Weihnachtspostkarte; vierte Welle: Vorankündigungsschreiben) mehr Personen zur Teilnahme motivieren.

# 5 Rücklauf

## 5.1 Erstbefragung

Die erste Befragungswelle von Nacaps 2018 erfolgte an 53 deutschen Hochschulen, wobei alle 155 promotionsberechtigten Hochschulen in Deutschland angefragt worden waren, an der Befragung teilzunehmen. Auf Hochschulebene fand eine Vollerhebung statt, d. h. alle zum Stichtag 1. Dezember 2018 registrierten Promovierenden wurden von den kooperierenden Hochschulen eingeladen.

Tabelle 4: Beteiligung der Hochschulen

Beteiligte	Anzahl
Hochschulen mit Promotionsrecht	155
Teilnehmende Hochschulen	53

Insgesamt wurden 77.509 Einladungen zur Befragung erfolgreich verschickt. Mit 28.368 Fragebögen, bei denen die Befragten den Datenschutzbestimmungen zugestimmt haben, ergibt sich somit eine (bereinigte) Rücklaufquote von 36,6 Prozent (Tabelle 5).

Tabelle 5: Brutto- und Nettorücklaufquote der ersten Nacaps-Befragungswelle

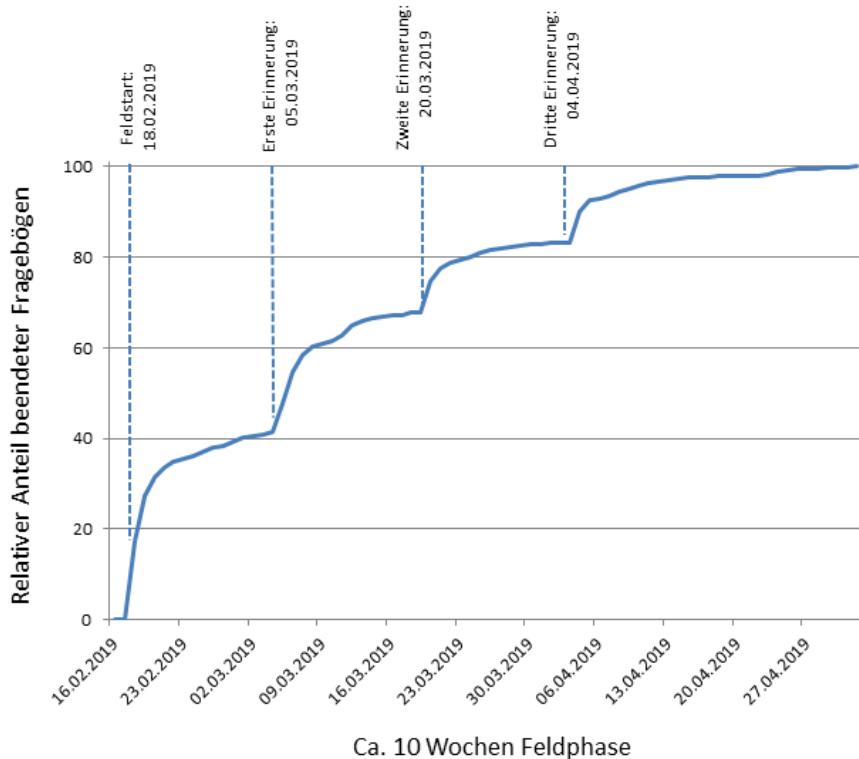
Welle 1	
Versendete Einladungen [1]*	80.086
Unzustellbar*	2.577
Bereinigte versendete Einladungen [2]*	77.509
Fragebogenlink geöffnet [3]	29.842
Datenschutz zugestimmt [4]	28.368
Fragebogen inhaltlich abgeschlossen [5]	23.612
Rücklaufquote 1 ([3]/[1])	37,3 %
Rücklaufquote 2 ([3]/[2])	38,5 %
Rücklaufquote 3 ([4]/[1])	35,4 %
Rücklaufquote 4 ([4]/[2])	36,6 %
Rücklaufquote 5 ([5]/[1])	29,5 %
Rücklaufquote 6 ([5]/[2])	30,5 %

\* Zu beachten ist, dass die hier genannte Anzahl auch Fälle enthält, die für andere Zwecke zusätzlich zur Befragung eingeladen wurden.

Darauf hingewiesen sei, dass im Panelverlauf einige Befragte um komplette Löschung ihrer Daten gebeten haben, weshalb die Fallzahl in den veröffentlichten Daten etwas niedriger ist als in den Rücklaufberechnungen in Tabelle 5.

In Abbildung 3 ist die Entwicklung des Rücklaufs über den gesamten Verlauf der Feldzeit als relativer Anteil beendeter Fragebögen dargestellt. Zusätzlich sind die Zeitpunkte des Versands der drei Erinnerungsschreiben gekennzeichnet, welche insgesamt eine rücklaufsteigernde Wirkung zeigen.

Abbildung 3: Relativer Anteil beendeter Fragebögen (in %) über den Verlauf der Feldzeit



## 5.2 Folgebefragungen

Der Rücklauf der Folgebefragungen wurde in der Tabelle zur Berechnung der Rücklaufquote der American Association for Public Opinion Research (AAPOR)<sup>5</sup> erfasst. In Tabelle 6 sind die wesentlichen Rücklaufzahlen und die Rücklaufquote nach AAPOR RR1 (minimale Quote) und RR2 (inkl. teilweise ausgefüllter Fragebögen) dargestellt.

In Welle 2 wurden Einladungen an alle Teilnehmer\*innen aus Welle 1 versandt, die in Welle 1 ihre Einwilligung für weitere Befragungen sowie eine gültige E-Mail-Adresse angegeben und nicht in der Zwischenzeit ihre Teilnahme widerrufen hatten. Für das weitere Panel gibt es keine generelle Ausfallregel, d. h. Befragte werden auch zur nächsten Folgebefragung eingeladen, wenn sie zwischenzeitlich an keiner der Wellen teilgenommen haben. Jedoch kommt es dennoch zu Ausfällen zwischen den Befragungswellen, da vereinzelt Personen ihre Teilnahme an der Studie widerrufen haben und somit nicht mehr kontaktiert wurden.

Es sei darauf hingewiesen, dass Fälle, die zum Zeitpunkt der Rücklaufberechnungen für Tabelle 6 bereits um komplette Löschung ihrer Daten gebeten hatten (vgl. Kapitel 5.1), nicht mehr in die Berechnungen eingegangen sind. Gleichzeitig muss berücksichtigt werden, dass im Anschluss an die

<sup>5</sup> <https://aapor.org/response-rates/> (zugegriffen am: 20.11.2025)

Rücklaufberechnungen auch noch weitere Datenlöschungsfälle hinzugekommen sind, sodass Differenzen in den Fallzahlen in Tabelle 6 und den veröffentlichten Daten bestehen können.

Tabelle 6: Brutto- und Nettorücklaufquote der Nacaps-Folgebefragungswellen (nach AAPOR 2016)

Daten	2. Welle	3. Welle	4. Welle	5. Welle	6. Welle
Versendete Einladungen [1]	17.247	17.068	16.893	16.806	16.083
Unzustellbar	157	271	486	488	210
Widerrufe (während der Feldphase)	20	24	46	54	49
Befragung geöffnet, aber nicht ausgefüllt [2]	473	494	237	311	246
Teilweise ausgefüllt (abgebrochen ab A01 und vor C01) [3]	626	527	387	284	370
Vollständig ausgefüllt (abgebrochen ab C01 bzw. nicht abgebrochen) [4]	8.893	6.994	6.915	5.420	4.885
Anzahl Fälle ([2]+[3]+[4])	9.992	8.015	7.539	6.015	5.501
davon verwertbare Fälle ([3]+[4])	9.519	7.521	7.302	5.704	5.255
Response Rate 1 (RR1) is the minimum response rate. ([4]/[1])	51,6 %	41,0 %	40,9 %	32,3 %	30,4 %
Response Rate 2 (RR2) counts partial interviews as respondents. ([3]+[4]/[1])	55,2 %	44,1 %	43,2 %	33,9 %	32,7 %

# 6 Datenaufbereitung

## 6.1 Datenübertragung

Die Befragung wurde mit dem DZHW-eigenen Onlinebefragungssystem *Zofar* durchgeführt. Die Angaben aus der Online-Befragung wurden von *Zofar* als .csv-Datei an das Nacaps-Team geliefert und daraufhin als Stata-Datei (.dta) weiterverarbeitet. Der Rohdatensatz wurde auf einem geschützten Server gespeichert.

## 6.2 Datenprüfung und -bereinigung

Im Anschluss an die Datenübertragung erfolgte eine umfassende Prüfung der Daten mit Hilfe des Statistikprogramms Stata. Folgende Prüfungen wurden vorgenommen:

**[Prüfung von Wertebereichen numerischer offener Angaben während der Befragung]** In der Erstbefragung wurden bei numerischen offenen Angaben – um von vorneherein das Ausmaß an unplausiblen Werten einzuschränken – innerhalb der Fragebogenprogrammierung Wertebereiche definiert, die nicht unter- bzw. überschritten werden sollten. Im Falle einer Unter- bzw. Überschreitung der Wertebereiche wurden die Befragten in roter Schrift darauf hingewiesen, dass ihre Eingabe unplausibel erscheint und sie wurden darum gebeten, den Wert zu korrigieren. Erfolgte keine Korrektur nach dem Hinweis, konnte die Befragung dennoch fortgesetzt werden. In den Folgewellen wurde auf die Definition gültiger Wertebereiche verzichtet und ausschließlich Typen (numerisch oder string) festgelegt, in der dritten Welle erfolgte gar keine Plausibilisierung in der laufenden Befragung.

