

Karsten Becker / Florence Baillet / Anne Weber

21. Sozialerhebung

Daten- und Methodenbericht zu der Erhebung der
wirtschaftlichen und sozialen Lage der Studierenden 2016

Daten- und Methodenbericht

April 2019

Autor(inn)en:
Karsten Becker
Florence Baillet, Anne Weber (Kapitel II und 8)

Der Bericht wurde dankenswerterweise durch unsere studentische Hilfskraft Birgit Schneider tatkräftig unterstützt.

Der vorliegende Bericht soll folgendermaßen zitiert werden:
Becker, K., Baillet, F. & Weber, A. (2019). 21. Sozialerhebung. Daten- und Methodenbericht zu der Erhebung der wirtschaftlichen und sozialen Lage der Studierenden 2016. Hannover: FDZ-DZHW.

Förderkennzeichen des Verbundprojektes von DSW und DZHW „21. Sozialerhebung des Deutschen Studentenwerks durchgeführt vom Deutschen Zentrum für Hochschul- und Wissenschaftsforschung“ mit Förderung des Bundesministeriums für Bildung und Forschung: M517000 und M517100

Herausgeber:
Deutsches Zentrum für Hochschul- und Wissenschaftsforschung GmbH (DZHW)
Lange Laube 12 | 30159 Hannover | www.dzhw.eu
Tel.: +49 511 450670-0 | Fax: +49 511 450670-960 | info@dzhw.eu

Geschäftsführung:
Prof. Dr. Monika Jungbauer-Gans
Karen Schlüter

Vorsitzender des Aufsichtsrats:
Ministerialdirigent Peter Greisler

Registergericht:
Amtsgericht Hannover | B 210251

Dieses Werk steht unter der Creative Commons "Namensnennung – Nicht kommerziell – Weitergabe unter gleichen Bedingungen 3.0 Deutschland Lizenz" (CC-BY-NC-SA)
<https://creativecommons.org/licenses/by-nc-sa/3.0/>



Inhaltsverzeichnis

Abbildungsverzeichnis	IV
Tabellenverzeichnis	IV
I Übersicht zur 21. Sozialerhebung (2016)	1
II Datennutzungshinweise	5
1 Inhalt und Anlage der Studie	8
2 Erhebungsinstrument	12
2.1 Inhalt des Erhebungsinstruments	12
2.2 Pretest	14
3 Grundgesamtheit und Stichprobenverfahren	16
4 Durchführung der Erhebung	18
5 Rücklauf	20
6 Datenaufbereitung	22
6.1 Datenübertragung	22
6.2 Codierung offener Angaben	22
6.3 Datenprüfung und Datenbereinigung	23
6.4 Generierung von Variablen	24
6.5 Datenstruktur und Dateiformat	25
6.6 Labels, Variablennamen und Missings	25
7 Gewichtung	28
8 Anonymisierung	31
9 Literaturverzeichnis	37

Abbildungsverzeichnis

Abbildung 1: Datenzugangswege und Analysepotential.....	6
Abbildung 2: Erinnerungskonzept der 21. Sozialerhebung (2016)	19
Abbildung 3: Entwicklung des unbereinigten Rücklaufs der 21. Sozialerhebung (2016) im Zeitverlauf (kumulativ, absolut und in %).....	21
Abbildung 4: Datenzugangsweg, statistischer Anonymisierungsgrad und Analysepotential der Daten der 21. Sozialerhebung (2016)	32

Tabellenverzeichnis

Tabelle 1: Themenfelder der 21. Sozialerhebung (Deutsche/Bildungsinländer(innen))	13
Tabelle 2: Brutto- und Nettostichproben sowie Rücklaufquoten der 21. Sozialerhebung (2016)....	20
Tabelle 3: Vercodete Merkmale und verwendete Codierlisten in der 21. Sozialerhebung	23
Tabelle 4: Themengebiete im Variablennamen der 21. Sozialerhebung (2016).....	25
Tabelle 5: Systematik fehlender Werte der 21. Sozialerhebung	27
Tabelle 6: Bereitgestellte Gewichte zur 21. Sozialerhebung (2016)	29
Tabelle 7: Maßnahmen der statistischen Anonymisierung der Daten der 21. Sozialerhebung (2016) nach Zugangsweg.....	34

I Übersicht zur 21. Sozialerhebung (2016)

Studienreihe	Sozialerhebung
Erhebung (Jahr)	21. Sozialerhebung (2016)
Erhebende Institution	Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (DZHW)
Gefördert von	Bundesministerium für Bildung und Forschung (BMBF)
Projektmitarbeiter(innen) (Projektleitung)	Beate Apolinarski; Karsten Becker; Philipp Bornkessel; Tasso Brandt; Sonja Heißenberg; <u>Elke Middendorff</u> ; Heike Naumann; Jonas Poskowsky
Themen	Studienmerkmale und Hochschulzugang Soziale Zusammensetzung der Studierenden Finanzierung des Lebensunterhalts Förderung nach dem BAföG Auslandsmobilität der Studierenden Zeitaufwand für Studium und Erwerbstätigkeit Studentische Erwerbstätigkeit Wohnsituation Ernährung und Mensa Gesundheitliche Beeinträchtigung Studium mit Kind Studierende mit Migrationshintergrund Informations- und Beratungsbedarf
Erhebungsdesign	Querschnitt
Erhebungsdatentyp	Quantitative Daten
Grundgesamtheit	Studierende, die im Sommersemester 2016 an einer staatlichen oder staatlich anerkannten Hochschule in Deutschland immatrikuliert waren. Ausgenommen sind Studierende der Verwaltungsfachhochschulen und der Universitäten der Bundeswehr, Gasthörer(innen), beurlaubte, Fern- und deutsche/bildungsinländische Promotionsstudierende sowie Studienkollegiat(inn)en und so genannte Sprachkursstudent(inn)en.
Teilnehmende Hochschulen	248 teilnehmende von 371 eingeladenen Hochschulen (66,8 %)

Stichprobe	Einfach (disproportional) geschichtete Zufallsstichprobe
Erhebungsmethode	standardisierte Online-Befragung
Feldzeit	23. Mai bis 31. August 2016
Nettostichprobe	n = 55.211 (Deutsche/Bildungsinländer(innen)) n = 3.586 (Bildungsausländer(innen))
Rücklaufquote	16,2 % (Deutsche/Bildungsinländer(innen)) 10,2 % (Bildungsausländer(innen))
Datenprodukt und Zugangsweg	CUF: Download SUF: Download, Remote-Desktop, On-Site
Datensatzstruktur	Personendatensätze im wide-Format
DOI	10.21249/DZHW:ssy21:2.0.0
Zitation	Middendorff, E., Apolinarski, B., Becker, K., Bornkessel, P., Brandt, T., Heißenberg, S. & Poskowsky, J. (2019): 21. Sozialerhebung. Datenerhebung 2016. Version: 2.0.0. Hannover: FDZ-DZHW. Datenkuratierung: Baillet, F. & Weber, A. Datensatz: DATENSATZNAME ¹ . doi: 10.21249/DZHW:ssy21:2.0.0.
Anmerkungen	<p>Die 21. Sozialerhebung ist ein Verbundprojekt des Deutschen Studentenwerks (DSW) und des Deutschen Zentrums für Hochschul- und Wissenschaftsforschung (DZHW), gefördert vom Bundesministerium für Bildung und Forschung (BMBF). Der Originaltitel der Studie lautet: „21. Sozialerhebung des Deutschen Studentenwerks durchgeführt vom Deutschen Zentrum für Hochschul- und Wissenschaftsforschung“.</p> <p>Im Rahmen der Studie wurden zwei Teilerhebungen für Deutsche/Bildungsinländer(innen) bzw. Bildungsausländer(innen) durchgeführt, sodass zwei getrennte Datensätze für diese Gruppen bereitgestellt werden.</p>
Weitere Informationen	https://metadata.fdz.dzhw.eu http://www.sozialerhebung.de

¹ An dieser Stelle bitte den genauen Namen der verwendeten Datensatzversion angeben, z. B. ssy21_bi_d_2-0-0 für das Download-SUF des Datensatzes für Deutsche/Bildungsinländer(innen) in der Version 2-0-0 (s. dazu [https://metadata.fdz.dzhw.eu/#/de/data-sets/dat-ssy21-ds1\\$](https://metadata.fdz.dzhw.eu/#/de/data-sets/dat-ssy21-ds1$) für Deutsche/Bildungsinländer(innen) und [https://metadata.fdz.dzhw.eu/#/de/data-sets/dat-ssy21-ds2\\$](https://metadata.fdz.dzhw.eu/#/de/data-sets/dat-ssy21-ds2$) für Bildungsausländer(innen)).