**[Nachträgliche Bereinigung numerischer offener Angaben]** Im Zuge der Datenaufbereitung erfolgte i. d. R. keine nachträgliche Bereinigung von Wertebereichen bei numerischen offenen Angaben. Unplausible oder unmögliche Angaben wurden beibehalten, damit von Nutzer\*innenseite entschieden werden kann, wie damit umgegangen werden soll.<sup>6</sup> Eine Ausnahme besteht bei Variablen, bei denen laut Ausfüllanweisung nur ganze Zahlen bzw. eine Dezimalstelle angegeben werden sollten. In diesen Fällen wurden die abweichenden Werte nachträglich gerundet und die Originalangaben überschrieben. Zudem erfolgte bei offenen Angaben, die eine Eingabe numerischer Werte erforderten, aber Textangaben enthielten<sup>7</sup>, eine Recodierung zu einem numerischen Wert, wenn eine eindeutige Zuordnung möglich war. Ansonsten wurde der Missingwert „-966 nicht bestimmbar“ (vgl. Kapitel 6.6) zugewiesen.

**[Filterprüfung]** Weiterhin wurden einerseits Vollständigkeitsprüfungen hinsichtlich der Filterführung vorgenommen, andererseits auf Filterverstöße getestet (Prüfung der Einhaltung der Filterführung). Da es in der Befragung möglich war, durch Vor- und Zurückspringen Angaben nachträglich zu ändern, ist in den Daten nicht zu erkennen, welche Angabe tatsächlich korrekt ist. Daher wurden keine Änderungen in den Angaben vorgenommen und die Entscheidung, welche Angabe für die Analysen verwendet werden soll, kann von den Datennutzer\*innen selbst getroffen werden.<sup>8</sup> Eine Ausnahme

<sup>6</sup> Darauf hingewiesen sei jedoch, dass die offen erfassten Informationen zu den Kindern (Anzahl und Geburtsjahre) für ihre Aggregation im Zuge der Anonymisierung (vgl. Kapitel 8) zum Teil zuvor plausibilisiert wurden (ohne Änderung der Originalvariablen).

<sup>7</sup> Da in der Befragungsprogrammierung der dritten Welle vorab keine Antworttypen (numerisch vs. string) definiert wurden, ist dies in der dritten Welle häufiger vorgekommen als in den übrigen Wellen.

<sup>8</sup> Eine Möglichkeit Filterverstöße für die Datennutzer\*innen zu markieren, besteht in der Erstellung von Flag-Variablen. Zwar kamen bei einem Großteil der Variablen Filterverstöße vor, die gleichzeitig aber stets nur vereinzelte Fälle betreffen. Da diese zusätzlichen

besteht bei Fällen, bei denen zunächst Angaben bei den Gesundheitsfragen gemacht, der entsprechende Informed Consent durch Zurückspringen jedoch nachträglich zurückgezogen wurde. In diesen Fällen wurden die Originalangaben nachträglich zu „-989 filterbedingt fehlend“ (vgl. Kapitel 6.6) recodiert.

**[Lösung von Fällen bzw. Angaben]** Im Panelverlauf wurden einige Fälle bzw. Angaben aus dem Datensatz entfernt. Wenn Befragte nur Angaben aus einzelnen Wellen gelöscht haben wollten, wurden die Angaben für diese Wellen auf „-995 keine Teilnahme am Panel“ (vgl. Kapitel 6.6) recodiert, aber die Fälle grundsätzlich im Datensatz belassen. Bei einem Wunsch nach vollständiger Lösung der Daten wurde dieser Fall komplett entfernt (vgl. Kapitel 5).

## 6.3 Generierung von Variablen

Im Rahmen der Datenaufbereitung wurden auch zusätzliche Variablen generiert (gekennzeichnet über das Suffix „\_g#“ im Variablenamen, vgl. Kapitel 6.5). Dabei handelt es sich erstens um Variablen mit numerischen Codierungen von ursprünglich offenen Nennungen<sup>9</sup>, zweitens um bereinigte Variablen (vgl. Kapitel 6.2) sowie drittens um Variablen, die aus Datenschutz-/Anonymisierungsgründen erstellt wurden (vgl. Kapitel 8). Das Vorgehen bei der Generierung kann für jede in den Daten enthaltene generierte Variable dem Datensatzreport entnommen werden. Dabei wird teilweise Bezug auf bestimmte Listen genommen, die hier zu finden sind:

- Bei Variablen zur Hochschule:
  - Wellen 1-4: Destatis-Schlüsselverzeichnis für die Studenten- und Prüfungsstatistik, Promovierendenstatistik WS 2018/2019 und SS 2019 (Statistisches Bundesamt, 2019a): [https://www.statistik-nord.de/fileadmin/Dokumente/Erhebungen/Hochschulstatistiken/1\\_Studenten\\_2\\_Pr%C3%BCfungen/Schluesselverzeichnis\\_WS\\_2018.xlsx](https://www.statistik-nord.de/fileadmin/Dokumente/Erhebungen/Hochschulstatistiken/1_Studenten_2_Pr%C3%BCfungen/Schluesselverzeichnis_WS_2018.xlsx), Schlüssel 2.2 (Hochschulen nach Hochschularten und Bundesländern) und Schlüssel 2.3 (Hochschulen, numerisch) (abgerufen am 16.04.2025)
  - Welle 5: Destatis-Schlüsselverzeichnis für die Studenten- und Prüfungsstatistik, Promovierendenstatistik und Gasthörerstatistik WS 2021/2022 und SS 2022 (Statistisches Bundesamt, 2022a): [https://www.statistik-nord.de/fileadmin/Dokumente/Erhebungen/Hochschulstatistiken/1\\_Studenten\\_2\\_Pr%C3%BCfungen/Schluesselverzeichnis\\_WS\\_2021.xlsx](https://www.statistik-nord.de/fileadmin/Dokumente/Erhebungen/Hochschulstatistiken/1_Studenten_2_Pr%C3%BCfungen/Schluesselverzeichnis_WS_2021.xlsx), Schlüssel 2.2 (Hochschulen nach Hochschularten und Bundesländern) und Schlüssel 2.3 (Hochschulen, numerisch) (abgerufen am 03.07.2025)
  - Welle 6: Destatis-Schlüsselverzeichnis für die Studenten- und Prüfungsstatistik, Promovierendenstatistik und Gasthörerstatistik WS 2023/2024 und SS 2024 (Statistisches Bundesamt, 2024): [https://www.statistik-nord.de/fileadmin/Dokumente/Erhebungen/Hochschulstatistiken/1\\_Studenten\\_2\\_Pr%C3%BCfungen/Schluesselverzeichnis\\_WS\\_2023.xlsx](https://www.statistik-nord.de/fileadmin/Dokumente/Erhebungen/Hochschulstatistiken/1_Studenten_2_Pr%C3%BCfungen/Schluesselverzeichnis_WS_2023.xlsx), Schlüssel 2.2 (Hochschulen nach Hochschularten und Bundesländern) und Schlüssel 2.3 (Hochschulen, numerisch) (abgerufen am 03.07.2025)
- Bei Variablen zum Studien- und Promotionsfach:
  - Wellen 1-4: Destatis-Fächersystematik WS 2017/2018 (Statistisches Bundesamt, 2018): [https://www.statistischebibliothek.de/mir/receive/DEHeft\\_mods\\_00092410](https://www.statistischebibliothek.de/mir/receive/DEHeft_mods_00092410), S. 434-436 (abgerufen am 29.04.2025)
  - Welle 5: Destatis-Fächersystematik WS 2021/2022 (Statistisches Bundesamt, 2022b): [https://www.statistischebibliothek.de/mir/receive/DEHeft\\_mods\\_00143900](https://www.statistischebibliothek.de/mir/receive/DEHeft_mods_00143900), S. 424-426 (abgerufen am 03.07.2025)

Flag-Variablen die Datensätze stark aufblähen würden, wurde nach Abwägung von Aufwand und Nutzen auf die Kennzeichnung von Filterverstößen mit Flag-Variablen verzichtet.

<sup>9</sup> Darauf hingewiesen sei in diesem Zusammenhang, dass halboffene Angaben grundsätzlich nicht codiert wurden und auch bei einigen weiteren offenen Angaben auf eine Codierung verzichtet wurde. Teilweise dienten sie vor allem dazu, mit Blick auf zukünftige Befragungen zu prüfen, welche Antwortkategorien ggf. bei einer geschlossenen Frage sinnvoll wären; teilweise waren sie ausschließlich Hilfsvariablen, die bei der Codierung anderer offener Angaben unterstützende Informationen liefern sollten. Für die Codierung der offenen Angaben zum Beruf der Eltern in Welle 1 (Frage E14) gab es bislang keine ausreichenden Kapazitäten.

- Welle 6: Destatis-Fächersystematik WS 2022/2023 (Statistisches Bundesamt, 2023a): [https://www.statistischebibliothek.de/mir/receive/DEHeft\\_mods\\_00149014](https://www.statistischebibliothek.de/mir/receive/DEHeft_mods_00149014), S. 79-81 (abgerufen am 03.07.2025)
- Referenzliste von Destatis-Studienbereichen zu einer projekteigenen Fächergruppeneinteilung, siehe Anhang 10.2
- Bei Variablen zu Land und Staatsangehörigkeit:
  - Wellen 1-4: Destatis-Staats- und Gebietssystematik (Stand 01.04.2019) (Statistisches Bundesamt, 2019b): <https://statistik.arbeitsagentur.de/DE/Statischer-Content/Grundlagen/Klassifikationen/Staats-und-Gebietssystematik/Generische-Publikationen/2019-Staats-und-Gebietssystematik0419-pdf.pdf> (abgerufen am 16.04.2025)
  - Wellen 5-6: Destatis-Staats- und Gebietssystematik (Stand 01.01.2023) (Statistisches Bundesamt, 2023b): <https://statistik.arbeitsagentur.de/DE/Statischer-Content/Grundlagen/Klassifikationen/Staats-und-Gebietssystematik/Generische-Publikationen/2023-Staats-und-Gebietssystematik-pdf.pdf> (abgerufen am 03.07.2025)
  - NEPS-recoding map of countries (Koberg, 2016b): [https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/6-0-0/SC5\\_6-0-0\\_Anonymisation.pdf](https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/6-0-0/SC5_6-0-0_Anonymisation.pdf), S. 33-34 (abgerufen am 16.04.2025)
- Bei Variablen zur Währung des Einkommens:
  - ISO 4217 Currency Codes: <https://www.iso.org/iso-4217-currency-codes.html> (abgerufen am 16.04.2025)
  - Wechselkurs-/Währungsrechner der Europäischen Kommission: [https://commission.europa.eu/funding-tenders/procedures-guidelines-tenders/information-contractors-and-beneficiaries/exchange-rate-inforeuro\\_de](https://commission.europa.eu/funding-tenders/procedures-guidelines-tenders/information-contractors-and-beneficiaries/exchange-rate-inforeuro_de) (abgerufen am 16.04.2025)

## 6.4 Datenstruktur und Datenformat

Die Daten der ersten bis sechsten Befragungswelle wurden zwar separat erhoben, jedoch in einem Datensatz zusammengeführt. Die Zuordnung der Fälle erfolgte über die im Rahmen der ersten Feldphase vergebenen Identifikationsnummern der Befragten. Für jede befragte Person existiert eine Datenzeile (wide-Format). Die Reihenfolge der Variablen orientiert sich an der Reihenfolge der zugehörigen Fragen im Fragebogen. Generierte Variablen (vgl. Kapitel 6.3) wurden soweit möglich hinter der/den jeweiligen Ausgangsvariablen positioniert. Der Datensatz wird sowohl im Stata- als auch im SPSS-Format bereitgestellt.<sup>10</sup>

## 6.5 Vergabe von Variablennamen und Variablen-/Wertelabels

**[Variablen- und Wertelabelvergabe]** Für Variablen- und Wertelabels wurden Kurzformen der Fragebogenformulierungen gewählt. Die Variablen- und Wertelabel liegen auf Deutsch und auf Englisch vor. Im Stata-Format wurden zweisprachige Labels im gleichen Datensatz hinterlegt.<sup>11</sup> Im SPSS-Format existiert für jede Sprache ein eigener Datensatz.

**[Variablenbenennung]** Für die Variablenbenennung wurde ein Präfix-Stamm-Suffix-Schema verwendet. Das *Präfix* enthält die Wellenkennung anhand der Buchstaben a (Welle 1), b (Welle 2), c (Welle 3), d (Welle 4), e (Welle 5) und f (Welle 6). Aus dem *Stamm* geht erstens der Themenbereich, dem die Variable zugeordnet ist, hervor (vgl. Tabelle 7 für die verschiedenen Themenbereiche sowie das zugehörige Kürzel für den Stamm des Variablenamens). Zweitens sind die Variablen innerhalb der definierten Themenbereiche nummeriert. Mit Hilfe eines Buchstabens am Ende des Stamms sind drittens

<sup>10</sup> Bei Nutzung von Stata/BE (maximale Variablenanzahl: 2.048) können Wellen-Subsets der Daten mit Hilfe der Wellenpräfixe in den Variablennamen (vgl. Kapitel 6.5) geladen werden, z. B. die ID-Variable plus alle Variablen der ersten Welle über: use pid a\* using "datasetname.dta"

<sup>11</sup> Mit dem Befehl „label language en“ erfolgt ein Wechsel zu den englischen Labels. Über „label language de“ kann zu den deutschen Labels zurückgewechselt werden.

Variablen differenziert, die zur gleichen Frage gehören und dadurch die gleiche Themendifferenzierung und Nummerierung aufweisen (z. B. bei Itembatterien, Mehrfachnennungen oder Fragen, in denen geschlossene und offene Fragen kombiniert werden). Mit dem *Suffix „\_g#“* werden darüber hinaus generierte Variablen (vgl. Kapitel 6.3) gekennzeichnet.

Tabelle 7: Themengebiete und Kürzel für Variablennamen in Nacaps 2018

Themen-gebiete-kürzel	Themengebiet (englisch)	Themengebiet (deutsch)
ict*	informed consent	Informierte Einwilligung
dbi	basic information on doctoral phase	Eckdaten zur Promotion
did	interruption/drop-out	Promotionsunterbrechung und -abbruch
dtc	thesis topic and motives for PhD/scientific activities	Promotionsthema und -motive/Wissenschaftliche Aktivitäten
dcd	conditions during doctoral phase	Promotionsbedingungen
dsv	supervisor	Betreuer*in
fin	finances	Finanzierung
emp	employment	Beschäftigungssituation
dwr	daily scientific working routine	Wissenschaftlicher Arbeitsalltag
abr	international mobility	Auslandsaufenthalte/Mobilität
psy	personality/psychic traits	Persönlichkeit/psychische Eigenschaften
lcd	living conditions	Persönliche Lebenssituation
hea	health	Gesundheit
crg	career goals	Karriereziele
dem	demography	Demographie
par	social background	Soziale Herkunft
edb	educational biography	Bildungsbiographie
fed	further education	Berufsständische Weiterbildung
pop	political participation	Politische Partizipation
kom	science communication	Wissenschaftskommunikation
sum	panel commitment	Panel Commitment
inf	information	Information
cmt*	comment	Kommentar

\*In den veröffentlichten Daten sind Variablen mit diesem Stamm nicht mehr enthalten.

## 6.6 Codierung fehlender Werte

Zur Codierung fehlender Werte wurde im FDZ-DZHW eine übergreifende Systematik erstellt, um über verschiedene Datensätze des DZHW hinweg eine einheitliche Missingcodierung gewährleisten zu können. Fehlende Angaben werden dabei durch dreistellige negative Werte codiert. Die Missingcodierung der Nacaps-Daten basiert auf dieser Systematik.

Die in Nacaps 2018 verwendeten Missingcodierungen sind in Tabelle 8 fett hervorgehoben. Sie lassen sich drei verschiedenen Gruppen zuordnen.<sup>12</sup> In den ersten beiden Gruppen wird zwischen fehlenden Werten aufgrund von Nicht-Beantwortung von Fragen seitens der Befragten (Item-Nonresponse, Interviewabbruch oder Panel-Nonresponse) und fehlenden Werten aufgrund der Filterführung bzw. Irrelevanz der Frage für die/den Befragte\*n unterschieden (nicht zutreffend). Die dritte Gruppe beinhaltet Missingcodierungen, die durch das Primärforschungsprojekt oder das FDZ im Zuge der Datenaufbereitung vergeben wurden (editierter fehlender Wert). Zu dieser Gruppe gehört auch die Codierung, die aufgrund von Anonymisierungsmaßnahmen (vgl. Kapitel 8) für bestimmte Variablen vergeben wurde.

Schließlich sei darauf hingewiesen, dass in den aus dem Befragungssystem exportierten Rohdaten zunächst nur zwischen zwei Missingkategorien unterschieden wird: (1) Frage wurde gesehen, aber nicht beantwortet und (2) Frage wurde nicht gesehen. Fehlende Werte der ersten Kategorie werden im Zuge der Datenaufbereitung zu dem FDZ-Missingwert -998 „keine Angabe“ bzw. zu -988 „trifft nicht zu“ codiert. Für die zweite Missingkategorie (Frage nicht gesehen) muss unterschieden werden, ob die Frage nicht gesehen wurde, weil die Person nicht in die Frage hineingefiltert wurde oder weil die Person die Befragung abgebrochen hat. Zunächst werden in der Datenaufbereitung alle Missings dieser Kategorie zu -989 „filterbedingt fehlend“ codiert. Im Anschluss werden die Interviewabbrüche codiert. Dafür wird auf Prozessdaten zurückgegriffen, die Informationen zur maximal erreichten sowie zur zuletzt aufgerufenen Seite im Fragebogen beinhalten. Durch diese nachträgliche, händische Codierung kann es vereinzelt zu Unstimmigkeiten<sup>13</sup> kommen. Die Systematik der fehlenden Werte darf daher nicht als verlässliche, technische Gegebenheit verstanden werden.

Explizit hingewiesen sei außerdem auf die folgenden Sonderfälle bzw. Abweichungen von der Systematik:

- Bei den Variablen ccd37a bis ccd37g gibt es die Kategorie „keine Angabe (Antwortkategorie)“, die aktiv von den Befragten angeklickt werden konnte und der analog zur Fragebogenprogrammierung der Code 4 (nicht der Code -997) zugewiesen wurde.
- Bei den Variablen bdwr12a bis bdwr12d, cdwr19a bis cdwr19d, ddwr19a bis ddwr19d, edwr19a bis edwr19d und fdwr19a bis fdwr19d gibt es die Kategorie „trifft nicht zu (Antwortkategorie)“, die aktiv von den Befragten angeklickt werden konnte und der analog zur Fragebogenprogrammierung der Code 6 (nicht der Code -988) zugewiesen wurde.
- Bei den Variablen cdtc07a bis cdtc07d, cdtc19a, cdtc19b, ddtc07a bis ddtc07d, ddtc19a, ddtc19b, edtc07a bis edtc07d, edtc19a, edtc19b, fdtc07a bis fdtc07d, fdtc19a und fdtc19b gibt es die Kategorie „trifft nicht zu“ die aktiv von den Befragten angeklickt werden konnte und der analog zur Fragebogenprogrammierung der Code 2 (nicht der Code -988) zugewiesen wurde. Auf den Zusatz „(Antwortkategorie)“ im Wertelabel wurde hier verzichtet.