<https://fdz.dzhw.eu>

Projektpublikationen

- Apolinarski, B., & Brandt, T. (2018). *Ausländische Studierende in Deutschland 2016. Ergebnisse der Befragung bildungsausländischer Studierender im Rahmen der 21. Sozialerhebung des Deutschen Studentenwerks durchgeführt vom Deutschen Zentrum für Hochschul- und Wissenschaftsforschung*. Berlin: Bundesministerium für Bildung und Forschung (BMBF).
- Bornkessel, P. (Hrsg.). (2018). *Erfolg im Studium. Konzeptionen, Befunde und Desiderate*. Bielefeld: wbv.
- Bornkessel, P. (2018). Einleitung. In P. Bornkessel (Hrsg.), *Erfolg im Studium. Konzeptionen, Befunde und Desiderate* (S. 7–29). Bielefeld: wbv.
- Bornkessel, P., Heißenberg, S., & Becker, K. (i.E.). Studentische Heterogenität im Spiegel hochschulischer Homogenitätsorientierung – zur Komposition Studierender verschiedener Bildungswege und Studienberechtigungen und deren Implikationen für das Studium. In C. Driesen & A. Ittel (Hrsg.), *Erfolgreichen ankommen – Strategien, Strukturen und Best Practice deutscher Hochschulen für den Übergang Schule–Hochschule*. Münster: Waxmann.
- Dahm, G., Becker, K., & Bornkessel, P. (2018). Determinanten des Studienerfolgs nicht-traditioneller Studierender – zur Bedeutung der sozialen und akademischen Integration, der Lebensumstände und des Studienkontextes für die Studienabbruchneigung beruflich qualifizierter Studierender ohne Abitur. In P. Bornkessel (Hrsg.), *Erfolg im Studium. Konzeptionen, Befunde und Desiderate* (S. 108–174). Bielefeld: wbv.
- Middendorff, E., Apolinarski, B., Becker, K., Bornkessel, P., Brandt, T., Heißenberg, S., & Poskowsky, J. (2017a). *Die wirtschaftliche und soziale Lage der Studierenden in Deutschland 2016. 21. Sozialerhebung des Deutschen Studentenwerks – durchgeführt vom Deutschen Zentrum für Hochschul- und Wissenschaftsforschung*. Berlin: Bundesministerium für Bildung und Forschung (BMBF).
- Middendorff, E., Apolinarski, B., Becker, K., Bornkessel, P., Brandt, T., Heißenberg, S., & Poskowsky, J. (2017b). *Die wirtschaftliche und soziale Lage der Studierenden in Deutschland 2016. Zusammenfassung zur 21. Sozialerhebung des Deutschen Studentenwerks – durchgeführt vom Deutschen Zentrum für Hochschul- und Wissenschaftsforschung*. Berlin: Bundesministerium für Bildung und Forschung (BMBF).
- Middendorff, E., Apolinarski, B., Becker, K., Bornkessel, P., Brandt, T., Heißenberg, S., & Poskowsky, J. (2017c). *The Economic and Social Situation of Students in Germany 2016. Summary of the 21st Social Survey of Deutsches Studentenwerk, conducted by the German Centre for Higher Education Research and Science Studies*. Federal Ministry of Edu-
-

cation and Research (BMBF), Berlin.

Schirmer, H. (2017). *Die wirtschaftliche und soziale Lage der Studierenden in Nordrhein-Westfalen 2016. Regionalauswertung der 21. Sozialerhebung des DSW durchgeführt vom DZHW für die Arbeitsgemeinschaft Studierendenwerke NRW*. Bochum: Arbeitsgemeinschaft Studierendenwerke Nordrhein-Westfalen.

Schirmer, H. (2018). *Die wirtschaftliche und soziale Lage der Studierenden in Potsdam 2016. Regionalauswertung der 21. Sozialerhebung durchgeführt vom DZHW für das Studentenwerk Potsdam*. Potsdam: Studentenwerk Potsdam.

Schirmer, H. (2018). *So leben Studierende in Hamburg. Zur wirtschaftlichen und sozialen Lage der Studierenden in Hamburg 2016. Online-Befragung an Hamburger Hochschulen*. Hamburg: Studierendenwerk Hamburg.

Schirmer, H. (2018). *Die wirtschaftliche und soziale Lage der Studierenden in Hannover 2016. Regionalauswertung der 21. Sozialerhebung des DSW durchgeführt vom DZHW für das Studentenwerk Hannover*. Hannover: Studentenwerk Hannover.

Schirmer, H. (2018). *Die wirtschaftliche und soziale Lage der Studierenden in Berlin 2016. Regionalauswertung der 21. Sozialerhebung durchgeführt vom DZHW für das studierendenWERK BERLIN*. Berlin: studierendenWERK BERLIN.

Schirmer, H. (im Erscheinen). *Die wirtschaftliche und soziale Lage der Studierenden in Köln 2016. Regionalauswertung der 21. Sozialerhebung des DSW durchgeführt vom DZHW für das Kölner Studierendenwerk*. Köln: Kölner Studierendenwerk.

Schirmer, H., & Bröker, L.-M. (im Erscheinen). *Die wirtschaftliche und soziale Lage der Studierenden in Sachsen 2016. Regionalauswertung der 21. Sozialerhebung durchgeführt vom DZHW für die Studentenwerke in Sachsen*. Dresden: Studentenwerk Dresden.

Schirmer, H. (im Erscheinen). *Die wirtschaftliche und soziale Lage der Studierenden in Baden-Württemberg 2016. Regionalauswertung der 21. Sozialerhebung durchgeführt vom DZHW für die Studierendenwerke in Baden-Württemberg*. Karlsruhe: Studierendenwerk Karlsruhe.

Schirmer, H. (im Erscheinen). *Studieren in München. Die wirtschaftliche und soziale Lage der Studierenden in München 2016. Regionalauswertung der 21. Sozialerhebung durchgeführt vom DZHW für das Studentenwerk München*. München: Studentenwerk München.

Weber, A., Daniel, A., Becker, K., & Bornkessel, P. (2018). Proximale Prädiktoren objektiver wie subjektiver Studienerfolgskriterien. In P. Bornkessel (Hrsg.), *Erfolg im Studium. Konzeptionen, Befunde und Desiderate* (S. 59–107). Bielefeld: wbv.