<sup>12</sup> Eine mögliche vierte Gruppe würde spezielle Missingcodierungen umfassen, die im Rahmen der Datenaufbereitung eines konkreten Datensatzes nur für einzelne Items vergeben wurden. Zudem gibt es auch eine fünfte Gruppe für „andere fehlende Werte“ wie Datenverlust.

<sup>13</sup> Bei den Interviewabbrüchen kann es vorkommen, dass Werte als -989 „filterbedingt fehlend“ codiert sind, die eigentlich ein Interviewabbruch sind, oder dass filterbedingt fehlende Werte fälschlicherweise als -966 „Interviewabbruch“ codiert sind. Es empfiehlt sich daher, für die Nachvollziehbarkeit der Filterführung immer auch den Fragebogen zu konsultieren und Interviewabbrüche nicht auszuwerten.

Tabelle 8: Systematik des FDZ-DZHW für fehlende Werte

Wertebereich	Code	Wertelabel
<b>-999 bis -990: Nonresponse</b>	-999	weiß nicht
	-998	<b>keine Angabe</b>
	-997	keine Angabe (Antwortkategorie)
	-996	<b>Interviewabbruch</b>
	-995	<b>keine Teilnahme (Panel)</b>
	-994	verweigert
<b>-989 bis -970: Nicht zutreffend</b>	-989	<b>filterbedingt fehlend</b>
	-988	<b>trifft nicht zu</b>
	-987	designbedingt fehlend (Fragebogensplit)
	-986	designbedingt fehlend (Welle) <sup>a</sup>
	-985	designbedingt fehlend (Kohorte) <sup>b</sup>
<b>-969 bis -950: Editierter fehlender Wert</b>	-969	<b>unbekannter fehlender Wert<sup>c</sup></b>
	-968	unplausibler Wert <sup>d</sup>
	-967	<b>anonymisiert</b>
	-966	<b>nicht bestimmbar<sup>e</sup></b>
	-965	ungültige Mehrfachnennung
<b>-949 bis -930: Item-spezifische fehlende Werte</b>	<i>(nicht vergeben)</i>	
<b>-929 bis -920: Andere fehlende Werte</b>	-929	Datenverlust

<sup>a</sup> Dieser Wert wird nur für Datensätze im Long-Format vergeben.

<sup>b</sup> Dieser Wert wird nur in gepoolten Datensätzen vergeben.

<sup>c</sup> Dieser Wert wird vergeben, wenn keinerlei Ursache rekonstruiert werden kann.

<sup>d</sup> Dieser Wert wird vergeben, wenn Angaben aufgrund unterschiedlicher Faktoren als nicht plausibel eingestuft werden.

<sup>e</sup> Dieser Wert wird vergeben, wenn eine eindeutige Codierung nicht möglich ist, z. B. wenn eine offene Angabe nicht vercodet werden konnte.

# 7 Repräsentativität der Daten

**[Vergleich zur Grundgesamtheit]** Die Anpassung der Daten der realisierten Stichprobe an ausgewählte Merkmale der Grundgesamtheit mittels Gewichtung ist ein übliches Vorgehen, um systematischen Selektionsprozessen entgegenzuwirken und damit die Repräsentativität der Daten zu optimieren. Aufgrund der Datenlage ist eine Anpassungsgewichtung an die Grundgesamtheit jedoch nicht möglich, da es keine zeitlich passenden, unverzerrten offiziellen amtlichen Daten zu den Promovierenden in Deutschland gibt. Zwar wurden zum 1. Dezember 2017 erstmalig die registrierten Promovierenden vom Statistischen Bundesamt erfasst, allerdings waren die Promovierenden in dieser Statistik untererfasst, da die erforderlichen organisatorischen und technischen Strukturen, die eine unverzerrte Erfassung ermöglicht hätten, zu diesem Zeitpunkt noch nicht vorhanden waren (vgl. Vollmar, 2019). Auch die zweite Erhebung zum Stichtag 1. Dezember 2018 weist weiterhin eine Untererfassung auf (Statistisches Bundesamt, 2019c). Um dennoch einen Überblick über eine mögliche ungleiche Verteilung zwischen Stichprobe und der Grundgesamtheit zu erhalten, wurde die Verteilung der Personen nach Fächergruppen zwischen der realisierten Stichprobe mit den Daten des Statistischen Bundesamtes (Destatis) (2019c) verglichen (Tabelle 9).<sup>14</sup>

Tabelle 9: Verteilung der Promovierenden nach Destatis 2019 und in Nacaps 2018, 1. Welle

Fächergruppen	Destatis 2019		Nacaps 2018, 1. Welle		
	Absolut	in %	Absolut	in %	Diff.
Geisteswissenschaften	21.688	12,5	2.919	11,8	-0,7
Sport	1.099	0,6	165	0,7	0,1
Rechts-, Wirtschafts- und Sozialwissenschaften	30.049	17,3	4.884	19,7	2,4
Mathematik, Naturwissenschaften	41.810	24,1	7.024	28,3	4,2
Humanmedizin/Gesundheitswissenschaften	36.925	21,2	3.666	14,8	-6,4
Agrar-, Forst- und Ernährungswissenschaften, Veterinärmedizin	5.158	3,0	738	3,0	0,0
Ingenieurwissenschaften	30.944	17,8	4.318	17,4	-0,4
Kunst, Kunstwissenschaft	3.787	2,2	494	2,0	-0,2
Außerhalb der Studienbereichsgliederung	2.319	1,3	578	2,3	1,0
	173.779	100,0	24.786*	100,0	

\* Die Abweichung zur in Kapitel 5.1 genannten Fallzahl ist im Wesentlichen auf fehlende Werte in der Fächervariable zurückzuführen.

Demnach sind Personen, die in den Rechts-, Wirtschafts- und Sozialwissenschaften, in der Fächergruppe Mathematik, Naturwissenschaften sowie in Fächern außerhalb der Studienbereichsgliederung promovieren, in der Erhebung leicht überrepräsentiert. Dagegen sind Promovierende aus der

<sup>14</sup> Im Herbst 2020 wurde die erste offizielle Promovierendenstatistik vorgelegt (Statistisches Bundesamt, 2020). In dieser Statistik ist zwar die offizielle Zahl der Promovierenden gegenüber der Mitteilung aus 2019 noch einmal deutlich auf 182.778 Personen angestiegen, allerdings sind die Verschiebungen in den Anteilen zwischen den Fächergruppen eher gering.

Fächergruppe Humanmedizin/Gesundheitswissenschaften etwas unterrepräsentiert. In den anderen Fächergruppen sind die Verteilungen in beiden Datenquellen nahezu gleich.

Ein Vergleich der Verteilungen nach den Merkmalen Geschlecht und Fächergruppe (Tabelle 10) lässt erkennen, dass in Nacaps Frauen aus der Fächergruppe Rechts-, Wirtschafts- und Sozialwissenschaften leicht überrepräsentiert sind. In der Fächergruppe Mathematik/Naturwissenschaften sind dagegen sowohl Männer als auch Frauen etwas überrepräsentiert. In der Fächergruppe Humanmedizin/Gesundheitswissenschaften sind sowohl Männer als auch Frauen unterrepräsentiert; der Unterschied ist bei Frauen jedoch deutlich größer als bei Männern. In den übrigen Fächergruppen gibt es keine nennenswerten Unterschiede. Ebenso sind die Unterschiede in der Verteilung nach dem Geschlecht insgesamt zwischen der Nacaps-Stichprobe (51 Prozent Männer, 48 Prozent Frauen, 1 Prozent divers) und den Informationen aus der amtlichen Statistik (53 Prozent Männer und 47 Prozent Frauen) eher gering.

**Tabelle 10:** Verteilung von männlichen und weiblichen Promovierenden nach Destatis 2019 und in Nacaps 2018, 1. Welle

Fächergruppen	Destatis 2019				Nacaps 2018, 1. Welle			
	absolut, männ- lich	in % männ- lich	absolut, weib- lich	in % weib- lich	absolut, männ- lich*	in % männ- lich	absolut, weib- lich*	in % weib- lich
Geisteswissenschaften	9.541	10,3	12.147	15,0	1.009	9,1	1.595	15,2
Sport	605	0,7	494	0,6	67	0,6	85	0,8
Rechts-, Wirtschafts- und Sozialwissenschaften	15.600	16,8	14.449	17,9	2.000	18,0	2.307	21,9
Mathematik, Naturwissenschaften	24.230	26,1	17.380	21,5	3.398	30,6	2.747	26,1
Humanmedizin/Gesundheitswissenschaften	13.876	15,0	23.049	28,5	1.206	10,9	1.881	17,9
Agrar-, Forst- und Ernährungswissenschaften, Veterinärmedizin	1.890	2,0	3.268	4,0	194	1,7	481	4,6
Ingenieurwissenschaften	24.418	26,3	6.526	8,1	2.873	25,9	845	8,0
Kunst, Kunstwissenschaft	1.297	1,4	2.490	3,1	114	1,0	339	3,2
Außerhalb der Studienbereichsgliederung	1.221	1,3	1.098	1,4	242	2,2	248	2,4
	92.678	100,0	80.901	100,0	11.103	100,0	10.528	100,0

\* Die Abweichung zur in Tabelle 9 genannten Fallzahl ist im Wesentlichen auf fehlende Werte in der Geschlechtsvariable zurückzuführen. Zudem wurden Personen mit der Angabe „divers“ bei dieser Auswertung nicht berücksichtigt.

Darüber hinaus können die Verteilungen mit weiteren Datenquellen abgeglichen werden, die jedoch ihrerseits jeweils Unsicherheiten bezüglich ihrer Repräsentativität aufweisen. Dazu zählt u. a. die Statistik zu abgeschlossenen Promotionen des Jahres 2018 (Statistisches Bundesamt, 2019d). Ein Problem bezüglich des Abgleichs mit dieser Datenquelle ist jedoch, dass der Abbruch von Promotionen sich möglicherweise zwischen Fächergruppen, Geschlechtern oder anderen Merkmalen systematisch unterscheidet. Insofern kann ein solcher Abgleich zwar vorgenommen werden, ist jedoch vermutlich mit größeren Unsicherheiten behaftet als ein Vergleich mit der oben vorgenommenen Statistik.<sup>15</sup> Ebenso ist ein Abgleich mit einer älteren Erhebung des Statistischen Bundesamtes zu Promovierenden in

<sup>15</sup> Nimmt man diesen Vergleich dennoch vor, sind Promovierende der Geisteswissenschaften (+5%) und der Rechts-, Wirtschafts- und Sozialwissenschaften (+5%) in Nacaps jeweils überrepräsentiert, Promovierende der Humanmedizin/Gesundheitswissenschaften dagegen unterrepräsentiert (-10%).

Deutschland möglich. Dabei handelt es sich um eine sog. §7-Erhebung, die außerhalb der regelmäßigen statistischen Erhebungen erfolgt (Statistisches Bundesamt, 2016). Nach §7 des Bundesstatistikgesetzes können Sondererhebungen (max. zweimalig) durchgeführt werden, um Daten für besondere Zwecke zu generieren. Da eine Promovierendenstatistik bis zur Novelle des Hochschulstatistikgesetzes im Jahr 2016 nicht existierte und die Studierendenstatistik eine systematische Untererfassung der Promovierenden aufweist, wurde die §7-Erhebung erstmalig für das Wintersemester 2010/11 und dann erneut für das Wintersemester 2014/15 durchgeführt. Damit ist auch die Erhebung für das Wintersemester 2014/15 einerseits etwas älteren Datums und andererseits basieren die Informationen auf Befragungen von Professor\*innen und Promovierenden. Insofern ist ein Vergleich zwischen den Informationen über die Grundgesamtheit aus dieser Erhebung und Nacaps ebenso mit Unsicherheiten behaftet.<sup>16</sup> Eine weitere Datenquelle für einen Abgleich der Verteilungen ist die amtliche Statistik über immatrikulierte Studierende, in der ebenfalls eingeschriebene Promovierende erfasst sind (Statistisches Bundesamt, 2019e). In dieser Datenquelle gibt es jedoch eine deutliche Untererfassung der Promovierenden gegenüber den anderen beiden Datenquellen über Promovierende, denn es sind hier nur rund 110.000 Promovierende verzeichnet. Somit ist auch hier zu erwarten, dass ein Abgleich zur Verteilung in Nacaps mit Unsicherheiten behaftet ist.<sup>17</sup>

Da derzeit keine vollkommen zuverlässigen Datenquellen zum Abgleich von Fächer- und Geschlechterverteilungen für Nacaps vorliegen, kann keine sichere Bewertung der Repräsentativität der Daten vorgenommen werden. Aus dem gleichen Grund können keine Gewichtungsvariablen erstellt werden. Die dargestellten statistischen Abgleiche sollen die Datennutzenden jedoch ermächtigen, eine erste Bewertung der Repräsentativität der Daten vornehmen zu können.

**[Verteilung der Hochschulen]** Ergänzend wurde zudem die Verteilung der teilnehmenden und nicht-teilnehmenden Hochschulen analysiert. Hinsichtlich der regionalen Verteilung gibt es nur geringfügige Unterschiede zwischen den Regionen Nord, Ost, Süd und West. Hochschulen aus dem Osten sind mit einer Teilnahmequote von 27,6 Prozent leicht unter-, Hochschulen aus dem Norden dagegen leicht überrepräsentiert (41,9 Prozent). Aus dem Westen haben 32,7 Prozent und aus dem Süden 34,9 Prozent der Hochschulen teilgenommen. Zudem sind größere Hochschulen etwas überrepräsentiert, kleinere Hochschulen dagegen etwas seltener vertreten. Gemessen an der Zahl der Professorinnen und Professoren haben 21,2 Prozent der kleineren Hochschulen, 37,3 Prozent der mittelgroßen und 44,2 Prozent der großen Hochschulen an Nacaps teilgenommen.<sup>18</sup> Für die Aussagekraft des Samples ist diese relativ hohe Teilnahmequote der großen Hochschulen vermutlich eher von Vorteil, da so die Gesamtzahl der teilnehmenden Promovierenden deutlich erhöht wird.

**[Ausfallanalysen]** Systematischer bzw. selektiver Panel-Nonresponse kann zu verzerrten Ergebnissen führen. Um zu überprüfen, ob Panelausfälle zwischen den Befragungswellen in systematischer Form auftreten, wurden Ausfallanalysen durchgeführt. Zu diesem Zweck wurden fünf logistische Regressionsanalysen gerechnet, die die Wahrscheinlichkeit einer Teilnahme an den Wellen 2 bis 6 schätzen (vgl. Tabelle 11 und Tabelle 12). Es sei darauf hingewiesen, dass Fälle, die zum Zeitpunkt der Berechnungen bereits um komplette Löschung ihrer Daten gebeten hatten (vgl. Kapitel 5), nicht in die Berechnungen eingegangen sind. Gleichzeitig muss berücksichtigt werden, dass im Anschluss an die Berechnungen auch noch weitere Datenlöschungsfälle hinzugekommen sind.

<sup>16</sup> Bei einem Vergleich dieser Erhebungen gibt es wiederum Unterschiede im Bereich der Geisteswissenschaften. Dieses Mal sind Promovierende aus diesen Fächergruppen in Nacaps gegenüber der herangezogenen Datenquelle jedoch unterrepräsentiert (-6%); in den Rechts-, Wirtschafts- und Sozialwissenschaften (+3%) und der Humanmedizin/den Gesundheitswissenschaften (+5%) sind Promovierende in Nacaps dagegen überrepräsentiert.

<sup>17</sup> Bei einem Abgleich sind allerdings nur geringe Abweichungen zu der Fächerverteilung in Nacaps zu beobachten. Lediglich im Bereich Humanmedizin/Gesundheitswissenschaften sind Promovierende in Nacaps überrepräsentiert (+4%).

<sup>18</sup> Zu berücksichtigen ist jedoch, dass angesichts der geringen Zahl an Hochschulen insgesamt bereits geringfügige Veränderungen der Teilnahmезahlen deutliche Auswirkungen auf die prozentuale Verteilung (insbesondere bei der regionalen Verteilung) haben. Aus diesem Grund wird u.a. darauf verzichtet, die Teilnahmequote auf Ebene der Bundesländer zu berichten.

Tabelle 11: Logistische Regression der Teilnahmewahrscheinlichkeit an Welle 2, 3 und 4