II Datennutzungshinweise

[Voraussetzungen der Datennutzung] Die Daten der 21. Sozialerhebung werden durch das FDZ des DZHW entsprechend der europäischen Datenschutzgrundverordnung (EU-DSGVO) anonymisiert bereitgestellt und ausschließlich zur wissenschaftlichen Nutzung freigegeben.² Das FDZ bietet ein *Scientific Use File* (SUF) für die wissenschaftliche Sekundärnutzung und ein *Campus Use File* (CUF) für Lehr- und Übungszwecke an.

Voraussetzungen für die Nutzung des SUF sind die Anstellung der Datennutzerin/des Datennutzers an einer wissenschaftlichen Einrichtung und der Abschluss eines Datennutzungsvertrags mit dem FDZ. Studierende oder Promovierende ohne eine Anstellung an einer wissenschaftlichen Einrichtung müssen gemeinsam mit einer/einem betreuenden Mitarbeiter(in) einen Datennutzungsvertrag abschließen. Im Zuge des Vertragsabschlusses wird durch das FDZ auch das Vorliegen eines wissenschaftlichen Nutzungsinteresses geprüft. Das Formular für den Datennutzungsantrag kann von der Website des FDZ heruntergeladen werden. Für die Nutzung des CUF ist eine Registrierung beim FDZ nötig. Danach wird das CUF durch das FDZ übermittelt. Ein Datennutzungsvertrag muss nicht abgeschlossen werden.

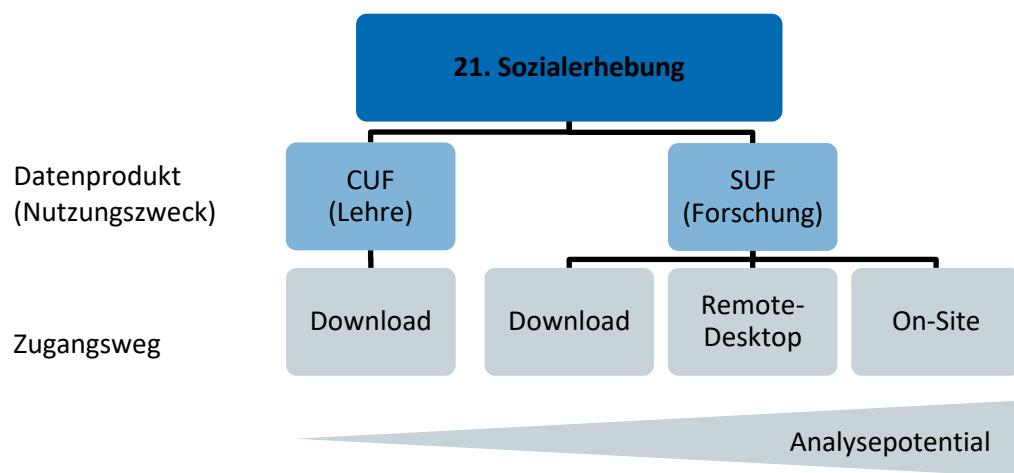
[Datenzugang] Das CUF der 21. Sozialerhebung kann nach Bereitstellung am lokalen Computer genutzt werden. Das SUF wird über drei Zugangswege angeboten, die hinsichtlich des Speicherortes, der Möglichkeit der eigenständigen Verknüpfung mit externen Daten und der Kontrollmöglichkeiten des FDZ unterschiedlich restriktiv sind.

- **Download:** Die Daten werden verschlüsselt auf der Website des FDZ zum Download bereitgestellt. Datennutzer(innen) können die Daten auf ihrem lokalen Computer speichern, falls gewünscht selbst mit Daten aus externen Quellen verknüpfen und die Daten mit eigener Software analysieren.
- **Remote-Desktop:** Die Daten werden auf einem Terminal-Server des FDZ bereitgestellt. Über eine besonders gesicherte Verbindung zwischen dem lokalen Computer der nutzenden Person und dem Terminal-Server des FDZ können die Daten mit der auf dem Terminal-Server vorhandenen Software analysiert werden. Das Übertragen der Daten auf den lokalen Computer ist nicht möglich. Analyseergebnisse werden erst nach einer Prüfung auf datenschutzrechtliche Unbedenklichkeit durch das FDZ freigegeben und zur Verfügung gestellt.
- **On-Site:** Die Daten werden in den Räumlichkeiten des FDZ in einer kontrollierten Umgebung an einem speziell gesicherten Computer zur Analyse bereitgestellt. Wie beim Remote-Desktop-Zugang werden Analyseergebnisse erst nach einer Prüfung auf datenschutzrechtliche Unbedenklichkeit durch das FDZ freigegeben und zur Verfügung gestellt.

² Das Datenschutzkonzept des FDZ ist angelehnt an den Portfolio-Ansatz von Lane, Heus und Mulcahy (2008, S. 6 ff.), an dem sich bereits das Leibniz-Institut für Bildungsverläufe (LIbBi) (Koberg, 2016, S. 699 ff.) und das FDZ der Bundesagentur für Arbeit im Institut für Arbeitsmarkt- und Berufsforschung (Hochfellner, Müller, Schmucker und Roß, 2012, S. 9 f.) orientieren. Das FDZ des DZHW hat diesen Ansatz an die Anforderungen der eigenen Datenbestände angepasst und nutzt vier Kategorien von Maßnahmen zur Sicherstellung des Datenschutzes, die in unterschiedlicher Weise kombiniert werden können: Rechtlich-institutionelle Maßnahmen, informationelle Maßnahmen, technische Maßnahmen und statistische Maßnahmen.

Die bereitgestellten Daten weisen je nach Zugangsweg einen unterschiedlich hohen Informationsgehalt auf und unterscheiden sich damit hinsichtlich ihres Analysepotentials (s. Abbildung 1). Dabei gilt: Je stärker der Datenzugriff der Nutzer(innen) durch technische und organisatorische Maßnahmen kontrolliert wird, desto mehr Informationen können für die Datennutzer(innen) bereitgestellt werden.³ Mit diesem Vorgehen wird ein Höchstmaß an Nutzbarkeit und gleichzeitig ein bestmöglicher Schutz der bereitgestellten Daten sichergestellt.

Abbildung 1: Datenzugangswege und Analysepotential



[Datenprodukte] Über den *Digital Object Identifier* (DOI) 10.21249/DZHW:ssy21:2.0.0 ist eine Website mit zentralen Informationen zur Studie, weiteren Dokumentationsmaterialien sowie einer Übersicht der zur Verfügung stehenden Datenprodukte zur Studie erreichbar.

[Gebühren der Datenbereitstellung] CUF und SUF werden derzeit (Stand: September 2018) kostenfrei zur Verfügung gestellt. Änderungen bzw. die aktuelle Gebührenordnung können auf der Website des FDZ (<https://fdz.dzhw.eu>) eingesehen werden.

[Pflichten der Datennutzer(innen)] Die Datennutzer(innen) sind verpflichtet, folgende Regeln⁴ einzuhalten:

- **Wissenschaftliche Nutzung:** Die Daten dürfen ausschließlich für wissenschaftliche Zwecke verwendet werden. Eine kommerzielle Nutzung ist untersagt.
- **De-Anonymisierungsverbot:** Jeder Versuch der Re-Identifikation von Analyseeinheiten (z. B. Personen, Haushalten, Institutionen) ist verboten.
- **Gebot zur Mitteilung von Sicherheitslücken:** Falls Datennutzer(innen) Kenntnis von Sicherheitslücken hinsichtlich Datenschutz bzw. Datensicherheit erlangen, müssen diese dem FDZ unverzüglich angezeigt werden.
- **Keine Weitergabe der Daten:** SUF dürfen nur durch die Person genutzt werden, die den Datennutzungsvertrag abgeschlossen hat. CUF dürfen ausschließlich im Rahmen der angegebenen Lehrveranstaltung weitergegeben werden.