		2. Welle	3. Welle	4. Welle	
Geschlecht (ref. weiblich)	männlich	0,761 (0,022)	***	0,759 (0,023)	***
	divers	0,432 (0,097)	***	0,609 (0,137)	*
	keine Angabe	0,642 (0,221)		0,671 (0,247)	0,517 (0,216)
Alter (ref. 1991-1993)	bis 1984	0,804 (0,036)	***	0,895 (0,042)	*
	1985-1990	0,892 (0,030)	***	0,938 (0,033)	0,958 (0,035)
	1994 und jünger	1,099 (0,074)		1,004 (0,072)	0,979 (0,072)
	keine Angabe	0,276 (0,053)	***	0,358 (0,072)	***
Land der HZB (ref. Deutschland)	Ausland	0,549 (0,021)	***	0,473 (0,020)	***
	keine Angabe	0,068 (0,028)	***	0,039 (0,020)	***
höchster Bildungsabschluss der Eltern (ref. kein HSA)	Promotion	1,069 (0,044)		1,024 (0,044)	1,143 (0,049)
	anderer HSA	1,105 (0,033)	***	1,095 (0,034)	**
	keine Angabe	0,470 (0,067)	***	0,575 (0,086)	***
Selbstwirksamkeit (ref. 3 „mittel“)	1-2 gering	0,998 (0,119)		1,045 (0,131)	1,080 (0,137)
	4-5 hoch	1,170 (0,052)	***	1,106 (0,051)	*
	keine Angabe	0,382 (0,098)	***	0,501 (0,130)	**
Promotionsstatus (ref. promovierend)	abgeschlossen	2,997 (0,770)	***	2,049 (0,536)	**
	unterbrochen	0,896 (0,093)		0,976 (0,106)	1,052 (0,113)
	abgebrochen	1,677 (0,487)		0,854 (0,264)	1,225 (0,376)
	keine Angabe	7,076 (7,861)		7,908 (8,990)	-
Fächergruppe (ref. Geisteswissenschaften)	Sport	0,989 (0,158)		0,883 (0,146)	1,035 (0,169)
	Rechts-, Wirtschafts- und Sozialwiss.	0,876 (0,043)	**	0,837 (0,043)	***
	Mathematik, Naturwiss.	0,727 (0,035)	***	0,790 (0,040)	***
	Humanmedizin, Gesundheitswiss.	0,572 (0,033)	***	0,549 (0,033)	***
	Agrar-, Forst- und Ernährungswiss.	0,660 (0,059)	***	0,724 (0,067)	***
	Ingenieurwiss.	0,654 (0,036)	***	0,654 (0,038)	***
	Kunst, Kunstwiss.	0,913 (0,092)		0,917 (0,096)	0,882 (0,093)
	keine Angabe	0,757 (0,073)	**	0,803 (0,081)	*
				0,724 (0,075)	**
Bundesland der Promotionshochschule (ref. Schleswig-Holstein)	Hamburg	1,053 (0,111)		1,007 (0,111)	1,031 (0,114)
	Niedersachsen	0,793 (0,087)	*	0,826 (0,095)	0,825 (0,095)
	Nordrhein-Westfalen	0,801 (0,074)	*	0,819 (0,079)	*
	Hessen	0,794 (0,079)	*	0,864 (0,090)	0,879 (0,092)
	Rheinland-Pfalz	0,747 (0,076)	**	0,780 (0,083)	*
	Baden-Württemberg	0,967 (0,091)		0,966 (0,095)	0,991 (0,098)
	Bayern	0,942 (0,087)		0,938 (0,091)	0,900 (0,088)
	Saarland	0,660 (0,078)	***	0,788 (0,097)	0,662 (0,084)
	Berlin	0,984 (0,102)		1,052 (0,114)	1,034 (0,112)
	Brandenburg	1,153 (0,134)		1,087 (0,131)	1,140 (0,138)
	Sachsen	1,010 (0,113)		1,088 (0,127)	1,022 (0,120)
	Sachsen-Anhalt	0,791 (0,109)		0,879 (0,127)	0,750 (0,112)
	Thüringen	0,751 (0,087)	*	0,830 (0,101)	0,936 (0,114)
	keine Angabe	-		-	-
erwerbstätig in Wissenschaft (ref. nein)	Ja	1,255 (0,039)	***	1,274 (0,042)	***
	keine Angabe	0,476 (0,191)		0,463 (0,211)	0,379 (0,204)
	Pseudo-R <sup>2</sup>	0,136		0,118	0,113
	N	28.349		28.349	28.147

Anmerkungen: Odds-Ratios, Standardfehler in Klammern; \*\*\* p<0,001, \*\* p<0,01, \* p<0,05

Tabelle 12: Logistische Regression der Teilnahmewahrscheinlichkeit an Welle 5 und 6

		5. Welle	6. Welle	
Geschlecht (ref. weiblich)	männlich	0,738 (0,024)	***	0,812 (0,027) ***
	divers	0,635 (0,157)		0,421 (0,126) **
	keine Angabe	0,735 (0,303)		0,808 (0,343)
Alter (ref. 1991-1993)	bis 1984	1,054 (0,054)		1,084 (0,057)
	1985-1990	1,041 (0,041)		1,062 (0,043)
	1994 und jünger	0,851 (0,070)	*	0,828 (0,070) *
	keine Angabe	0,413 (0,093)	***	0,351 (0,089) ***
Land der HZB (ref. Deutschland)	Ausland	0,457 (0,022)	***	0,461 (0,023) ***
	keine Angabe	0,030 (0,020)	***	0,048 (0,029) ***
höchster Bildungsabschluss der Eltern (ref. kein HSA)	Promotion	1,159 (0,054)	**	1,203 (0,057) ***
	anderer HSA	1,178 (0,040)	***	1,165 (0,041) ***
	keine Angabe	0,550 (0,097)	***	0,584 (0,106) **
Selbstwirksamkeit (ref. 3 „mittel“)	1-2 gering	1,021 (0,142)		1,014 (0,148)
	4-5 hoch	1,153 (0,059)	**	1,188 (0,063) **
	keine Angabe	0,493 (0,150)	*	0,591 (0,178)
Promotionsstatus (ref. promovierend)	abgeschlossen	2,042 (0,623)	*	1,900 (0,573) *
	unterbrochen	0,917 (0,108)		0,952 (0,115)
	abgebrochen	1,202 (0,418)		1,038 (0,362)
	keine Angabe	-	-	-
Fächergruppe (ref. Geisteswissenschaften)	Sport	0,824 (0,146)		0,862 (0,156)
	Rechts-, Wirtschafts- und Sozialwiss.	0,851 (0,046)	**	0,861 (0,048) **
	Mathematik, Naturwiss.	0,767 (0,041)	***	0,823 (0,046) ***
	Humanmedizin, Gesundheitswiss.	0,570 (0,037)	***	0,630 (0,042) ***
	Agrar-, Forst- und Ernährungswiss.	0,561 (0,059)	***	0,628 (0,067) ***
	Ingenieurwiss.	0,667 (0,042)	***	0,604 (0,039) ***
	Kunst, Kunsthiss.	0,722 (0,083)	**	0,849 (0,098)
	keine Angabe	0,805 (0,088)	*	0,694 (0,082) **
Bundesland der Promotionshochschule (ref. Schleswig-Holstein)	Hamburg	1,005 (0,119)		0,892 (0,106)
	Niedersachsen	0,894 (0,110)		0,808 (0,100)
	Nordrhein-Westfalen	0,861 (0,089)		0,740 (0,077) **
	Hessen	0,839 (0,094)		0,706 (0,080) **
	Rheinland-Pfalz	0,736 (0,086)	**	0,645 (0,076) ***
	Baden-Württemberg	0,994 (0,106)		0,914 (0,097)
	Bayern	0,912 (0,096)		0,834 (0,087)
	Saarland	0,738 (0,100)	*	0,642 (0,089) **
	Berlin	1,054 (0,123)		0,922 (0,108)
	Brandenburg	1,151 (0,149)		0,916 (0,120)
	Sachsen	1,177 (0,147)		0,960 (0,121)
	Sachsen-Anhalt	0,768 (0,124)		0,779 (0,126)
	Thüringen	0,832 (0,110)		0,707 (0,095) **
	keine Angabe	-	-	-
erwerbstätig in Wissenschaft (ref. nein)	Ja	1,155 (0,041)	***	1,242 (0,046) ***
	keine Angabe	0,320 (0,197)		0,217 (0,161) *
Pseudo-R <sup>2</sup>		0,100		0,096
N		28.134		28.134

Anmerkungen: Odds-Ratios, Standardfehler in Klammern; \*\*\* p<0,001, \*\* p<0,01, \* p<0,05

Um zu untersuchen, inwieweit die Teilnahmewahrscheinlichkeit an den Folgewellen mit den Merkmalen der Befragten in Zusammenhang steht (MAR), enthalten die Erklärungsmodelle als unabhängige Variablen demographische Eigenschaften (Geschlecht, Alter, Land der Hochschulzugangsberechtigung (HZB), höchster Bildungsabschluss der Eltern), Persönlichkeitsmerkmale (Selbstwirksamkeit) sowie kontextuelle Merkmale der Promotion (Promotionsstatus, Promotionsfach (Fächergruppe), Bundesland der Promotionshochschule, Tätigkeit in der Wissenschaft während der Promotion), die im Rahmen der Erstbefragung erhoben und von allen Statusgruppen erfragt wurden. Fehlende Werte auf den erklärenden Variablen wurden jeweils als eigene Kategorie in den Analysen berücksichtigt<sup>19</sup>. Dabei zeigt sich, dass Unit-Nonresponse auf den Modellvariablen aus der Erstbefragung größtenteils ein starker Prädiktor für Panelausfälle in den Folgebefragungen ist.

Im Vergleich zu weiblichen Befragten haben männliche Personen signifikant seltener an den Folgebefragungen teilgenommen. Eine geringere Teilnahmewahrscheinlichkeit lässt sich für die zweite, dritte und sechste Welle zudem auch für Personen mit der Geschlechtsidentität „divers“ finden. Auch für das Alter lässt sich ein Einfluss auf das Teilnahmeverhalten nachweisen. Personen, die vor 1991 geboren wurden, haben signifikant seltener an der zweiten Befragung teilgenommen; für die älteste Kategorie, also die bis einschließlich 1984 Geborenen, lässt sich ein negativer Effekt auch für die dritte Welle zeigen. Personen mit Geburtsjahrgängen ab 1994 haben eine geringere Teilnahmewahrscheinlichkeit in den Wellen 5 und 6. Eine Hochschulzugangsberechtigung, die im Ausland erworben wurde, verringert die Teilnahmewahrscheinlichkeit an allen fünf Folgebefragungen deutlich. Verglichen mit Befragten aus nicht-akademischen Elternhäusern begünstigt ein Hochschulabschluss einer oder beider Elternteile die Teilnahmewahrscheinlichkeit an allen fünf Folgebefragungen. Ein positiver Effekt auf die Teilnahmewahrscheinlichkeit ab der vierten Welle lässt sich zudem nachweisen, wenn mindestens ein Elternteil selbst eine Promotion abgeschlossen hat. Zudem erhöht eine hohe Selbstwirksamkeitserwartung die Wahrscheinlichkeit, sich an den Folgebefragungen zu beteiligen.

Befragte, die ihre Promotion zum Zeitpunkt der Erstbefragung bereits abgeschlossen haben, nehmen im Vergleich zu Promovierenden signifikant häufiger an den Folgebefragungen teil. Eine unterbrochene oder abgebrochene Promotion hat hingegen keinen Effekt auf die Teilnahmewahrscheinlichkeit. Gleichwohl lassen sich Fächereffekte beobachten. Im Vergleich zur Referenzkategorie der Befragten aus der Fächergruppe Geisteswissenschaften lässt sich – mit Ausnahme der Fächergruppen Sport und Kunst, Kunswissenschaften (außer in Welle 5) – für alle Fächergruppen ein signifikant negativer Effekt auf die Teilnahmewahrscheinlichkeit der Folgewellen nachweisen. Deutliche Unterschiede in der Beteiligung an den Folgewellen lassen sich auch bei den Bundesländern finden, in denen die Hochschulen angesiedelt sind. Im Vergleich zu Schleswig-Holstein, das hier als Referenzkategorie fungiert, lassen sich signifikant negative Effekte für Niedersachsen (zweite Welle), Nordrhein-Westfalen (zweite, dritte & sechste Welle), Hessen (zweite & sechste Welle), Rheinland-Pfalz (alle Wellen), Saarland (zweite, vierte, fünfte & sechste Welle) sowie Thüringen (zweite & sechste Welle) nachweisen. Schließlich hat auch das Tätigkeitsfeld während der Promotion einen Einfluss auf die Beteiligung an den Folgebefragungen. So begünstigt eine Erwerbstätigkeit in der Wissenschaft während der Promotion die Teilnahme an den Nacaps-Folgebefragungen.

---

<sup>19</sup> Eine Ausnahme bilden Fälle mit fehlenden Werten auf der Variable „Promotionsstatus“ in den Modellen zur Schätzung der Teilnahmewahrscheinlichkeiten an der vierten, fünften und sechsten Welle. Alle Befragten mit fehlenden Werten auf dieser Variablen haben an den entsprechenden Befragungen nicht teilgenommen, sodass Stata diese 202 Fälle aus den Analysen ausgeschlossen hat.

# 8 Anonymisierung

**[Datenschutzrechtlicher Rahmen]** Für personenbezogene Daten<sup>20</sup>, die in freiwilligen Befragungen durch das DZHW erhoben werden, gelten die EU-Datenschutz-Grundverordnung (EU-DSGVO) und das Bundesdatenschutzgesetz in seiner Neufassung vom 30. Juni 2017.<sup>21</sup> Danach sind personenbezogene Daten für die Weitergabe zur wissenschaftlichen Sekundärnutzung (ohne Vorliegen einer Einverständniserklärung zur Sekundärnutzung der personenbezogenen Daten) in der Regel derart aufzubereiten, dass „die personenbezogenen Daten ohne Hinzuziehung zusätzlicher Informationen nicht mehr einer spezifischen betroffenen Person zugeordnet werden können, sofern diese zusätzlichen Informationen gesondert aufbewahrt werden und technischen und organisatorischen Maßnahmen unterliegen, die gewährleisten, dass die personenbezogenen Daten nicht einer identifizierten oder identifizierbaren natürlichen Person zugewiesen werden“ können (Art. 4 Abs. 5 DSGVO; s. auch Art. 89 DSGVO sowie Erwägungsgrund 26 DSGVO).

**[Sicherstellung des Datenschutzes im FDZ-DZHW]** Im FDZ-DZHW wird der Datenschutz der Befragten über rechtlich-organisatorische Maßnahmen, technische Maßnahmen und statistische Maßnahmen sichergestellt<sup>22</sup>, die in unterschiedlicher Weise kombiniert werden können. Aus den Kombinationen ergeben sich vier potenzielle Zugangswege, die in Tabelle 13 im Überblick dargestellt sind.<sup>23</sup>

Tabelle 13: Zugangswege im FDZ-DZHW

	Campus Use File (CUF): Download	Scientific Use File (SUF): Download	Scientific Use File (SUF): Remote-Desktop	Scientific Use File (SUF): On-Site
rechtlich-organisatorische Maßnahmen	sehr gering	gering	moderat	hoch
technische Maßnahmen	sehr gering	gering	moderat	hoch
statistische Maßnahmen	sehr hoch	hoch	moderat	gering

Je stärker der Datenzugang rechtlich-organisatorisch und technisch kontrolliert und beschränkt wird, desto geringer ist das Risiko einer De-Anonymisierung der Daten. In der Folge müssen die Daten umso weniger mittels statistischer Maßnahmen anonymisiert, also um Informationen reduziert, werden – das heißt, desto größer bleibt ihr Analysepotenzial.

<sup>20</sup> „Personenbezogene Daten“ (sind) alle Informationen, die sich auf eine identifizierte oder identifizierbare natürliche Person (im Folgenden ‚betroffene Person‘) beziehen; als identifizierbar wird eine natürliche Person angesehen, die direkt oder indirekt, insbesondere mittels Zuordnung zu einer Kennung wie einem Namen, zu einer Kennnummer, zu Standortdaten, zu einer Online-Kennung oder zu einem oder mehreren besonderen Merkmalen identifiziert werden kann, die Ausdruck der physischen, physiologischen, genetischen, psychischen, wirtschaftlichen, kulturellen oder sozialen Identität dieser natürlichen Person sind“ (Art. 4 Abs. 1 DSGVO).

<sup>21</sup> Die DSGVO gilt grundsätzlich innerhalb der EU und somit ebenfalls für das DZHW. Das BDSG in seiner Neufassung vom 30. Juni 2017 (Gesetz zur Anpassung des Datenschutzrechts an die Verordnung (EU) 2016/679 und zur Umsetzung der Richtlinie (EU) 2016/680 (Datenschutz-Anpassungs- und Umsetzungsgesetz EU DSAnpUG-EU)) kommt teils zusätzlich zur Anwendung, da die DZHW GmbH juristisch als öffentliche Stelle des Bundes betrachtet wird (§ 2 Abs. 3 BDSG). Der Bund hält die absolute Mehrheit der Anteile der DZHW GmbH und das Institut erfüllt Aufgaben der öffentlichen Verwaltung des Bundes im weitesten Sinn.

<sup>22</sup> Das Datenschutzkonzept des FDZ ist angelehnt an den Portfolio-Ansatz von Lane, Heus und Mulcahy (2008, S. 6ff.), an dem sich bereits das Leibniz-Institut für Bildungsverläufe (LIfBi) (vgl. Koberg, 2016a, S. 699ff.) und das FDZ der Bundesagentur für Arbeit im Institut für Arbeitsmarkt- und Berufsforschung (vgl. Hochfellner, Müller, Schmucker & Roß, 2012, S. 9f.) orientieren.

<sup>23</sup> Für genauere Informationen vgl. <https://www.fdz.dzhw.eu/de/datennutzung>

Für die Daten von Nacaps 2018 gibt es alle Zugangswege. Im Folgenden werden die konkreten durchgeföhrten statistischen Anonymisierungsmaßnahmen genauer erläutert.

**[Statistische Anonymisierungsmaßnahmen]** Im Rahmen der statistischen Anonymisierungsmaßnahmen wurden zunächst alle Informationen darauf geprüft, ob sich über sie Personen direkt identifizieren lassen. Diese direkten Identifikatoren, wie Namen, Adressen und E-Mail-Adressen, wurden bereits während der Feldphase in einem separaten Datensatz erfasst und sind somit per se nicht in den Daten enthalten. Die Original-Identifikationsnummer wurde entfernt und durch eine neue zufällig vergebene Identifikationsnummer ersetzt.

Darüber hinaus wurde der Empfehlung von Ebel und Meyermann gefolgt, offene Angaben zu löschen „selbst wenn die jeweiligen Fragestellungen an sich unproblematisch sind. Denn es besteht die Gefahr, dass Studienteilnehmer/-innen bei eigentlich unbedenklichen Fragen mit offener Antwortmöglichkeit kritische Informationen preisgegeben haben, die zu einer Identifikation führen könnten“ (Ebel & Meyermann, 2015, S. 5).

Anschließend wurden die Quasi-Identifikatoren bestimmt, also Informationen, die in Kombination oder durch die Anspielung externer Informationen geeignet sind, eine Person indirekt zu identifizieren. Bei Nacaps 2018 sind dies beispielsweise bestimmte Informationen zu Promotionseckdaten (z. B. Promotionshochschule, -fach oder -note) bzw. zum Promotionsverlauf (z. B. Zeitpunkt des Abschlusses), zu Promotionsprogrammen und Promotionsförderung, zum/zur Betreuer\*in der Promotion, zu Auslandsaufenthalten und wissenschaftlichen Aktivitäten (z. B. Publikationen und Anträge), zur Erwerbstätigkeit (z. B. Branche, Stellung, Arbeitsort) und Weiterbildung, zu Bildungsetappen vor der Promotion, zur Soziodemographie (z. B. Geburtsjahr, Geburtsland/Staatsangehörigkeit oder Informationen zu den Eltern) und zur persönlichen Lebenssituation (z. B. Kinder, Hauptwohnsitz). Um eine eindeutige Zuordnung der Daten zu einer Person zu unterbinden, wurde für diese Schlüsselmerkmale – je nach Zugangsweg – festgelegt, ob sie ggf. nur in aggregierter Form oder gar nicht freigegeben werden dürfen.