³ Zu den verschiedenen Anonymisierungsgraden und Analysepotentialen des CUF und der verschiedenen SUF-Varianten s. Kapitel 8.

⁴ Der Datennutzungsvertrag regelt die Nutzungsbedingungen im Detail.

- **Löschungsgebot:** Download-SUF sind nach Ablauf der vereinbarten Nutzungsdauer (in der Regel 1,5 Jahre) von jeglichen Rechnern, Servern und Datenträgern zu löschen. Ebenso müssen alle Sicherungskopien, modifizierten Datensätze (z. B. Arbeits-, Auszugs- oder Hilfsdateien) sowie Ausdrücke vernichtet werden.
- **Bereitstellung/Meldung von Publikationen:** Jede Art von Publikation, die aus der Arbeit mit Daten des FDZ hervorgeht, ist dem FDZ unmittelbar nach Veröffentlichung anzuzeigen. Dabei ist dem FDZ eine elektronische Version der Druckfassung zur Verfügung zu stellen.
- **Zitationspflicht:** Die verwendeten Daten müssen in Veröffentlichungen, anderen Arbeiten (z. B. Abschlussarbeiten) und Vorträgen laut der Vorgaben des FDZ zitiert werden.

1 Inhalt und Anlage der Studie

[Studienreihe] Die Sozialerhebung des Deutschen Studentenwerks (DSW) ist eine seit 1951 bestehende Untersuchungsreihe zur wirtschaftlichen und sozialen Situation der Studierenden in Deutschland.⁵ Damit ist sie im Bereich der Studierendenforschung in Deutschland die am längsten fortgeführte Untersuchungsreihe. Die Sozialerhebung dient – in Ergänzung zur amtlichen Hochschulstatistik – dem nationalen Bildungsmonitoring. Anhand der langen Zeitreihen von Fragestellungen im Kernbestand der alle drei bis vier Jahre durchgeführten Studie ergibt sich die einmalige Möglichkeit die wirtschaftliche und soziale Lage von Studierenden in Deutschland in ihrer zeitlichen Entwicklung abzubilden (Daniel, Sarcletti & Vietgen, 2017, S. 8 f). Mit der 21. Sozialerhebung wird diese Zeitreihe fortgesetzt.

Die Sozialerhebung ist in mehrfacher Hinsicht eine besondere Studie im Feld der Studierendenforschung. Sie erlaubt in einmaliger Weise, differenzierte Aussagen über Studierende zu treffen und bedient eine Fülle an verschiedenen Merkmalen, wie z. B. Abschlussarten, Hochschulzugangsvoraussetzungen und Hochschularten. Lediglich Studierende an Verwaltungsfachhochschulen, Hochschulen des Fernstudiums und Universitäten der Bundeswehr werden aufgrund ihrer besonderen Studiensituation aktuell nicht einbezogen. Seit der 21. Sozialerhebung werden auch Studierende, die in einem Fernstudiengang einer Hochschule der Grundgesamtheit immatrikuliert sind, sowie deutsche und bildungsinländische Studierende im Promotionsstudium von der Studie ausgenommen (Middendorff et al., 2017a, S. 9).⁶

Während in den ersten sechs Sozialerhebungen Vollerhebungen durchgeführt wurden, erfolgt seit der 7. Sozialerhebung eine Stichprobenziehung. Auch die Definition der Grundgesamtheit veränderte sich über die Laufzeit der Studie. War die Studie bis zur 9. Sozialerhebung (1979) auf deutsche Studierende und Bildungsinländer(innen)⁷ begrenzt, wurden ab der 10. Sozialerhebung (1982) auch Bildungsausländer(innen)⁸ einbezogen und ab der 17. Sozialerhebung (2003) als eigene Stichprobe mit einem spezifischen deutsch/englischen Fragebogen erfasst (Middendorff, 2016, S. 5).

Hinsichtlich der Erhebungsmethode konnte in der 21. Sozialerhebung die Umstellung von einer Paper-Pencil-Befragung auf einen Online-Survey abgeschlossen werden. Ausschlaggebend hierfür waren die Vorteile der Online-Befragung, die sich in den Methodentests an jeweils einer Teil-Stichprobe der 18., 19. und 20. Sozialerhebung zeigten:

- Bedeutsame Erhöhung des Stichprobenumfangs
- Individuelle Gestaltung von Befragungspfaden
- Neuaufnahme bzw. Modifikation von Fragen und Themen
- Spezifische Befragung kleinerer Studierendengruppen (Middendorff et al., 2016, S. 1).

Angesichts dieser Vorteile wurde die 21. Sozialerhebung 2016 ausschließlich als Online-Befragung durchgeführt.

⁵ Für einen detaillierten Überblick über die Geschichte der Sozialerhebung s. Middendorff (2016).

⁶ Weitere Informationen hierzu können dem Glossar entnommen werden:

http://www.sozialerhebung.de/download/21/Soz21_glossar.pdf

⁷ Bildungsinländer(innen) sind Studierende mit ausländischer Staatsangehörigkeit, die ihre Hochschulzugangsberechtigung im deutschen Bildungssystem erworben haben.

⁸ Bildungsausländer(innen) sind Studierende mit ausländischer Staatsangehörigkeit, die ihre Hochschulzugangsberechtigung nicht im deutschen Bildungssystem erworben haben.

[Themenfelder der 21. Sozialerhebung] Die Themenfelder der 21. Sozialerhebung umfassen:

- Studienmerkmale und Hochschulzugang
- Soziale Zusammensetzung der Studierenden
- Finanzierung des Lebensunterhalts
- Förderung nach dem BAföG
- Auslandsmobilität der Studierenden
- Zeitaufwand für Studium und Erwerbstätigkeit
- Studentische Erwerbstätigkeit
- Wohnsituation
- Ernährung und Mensa
- Gesundheitliche Beeinträchtigung
- Studium mit Kind
- Studierende mit Migrationshintergrund
- Informations- und Beratungsbedarf (Middendorff et al. 2017a, S. 8)

Dabei sind Themen des Kernbereichs, die als traditioneller Bestandteil des Fragenkatalogs Zeitreihenanalysen ermöglichen, teils unverändert, teils modifiziert eingeflossen, und einzelne Themen früherer Sozialerhebungen wurden erneut in den Fragenkatalog aufgenommen. Darüber hinaus erfolgte eine Ergänzung des Fragenkatalogs durch neue Themenfelder (Middendorff et al., 2016, S. 12). Die Aufnahme neuer Themenfelder unterstützt die Absicht, mit der 21. Sozialerhebung die Vielfalt der Studierenden in Deutschland möglichst präzise abzubilden und einzelne Studierendengruppen noch differenzierter zu untersuchen. Dabei steht im Vordergrund, die Bedingungsfaktoren der Merkmale der sozialen und wirtschaftlichen Lage von Studierenden exakt zu beschreiben um deren Zusammenhang analysieren zu können (Middendorff et al., 2016, S. 12).

Bezüglich des sozialen Hintergrunds der Studierenden wurden die bisherigen Fragen zum Thema Bildungshintergrund ergänzt und modifiziert, ebenso wie der Themenbereich Migrationshintergrund. In Ergänzung zur amtlichen Statistik, können damit die Daten der Sozialerhebung dazu beitragen, diese Studierendengruppe besser ausdifferenzieren zu können. Die neu aufgenommenen Fragen zu sozialen Struktur- und Prozessmerkmalen erfassen Aktivitäten, die die lebensweltliche Orientierung kennzeichnen (z. B. Freizeitverhalten), und die ggf. hinsichtlich ihrer Einflussnahme auf Disparitäten im Hochschulsystem relevant sind (Middendorff et al., 2016, S. 13 f.).