Zuletzt wurde geprüft, ob in den Daten sensible Informationen, z. B. zur Gesundheit, sexuellen Orientierung und zu politischen Einstellungen, enthalten waren. Diese Informationen eignen sich zwar nicht zwingend zur Re-Identifikation von Individuen oder Institutionen, stellen jedoch im Falle einer De-Anonymisierung ein potenziell erhöhtes Schadensrisiko für die Befragten dar (vgl. Koberg, 2016a, S. 694) und sind daher besonders schützenswert (Art. 9 DSGVO, Erwägungsgrund 51 DSGVO). Bei Nacaps 2018 sind dies beispielsweise Informationen zur Gesundheit, zur politischen Partizipation, zu Konflikten, Diskriminierung, sexueller Belästigung und Mobbing. Auch für die sensiblen Informationen wurde – je nach Zugangsweg – festgelegt, ob sie ggf. nur in aggregierter Form oder gar nicht freigegeben werden dürfen.

Dem Datensatzreport kann entnommen werden, welche Variablen in den bereitgestellten Daten enthalten sind und über welche(n) Zugangsweg(e) sie jeweils nutzbar sind.

Für den Zugangsweg „Campus Use File (CUF): Download“ wurde neben der Informationsreduktion auf Variablenebene eine per Zufallsauswahl gewonnene Substichprobe der Daten (25 % der Befragten) gezogen, also auch die Anzahl der Fälle reduziert.

# 9 Literatur

- AAPOR (2016). Standard Definitions. Final Dispositions of Case Codes and Outcome Rates for Surveys. <https://aapor.org/wp-content/uploads/2022/11/Standard-Definitions20169theditionfinal.pdf> (zugegriffen am: 18.11.2025).
- Brandtstädtter, J. & Renner, G. (1990). Tenacious goal pursuit and flexible goal adjustment: Explication and age-related analysis of assimilative and accommodative strategies of coping. *Psychology and Aging*, 5(1), S. 58–67. doi: 10.1037/0882-7974.5.1.58.
- Ebel, T. & Meyermann, A. (2015). Hinweise zur Anonymisierung von quantitativen Daten. *Forschungsdaten Bildung informiert*. Bd.3. Frankfurt am Main: Deutsches Institut für Internationale Pädagogische Forschung. doi: 10.25656/01:21970.
- Hochfellner, D., Müller, D., Schmucker, A. & Roß, E. (2012). Datenschutz am Forschungsdatenzentrum. (FDZ-Methodenreport Nr. 6). Nürnberg: Institut für Arbeitsmarkt- und Berufsforschung (IAB).
- Koberg, T. (2016a). Disclosing the National Educational Panel Study. In H.-P. Blossfeld, J. v. Maurice, M. Bayer & J. Skopek (Hrsg.), *Methodological Issues of Longitudinal Surveys. The example of the National Educational Panel Study* (S. 691–708). Wiesbaden: Springer VS. doi: 10.1007/978-3-658-11994-2\_38.
- Koberg, T. (2016b). Starting Cohort 5: First-Year Students (SC5). SUF Version 6.0.0. Anonymization Procedures. [https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/6-0-0/SC5\\_6-0-0-Anonymisation.pdf](https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/6-0-0/SC5_6-0-0-Anonymisation.pdf) (zugegriffen am: 05.06.2020).
- Lane, J., Heus, P. & Mulcahy, T. (2008). Data access in a cyber world: Making use of cyberinfrastructure. *Transactions on Data Privacy*, 1(1), 2–16.
- Lenzner, T., Neuert, C. & Otto, W. (2015). Kognitives Pretesting. Mannheim: GESIS – Leibniz-Institut für Sozialwissenschaften (GESIS Survey Guidelines). doi: 10.15465/gesis-sg\_010.
- Prüfer, P. & Rexroth, M. (2000). Zwei-Phasen-Pretesting. (ZUMA-Arbeitsbericht 2000/08). Mannheim: Zentrum für Umfragen, Methoden und Analysen (ZUMA). <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-200864> (zugegriffen am: 16.04.2025).
- Rußmann, M., Netz, N. & Schwabe, U. (2025). Doctoral Students' Life Satisfaction Throughout the Covid-19 Pandemic: Inequalities by Parenthood and Gender. *Social Indicators Research*. doi: 10.1007/s11205-025-03648-0.
- Statistisches Bundesamt (2016). Promovierende in Deutschland. Wintersemester 2014/2015. Wiesbaden: Statistisches Bundesamt.
- Statistisches Bundesamt (2018). Bildung und Kultur. Studierende an Hochschulen. Wintersemester 2017/2018. (Fachserie 11/4.1). Fächersystematik (Fächergruppen, Studienbereiche und Studienfächer). Wiesbaden: Statistisches Bundesamt.
- Statistisches Bundesamt (2019a). Schlüsselverzeichnisse für die Studenten- und Prüfungsstatistik, Promovierendenstatistik WS 2018/2019 und SS 2019. Wiesbaden: Statistisches Bundesamt.
- Statistisches Bundesamt (2019b). Staats- und Gebietssystematik. Stand 01.04.2019. Wiesbaden: Statistisches Bundesamt.
- Statistisches Bundesamt (2019c). Bildung und Kultur. Promovierendenstatistik: Analyse zu Vollständigkeit und Qualität der zweiten Erhebung 2018. Wiesbaden: Statistisches Bundesamt.
- Statistisches Bundesamt (2019d). Bildung und Kultur. Prüfungen an Hochschulen 2018. (Fachserie 11/4.2). Wiesbaden: Statistisches Bundesamt.
- Statistisches Bundesamt (2019e). Bildung und Kultur. Studierende an Hochschulen. Wintersemester 2018/19. (Fachserie 11/4.1). Wiesbaden: Statistisches Bundesamt.

- Statistisches Bundesamt (2020). Statistik der Promovierenden 2019. Wiesbaden: Statistisches Bundesamt.
- Statistisches Bundesamt (2022a). Schlüsselverzeichnisse für die Studenten- und Prüfungsstatistik, Promovierendenstatistik und Gasthörerstatistik WS 2021/2022 und SS 2022. Wiesbaden: Statistisches Bundesamt.
- Statistisches Bundesamt (2022b). Bildung und Kultur. Studierende an Hochschulen. Wintersemester 2021/2022. (Fachserie 11/4.1). Fächersystematik (Fächergruppen, Studienbereiche und Studienfächer). Wiesbaden: Statistisches Bundesamt.
- Statistisches Bundesamt (2023a). Bildung und Kultur. Studierende an Hochschulen. Wintersemester 2022/2023. (Fachserie 11/4.1). Fächersystematik (Fächergruppen, Studienbereiche und Studienfächer). Wiesbaden: Statistisches Bundesamt.
- Statistisches Bundesamt (2023b). Staats- und Gebietssystematik. Stand 01.01.2023. Wiesbaden: Statistisches Bundesamt.
- Statistisches Bundesamt (2024). Schlüsselverzeichnisse für die Studenten- und Prüfungsstatistik, Promovierendenstatistik und Gasthörerstatistik WS 2023/2024 und SS 2024. Wiesbaden: Statistisches Bundesamt.
- Vollmar, M. (2019). Neue Promovierendenstatistik: Analyse der ersten Erhebung 2017. WISTA Wirtschaft und Statistik, 1, 68-79.
- Weichbold, M. (2019). Pretest. In N. Bauer & J. Blasius (Hrsg.), Handbuch Methoden der empirischen Sozialforschung (S. 349-356). Wiesbaden: Springer VS.
- Wegner, A. & Briedis, K. (2020). National Academics Panel Study – Die Längsschnittstudie zu Promovierenden und Promovierten in Deutschland. Wissenschaftliches Projektkonzept. [https://www.nacaps.de/files/downloads/nacaps\\_wisskonzept.pdf](https://www.nacaps.de/files/downloads/nacaps_wisskonzept.pdf) (zugegriffen am: 16.04.2025).

# 10 Anhang

## 10.1 Dokumentation der Herkunft sekundär genutzter Fragen

Aufgrund des Umfangs erfolgt die Dokumentation der Herkunft sekundär genutzter Fragen in einer separaten Datei. Diese ist unter dem folgenden Link abrufbar:

[https://metadata.fdz.dzhw.eu/public/files/data-packages/stu-nac2018\\$-3.0.0/attachments/nac2018\\_QuestionOrigin.xlsx](https://metadata.fdz.dzhw.eu/public/files/data-packages/stu-nac2018$-3.0.0/attachments/nac2018_QuestionOrigin.xlsx)

## 10.2 Referenzliste: Destatis-Studienbereiche zu projekteigener Fächergruppeneinteilung

Tabelle 14: Zuordnung der Destatis-Studienbereiche zu projekteigener Fächergruppeneinteilung

DZHW-Fächergruppe (Code)	DZHW-Fächergruppe (Label)	Destatis-Studienbereich (Code)
1	Agrar-, Forst-, Ern.wiss., Vet.	51; 57; 58; 59; 60
2	Architektur und Bauing.	66; 68
3	Biologie	42
4	Chemie	40
5	Elektrotechnik	64
6	Germanistik	9
7	Geschichte	5
8	Informatik	71
9	Kunst und Kunsthiss.	74; 75; 76; 77; 78
10	Maschinenbau	63
11	Mathematik	37
12	Medizin	48; 49; 50
13	Physik	39
14	Politik- und Sozialwiss.	23; 24; 25; 26; 27; 29; 34
15	Psych. und Erziehungswiss.	32; 33
16	Rechtswiss.	28
17	Sonstige Ingenieurwiss.	61; 62; 65; 67; 69; 70; 72
18	Sonstige Naturwiss.	36; 41; 43; 44
19	Sprach- und Kulturwiss.	1; 2; 3; 4; 6; 7; 8; 10; 11; 12; 13; 14; 18; 19
20	Wirtschaftswiss.	30; 31
21	Sport	22
-966	nicht bestimmbar	83