Erstmals in die Sozialerhebung aufgenommen wurden Fragen zum Themenfeld Studienperformanz und zu psychologischen Aspekten. Studienperformanz wird dabei nicht nur auf die Aspekte Studienleistungen und Studienabschluss beschränkt, sondern auch die persönliche Zufriedenheit mit den erbrachten Studienleistungen, der Grad der sozialen Integration und eine eventuell gegebene Studienabbruchintensio(n) wird in Ansatz gebracht, um aus dem Zusammenhang mit anderen Merkmalen die Struktur und den Zusammenhang verschiedener Bedingungsfaktoren für den Studienerfolg ableiten zu können (Middendorff et al., 2016, S. 18).

Durch die Aufnahme psychologischer Merkmale wird die Anschlussfähigkeit der Sozialerhebung an andere sozial- und erziehungswissenschaftliche Studien deutlich verbessert. Darüber hinaus werden dadurch Untersuchungen möglich, inwieweit intra-individuelle Persönlichkeitsmerkmale mit der wirtschaftlichen und sozialen Lage von Studierenden, ihren

spezifischen Merkmalen des Studienerfolgs und -verlaufs oder ihrem Informations- und Beratungsbedarf korrelieren (Middendorff et al., 2016, S. 14).

Schließlich wird das Themenfeld Gesundheit erweitert um Fragen nach dem allgemeinen Gesundheitszustand. Die subjektive Einschätzung dieser Merkmale kann im Zusammenspiel mit anderen Merkmalen eine Grundlage und Hinweise für die Umsetzung der UN-Behindertenrechtskonvention an den Hochschulen liefern (Middendorff et al., 2016, S. 15 f.).

[Analysepotential] Die Fragen des Kernbestandes der Sozialerhebung erlauben sowohl Analysen im Querschnitt als auch Zeitreihenanalysen zu den Themenbereichen Hochschulzugang, Studienmerkmale und -verlauf, Auslandsmobilität, sozio-demographische Merkmale, Bildungsherkunft, Migration, Studieren mit Kind, Studienfinanzierung und Lebenshaltungskosten, Erwerbstätigkeit, BAföG, Wohnsituation, Zeitaufwand für Studium und Erwerbstätigkeit und Student Services (Middendorff et al., 2017a, S. 8). Im Bericht der 21. Sozialerhebung⁹ erfolgt die Analyse der Einnahmen und Ausgaben Studierender hinsichtlich der Bezugsgruppe „Fokus-Typ“, die sich vom bis dahin verwendeten Typ „Normalstudierender“ in gewissem Umfang unterscheidet. Damit ist eine unmittelbare Vergleichbarkeit der Ergebnisse hinsichtlich dieser beiden Themen nicht mehr gegeben. Studierende, die der Gruppe „Fokus-Typ“ zugeordnet werden, sind nicht verheiratet, wohnen bzw. wirtschaften alleine (hierzu zählen auch Studierende in Wohngemeinschaften), haben noch keinen ersten Hochschulabschluss erlangt (außer Bachelor-Abschluss bei Master-Studierenden) und studieren in einem Vollzeit-Präsenz-Studium (Middendorff et al., 2017a, S. 39).¹⁰ Aufgrund der neu aufgenommenen Themenfelder sind über den Kernbestand hinaus Auswertungen zu sozialen Struktur- und Prozessmerkmalen, Studienperformanz und psychologischen Aspekte möglich.

[Neue Fokusgruppen] Im Rahmen der Umstellung auf einen Online-Survey konnte der Stichprobenumfang erheblich erhöht und eine differenzierte und zielgruppenspezifische Fragebogenführung umgesetzt werden. Daher wurden in der 21. Sozialerhebung auch kleinere Subgruppen, sogenannte „Fokusgruppen“, mit einer für eine belastbare Analyse ausreichenden Fallzahl erfasst (Middendorff et al., 2016, S. 3). Als Fokusgruppen sind z. B. zu nennen:

Besondere Studienformate:

- Studierende im Teilzeitstudium
- Studierende im dualen Studium
- Studierende im berufsbegleitenden Studium

Besonderheit des Hochschulzugangs:

- Studierende ohne schulische Hochschulzugangsberechtigung

Besondere Rahmenbedingungen oder Besonderer Hintergrund:

- Studierende mit Kind
- (einem bestimmten) Migrationshintergrund (z. B. Bildungsinländer(inn)en)
- Studierende mit gesundheitlicher Beeinträchtigung

⁹ http://www.sozialerhebung.de/archiv/soz_21_haupt

¹⁰ Informationen zur Differenzierung von „Fokus-Typ“ und „Normalstudierender“ s. http://www.sozialerhebung.de/download/21/Soz21_glossar.pdf

Besondere Haushaltstypen:

- Einzelhaushalte („Fokus-Typ“)
- Elternwohner(innen)
- Paarhaushalte
- Alleinerziehende

2 Erhebungsinstrument

Im Rahmen der 21. Sozialerhebung wurde sowohl für deutsche und bildungsinländische (nachfolgend als Instrument 1 bezeichnet) als auch für bildungsausländische Studierende (nachfolgend als Instrument 2 bezeichnet) ein standardisierter Online-Fragebogen in deutscher und für die bildungsausländischen Studierenden zusätzlich in englischer Sprache als Erhebungsinstrument eingesetzt.¹¹ Kapitel 2.1 stellt die zentralen Inhalte der Erhebungsinstrumente vor. Kapitel 2.2 beschreibt die zur Prüfung und Verbesserung der Fragebögen durchgeführten Pretests.

2.1 Inhalt des Erhebungsinstruments

[Charakteristika der Studienreihe] Im Fokus der 21. Sozialerhebung steht, wie bei den übrigen Erhebungen der Studienreihe, die soziale und wirtschaftliche Situation der Studierenden. Die Kernthemen sind dabei Strukturmerkmale des Studiums, die finanzielle Situation der Studierenden, ihre soziale Herkunft sowie demographische Angaben. Diese Kernthemen werden in allen Erhebungen abgefragt.

Die Fragen bezüglich der Strukturmerkmale des Studiums erfassen Informationen zum derzeitigen Studium (z. B. Studienfach, angestrebter Abschluss, Hochschul- und Fachsemester) (Instrument 1: Frage 5-22; Instrument 2: Frage: 4-13), zum Studienverlauf (z. B. bereits erworbene Hochschulabschlüsse, Studienunterbrechungen, Fach- und Hochschulwechsel) (Instrument 1: Fragen 23.1-43) sowie Angaben zum zeitlichen Aufwand für Studium und Erwerbstätigkeit (Instrument 1: Fragen 128-131 und 89-96; Instrument 2: 17-18 und 30-37). Zudem werden studienbezogene Auslandserfahrungen und deren Finanzierung, Hinderungsgründe bezüglich eines (Teil-)Studiums im Ausland sowie Sprachkenntnisse erfasst (Instrument 1: Fragen 132-147; Instrument 2: Frage 16, 19-21). Ebenso werden Fragen zur persönlichen Einschätzung des Stellenwerts des Studiums und zur subjektiv empfundenen zeitlichen Belastung gestellt (Instrument 1: Frage 161 bzw. 131; Instrument 2: Frage 77).

Im Fokus jeder Sozialerhebung steht die finanzielle Situation der Studierenden (Instrument 1: 73-83, Instrument 2: 28-29). Ausgaben für Lebenshaltung werden betragsgenau erfragt. Ebenfalls betragsgenau erfasst werden die Einnahmen aus eigener Erwerbstätigkeit, familiärer/partnerschaftlicher Unterstützung, Stipendien und weiteren Quellen. Detaillierte Angaben zur Nutzung der Ausbildungsförderung durch das Bundesausbildungsförderungsgesetz (BAföG) werden zusätzlich erhoben (Instrument 1: Frage 84-88.3). Angaben zu(r) Erwerbstätigkeit(en) im aktuellen Semester (Instrument 1: Frage 89-96, Instrument 2: Frage 30-37) werden durch Items zur subjektiven Einschätzung der finanziellen Situation und zu Motiven der eigenen Erwerbstätigkeit erweitert.

Darüber hinaus werden verschiedene sozio-demographische und bildungsbiographische Merkmale sowie der familiäre Hintergrund erhoben (Instrument 1: Frage 97-106; Instrument 2: Frage 46-57), insbesondere zur Bildungsherkunft der Eltern. Hinzu kommen Fragen zur Nutzung von Mensen und Cafeterien (Instrument 1: Fragen 162-167; Instrument 2: 38-39).

¹¹ Die Fragebögen sowie Filterführungsdiagramme können im Metadatensuchsystem des FDZ des DZHW (<https://metadata.fdz.dzhw.eu>) heruntergeladen werden.

[Spezifika der 21. Sozialerhebung] Erstmals abgefragt wurden in der 21. Sozialerhebung detaillierte Angaben zur Studienperformanz und zum Studienerfolg (Instrument 1: Fragen 57-61; Instrument 2: Fragen 73-76), wie beispielsweise die aktuelle Gesamtnote sowie die Zufriedenheit mit dieser Note.

Des Weiteren sind im Gegensatz zur 20. Sozialerhebung in der 21. Sozialerhebung Informationen zum Beratungs- und Informationsbedarf der Studierenden wieder abgefragt worden (Instrument 1: Fragen 152-155). Erstmals wurde nicht nur die Bildungsherkunft der Eltern, sondern auch der Großeltern erfragt (Instrument 1: Frage 126).

Tabelle 1: Themenfelder der 21. Sozialerhebung (Deutsche/Bildungsinländer(innen))

<i>Themen des Kernbestands (teilweise modifiziert)</i>	<i>Wiederaufgenommene Themen</i>	<i>Neue Themen</i>
<i>Studium & Studienzugang</i> Vorbildung & Hochschulzugang (T, M) Studienmerkmale & -verlauf (T, M) Auslandserfahrung (T, M)		
<i>Sozialer Hintergrund</i> sozio-demografische Merkmale (T, M) Bildungsherkunft (T, M) migrationsbezogene Angaben (T, M)		Soziale Prozessmerkmale
<i>Finanzielle Situation</i> Einnahmen & Ausgaben (T, M) BAföG/Studienförderung (T) Erwerbstätigkeit (T, M)		
<i>Wohnen</i> Wohnform (T, M)	Mobilität im Studienalltag	Multilokalität
<i>Zeit</i> Studienaufwand (T, M) Erwerbsaufwand (T, M) zeitliche Beanspruchung (T, M)		
<i>Student Services</i> Ernährung & Mensanutzung (T, M)	Informations- & Beratungsbedarf	
<i>Studienperformanz</i>		Studienerfolg/ Leistungs- stand Intention Studienabbruch soziale Integration Hoch- schule
<i>Psychologische Aspekte</i>		Persönlichkeitsmerkmale Stress- & Belastungsemp- finden akademische Passung
<i>Gesundheit</i> gesundheitsbedingte Studienbeein- trächtigungen (T, M)		allgemeiner Gesundheits- zustand

Themenstatus: T = traditionell, M = modifiziert

Darüber hinaus wurden im Rahmen der 21. Sozialerhebung unter anderem (erstmalig) psychologische Merkmale (z. B. Big-Five) (Instrument 1: Frage 156-157) sowie soziale Prozessmerkmale (studentische Freizeitaktivitäten) erfasst (Instrument 1: Frage 168).

Ebenfalls wurde in der 21. Sozialerhebung die bisherige Bildungsbiografie retrospektiv erhoben (Instrument 1: Frage 44-56). Demzufolge sind die jeweiligen Hochschulzugangswege der Studierenden detailliert abbildbar, so dass auch diejenigen, die mittels zweitem Bildungsweg (bspw. qua Abschluss an einem Abendgymnasium) oder über berufliche Qualifikation, auf dem so genannten dritten Bildungsweg (KMK, 2009) an die Hochschulen gelangen, mit den Daten der 21. Sozialerhebung analysiert werden können.

Ergänzend zur Staatsangehörigkeit der Studierenden (seit der 14. Sozialerhebung) und der Staatsangehörigkeit der Eltern der Studierenden, die seit der 19. Sozialerhebung vorliegt, wurde in der 21. Sozialerhebung erfragt, ob die Studierenden und deren Eltern in Deutschland geboren wurden (Instrument 1: Frage 110-112 bzw. Frage 118-120.2). Dadurch ist in der 21. Sozialerhebung eine genauere Erfassung des Migrationshintergrundes als bei früheren Sozialerhebungen möglich.

Diese Änderungen im Erhebungsinstrument dienen nicht zuletzt auch der Beschreibung bestehender Disparitäten zwischen einzelnen Studierenden- bzw. Fokusgruppen und damit ggf. der Herleitung möglicher Praxisimplikationen für deren soziale und wirtschaftliche Lage. Die Novellierung und Ausweitung der Themenfelder hat zum Ziel, die Bedingungsfaktoren, Zusammenhänge und Einflusshierarchien zwischen den Merkmalen der sozialen und wirtschaftlichen Lage der Studierenden präziser zu beschreiben und besser zu verstehen (s. Tabelle 1).

2.2 Pretest

[Ziel und Verfahren] Aufgrund der Umstellung der 21. Sozialerhebung von einer standardisierten postalischen Befragung auf eine standardisierte Online-Befragung war die empirische Testung der an den Online-Erhebungsmodus angepassten sowie überarbeiteten und neu konstruierten Fragen umsetzungsorientierter und feldvorbereitender Bestandteil der Konzeptionsphase des Projekts.

[Probanden] Insgesamt wurden 3.012 deutsche und bildungsinländische Studierende in drei Pretestgruppen mittels DZHW-HISBUS-Online-Studierendenpanel¹² zwischen dem 25.11.2015 und 31.01.2016 befragt. Darüber hinaus haben 31 internationale Studierende Anfang März 2016 das Instrument für bildungsausländische Studierende getestet.

[Durchführung] Das Erhebungsinstrument wurde im Vorfeld der Erhebung durch zwei zielgruppenorientierte Settings getestet:

- Instrument 1: Für die Fragen an die deutschen und bildungsinländischen Studierenden wurde eine Testbefragung des HISBUS-Panels, dem Online-Access-Studierendenpanel des Deutschen Zentrums für Hochschul- und Wissenschaftsforschung, durchgeführt.
- Instrument 2: Fragen, die sich an die Bildungsausländer(innen) richten, wurden mittels kognitiver Pretests durchgeführt.

¹² DZHW-HISBUS-Online-Studierendenpanel ist ein Online-Access-Panel, also eine aus den Studien des Deutschen Zentrums für Hochschul- und Wissenschaftsforschung (DZHW) auf freiwilliger Basis teilnehmende Gruppe von Studierenden. Dieses Instrument ist zum Zweck der Politikberatung und Studierendenforschung mit Mitteln des Bundesministeriums für Bildung und Forschung (BMBF) und unter methodischer Kontrolle des ZUMA Mannheims etabliert worden. Es besteht seit 2001. Das HISBUS-Panel wird mehrmals im Jahr zu verschiedenen Themen befragt.

Im Rahmen der Testung mittels Instrument 1 kamen drei Verfahren zum Einsatz:

- (1) Das Random-Probe-Verfahren: Hierbei werden Zusatz- und Verständnisfragen zu ausgewählten Items geschlossener Fragen oder ausgewählten Begriffen gestellt, um zu erkunden, wie die jeweilige Frage/der jeweilige Begriff von den Tester(inne)n verstanden bzw. beantwortet wird.
- (2) Das Open-Comment-Verfahren: Hierbei wird den Befragten anhand offener Kommentarfelder die Möglichkeit eingeräumt, zu jeder einzelnen Frage individuelle Anmerkungen zu machen.
- (3) Die Split-Ballot-Methode: Hierbei werden verschiedene Varianten einer Frage eingesetzt und z. B. in Bezug auf ihre Antwortverteilungen miteinander verglichen.

Ferner wurden die getesteten Fragen nach minimal oder nicht besetzten Antwortoptionen, extremen Antwortverteilungen sowie nach der Häufigkeit von Residual- (z. B. „Sonstiges“) und Ausweichkategorien (z. B. „weiß nicht“) analysiert. Überprüft wurden damit die technische Umsetzungsfähigkeit, die Verständlichkeit und Beantwortbarkeit alter und neuer Fragetypen im Online-Design sowie medienspezifischer Erhebungsformen (z. B. Tableau).

Für den kognitiven Pretest, der sich an die Bildungsausländer(innen) richtete, wurde ein anderes Setting gewählt: Hierbei beantworteten internationale Studierende den Online-Fragebogen am PC in einem Testlabor. Dabei waren sie angehalten, Feedback sowohl durch ein Kommentierungsfeld auf jeder Fragebogenseite direkt anzubringen, als auch im anschließenden Gruppengespräch zu äußern. Die Gruppengespräche fanden in kleinen Gruppen von fünf bis sechs Teilnehmer(inne)n statt und wurden anhand eines Interviewleitfadens und interaktiven Diskussionen durchgeführt. Diskutiert wurden unter anderem ihre Erfahrungen mit dem Survey, ihr Gesamteindruck (Länge/Beantwortungsdauer, Verständlichkeit, Interessantheit etc.), ihr Verständnis von einzelnen Fragen, ihr Vorgehen bei der Beantwortung bestimmter Fragen sowie Gründe für Antwortschwierigkeiten oder -verweigerungen.

Auf Grundlage der Pretestergebnisse wurden die Formulierungen verschiedener Frage- bzw. Itemtexte präzisiert, die Reihenfolgen von einzelnen Fragen bzw. Items und Antwortkategorien überarbeitet, einzelne Fragen bzw. Items gestrichen oder neu aufgenommen sowie das Layout entsprechend angepasst.

3 Grundgesamtheit und Stichprobenverfahren

[Grundgesamtheit] Die Grundgesamtheit der 21. Sozialerhebung bilden alle Studierende, die im Sommersemester 2016 an einer staatlichen bzw. staatlich anerkannten, auch kirchlichen Hochschule in Deutschland immatrikuliert waren. Im Rahmen der 21. Sozialerhebung wurden 371 staatliche und staatlich-erkannte Hochschulen gebeten, die Befragung zu unterstützen. 248 Hochschulen (66,8%) sind dieser Bitte gefolgt.

Ausgenommen sind, wie auch in vorherigen Sozialerhebungen, die Universitäten der Bundeswehr, Verwaltungsfach- und Fernhochschulen sowie Fern-, gasthörende und beurlaubte Studierende, sowie Studienkollegiat(inn)en und so genannte Sprachkursstudent(inn)en.¹³ Zudem wurde im Rahmen der 21. Sozialerhebung erstmals auf die Einbeziehung deutscher und bildungsinländischer Promotionsstudierender verzichtet. Schließlich wären, um die Studien- und Lebenssituation der verschiedenen nicht einbezogenen Studierendengruppen adäquat erfassen und abbilden zu können, spezifisch auf sie zugeschnittene Erhebungsinstrumente notwendig, wie sie in speziellen Untersuchungsreihen gegeben sind (bildungsausländische Promotionsstudierende hingegen verbleiben Teil der Grundgesamtheit, da sie mit zwölf Prozent aller bildungsausländischen Studierenden in Deutschland eine relevante Gruppe bilden (Apolinarski & Brandt, 2018, S. 23), die nicht anderweitig erfasst wird (Middendorff et al., 2016, S. 25).

Im Rahmen der Stichprobenziehung wird auf aktuelle (semestergenaue) Studierendenlisten der teilnehmenden Hochschulen zurückgegriffen. Die Studierendenzahlen nicht-teilnehmender, aber gemäß der definierten Grundgesamtheit zu berücksichtigenden Hochschulen werden aus der amtlichen Statistik ermittelt. Entsprechend dieser Vorgaben weist die amtliche Statistik an den 371 deutschen Hochschulen rund 2,5 Millionen Studierende aus, welche sich aus Deutschen, Bildungsinländer(inne)n und Bildungsausländer(inne)n zusammensetzen. An den 248 teilnehmenden Hochschulen (Auswahlrahmen) sind mit etwa 2.277.000 Studierenden 92,7 Prozent der Grundgesamtheit repräsentiert.

[Stichprobe] Die Daten der 21. Sozialerhebung basieren auf einer disproportional geschichteten Zufallsstichprobe. Bei einer Ziehungsquote von 16,67 Prozent der ca. 2.277.000 Studierenden an den 248 teilnehmenden Hochschulen wurden ca. 388.000 Studierende, darunter ca. 353.000 deutsche bzw. bildungsinländische sowie ca. 35.000 bildungsausländische Studierende, zur Teilnahme an der 21. Sozialerhebung eingeladen. Acht Hochschulen, zumeist Fachhochschulen, wünschten sich eine Vollerhebung, um weniger Aufwand mit der Stichprobenziehung zu haben bzw. um eine höhere Fallzahl (>300) zu erzielen, die eine hochschulbezogene Grundauszählung gestattet.

Die Hochschulen, als Adressmittlungsstelle zwischen DZHW und Studierenden, zogen zu einem bestimmten Zeitpunkt die Stichprobe aus ihren Studierenden, an die sie dann unter Anleitung des DZHW die Einladungs- und Erinnerungsschreiben des DZHW mit den personalisierten Zugangslinks versandten. Durch dieses Vorgehen wird die Datenqualität der Stichprobe gesichert, da ausschließlich Studierende einbezogen werden und Mehrfachteilnahmen

¹³ Nach dieser Definition zählen deutsche und bildungsinländische Studierende, die sich studienbezogen im Ausland befinden (SiA), nicht zur Grundgesamtheit. Aus diesem Grund wird die Gruppe der SiA zwar in die Befragung einbezogen, zu Beginn der Befragung jedoch mithilfe des Erhebungsinstruments identifiziert und in einen eigenen für diese spezifische Gruppe passenden Befragungsstrang weitergeleitet. Die Daten dieses Befragungsstrangs sind allerdings nicht im SUF/CUF enthalten.

durch die personalisierten Zugangslinks (nahezu) ausgeschlossen werden können. Demzufolge setzt die Durchführung der Sozialerhebung die grundsätzliche Beteiligung und Mithilfe der Hochschulen voraus, da diese die Studierenden zur Befragung einladen. Dieses Verfahren impliziert, dass Studierende an Hochschulen, die ihre Unterstützung an der Befragung verwehren, aus der jeweiligen Ziehungseinheit vollständig ausfallen (undercoverage). Um dieses Risiko zu minimieren, wurde der hochschulseitige Unterstützungsaufwand im Rahmen der im Sommersemester 2016 zuletzt durchgeführten Onlinebefragung verringert, indem keine separaten Ziehungsquoten für bildungsinländische und -ausländische Studierende zum Einsatz kamen (vorher: Oversampling der bildungsausländischen Studierenden), sondern generell ein Sechstel der Studierenden der Grundgesamtheit eingeladen wurde. Außerdem wurde ein seitens der Hochschulen vorab zu leistender Ausschluss bestimmter Studierendengruppen aus der zu ziehenden Stichprobe vermieden, indem z. B. Fern- und Promotionsstudierende per Selbstauskunft zu Beginn der Befragung identifiziert und exkludiert wurden.

Durch die standardisierte Online-Befragung wird die Realisierung eines großen Stichprobenumfangs technisch ermöglicht. Damit kann einerseits der Diversität unter Studierenden Rechnung getragen werden, und auch kleinere Studierendengruppen können in der Auswertung adäquat berücksichtigt werden. Andererseits können die Daten der Auswertung kleiner regionaler Einheiten (Länder, Regionen, Studierendenwerke, Hochschulen, Hochschulstandorte) dienen (Middendorff et al., 2016, S. 26).

Weiterhin wird nur die Menge von Daten erhoben, die für eine belastbare, bildungspolitisch relevante Analyse benötigt werden. Eine Erhebung darüber hinaus ist nicht vertretbar, da Studierende durch die Vielzahl verschiedener (wissenschaftlicher) Befragungen überbeansprucht werden könnten und der erhöhte Aufwand bei der Datenerhebung und Datenverarbeitung unnötig öffentlich geförderte personelle und finanzielle Ressourcen verbraucht. Auch der zeitliche Rahmen des Projektes würde dadurch erheblich ausgedehnt (Middendorff et al., 2016, S. 26).

4 Durchführung der Erhebung

[Kontaktaufnahme] Die Bereitschaft der Hochschulen sich an der 21. Sozialerhebung zu beteiligen, war von entscheidender Bedeutung für den Erfolg der Feldphase und damit der gesamten Befragung, da das DZHW keine Kontaktdaten der Studierenden hat und somit die Hochschulen als Schnittstelle im Adressmittlungsverfahren zwischen den zu kontaktierenden Studierenden und dem DZHW fungieren. In den letzten Jahren hat sich die Gewinnung der Hochschulen als zunehmend komplex, voraussetzungs- und sehr langwierig herausgestellt. Die Gründe liegen einerseits in der gestiegenen Anzahl an hochschulinternen und -externen Befragungen und dem damit in Verbindung stehenden Verwaltungsaufwand der Hochschulen. Andererseits ist das gestiegene Eigenverwertungsinteresse (z. B. für Qualitätsmanagement) der Hochschulen an hochschulexternen Studierendenbefragungen ursächlich für diese Entwicklung. Aus diesem Grund ist eine frühzeitige Kontaktaufnahme zu den Hochschulen erforderlich. Nur so steht ausreichend Zeit zur Verfügung, um die technischen Voraussetzungen und organisatorischen Bedingungen für die Hochschulteilnahme festzustellen, um zu erkunden, welche Informations- und Abstimmungsbedarfe die einzelnen Hochschulen haben, um diesen Bedarfen zu entsprechen. Die Hochschulen ihrerseits benötigen genug zeitlichen Vorlauf, um ihre Unterstützung der Sozialerhebung (u. a. Stichprobenziehung, Versand der Einladungs- und Erinnerungsmails) einzuplanen und zu organisieren. Die Gewinnung der Hochschulen und somit auch der Studierenden ist somit ein mehrstufiger Prozess.

Alle Hochschulen mit Studierenden der Grundgesamtheit wurden per E-Mail vom DZHW kontaktiert und um Teilnahme an der 21. Sozialerhebung gebeten. Im Anschreiben wurde auf die Relevanz der Studie und ihre Geschichte aufmerksam gemacht. Zudem teilte das DZHW den Hochschulen die Kriterien mit, anhand derer sie die Zielpersonen für die 21. Sozialerhebung identifizieren sollten.¹⁴ 248 der 371 angeschriebenen Hochschulen (66,8 Prozent) erklärten sich zur Teilnahme bereit.¹⁵ Da die Hochschulen die Kontaktdaten ihrer Studierenden aus Datenschutzgründen nicht herausgeben dürfen, wurden im Rahmen der 21. Sozialerhebung die Einladungs- und Erinnerungs-E-Mails mit dem individualisierten Zugangslink durch die Hochschulen an die Studierenden versendet. Das DZHW ermittelte dafür jeweils vorab die Anzahl der benötigten Zugangslinks zur Online-Befragung¹⁶ und sendete diese an die teilnahmebereiten Hochschulen. Diese wiederum zogen aus ihrem Studierendenverzeichnis die E-Mail-Adressen der Studierenden und kontaktierten sie in der Folge mittels den vom DZHW bereitgestellten Einladungs- und Erinnerungsschreiben (inkl. Zugangslinks) per E-Mail.

[Erhebungsunterlagen] Die Befragung wurde mittels der DZHW-Onlinebefragungssoftware „Zofar“ programmiert. Die Erhebungsunterlagen bestanden aus einem Anschreiben mit einem Link zur Online-Befragung sowie einem individuellen Passwort (Token). Die (maximal) drei Erinnerungsschreiben wurden gezielt nur an diejenigen verschickt, die sich noch nicht an der Online-Erhebung beteiligt hatten.

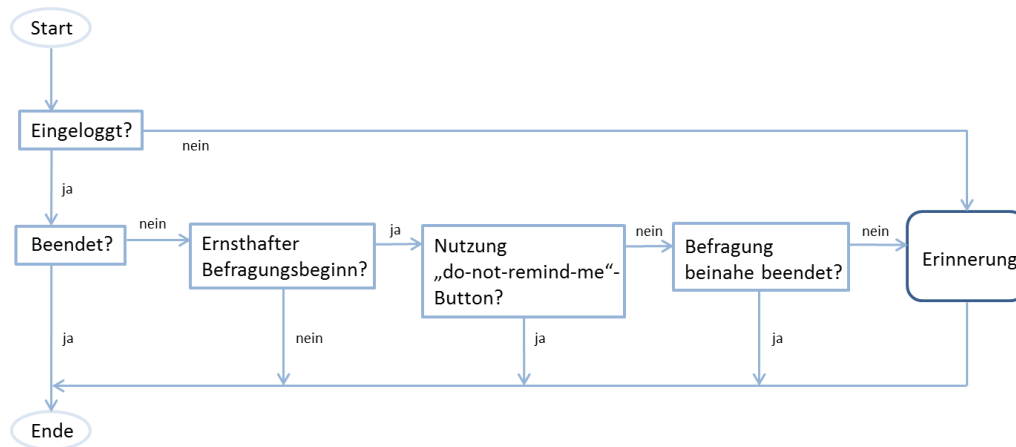
¹⁴ In der Regel ein Sechstel der im Sommersemester immatrikulierten (deutschen/bildungsinländischen und bildungsausländischen) Studierenden der Grundgesamtheit.

¹⁵ Kleinere, Kunst- und Musik- sowie private Hochschulen gaben überproportional keine positive Rückmeldung. Bezogen auf die Studierendenzahl sind dennoch knapp 93 Prozent an den 248 (von 371) Hochschulen, die sich zur Teilnahme bereit erklärt haben, immatrikuliert.

¹⁶ Dazu wurden die zu diesem Zeitpunkt verfügbaren Zahlen der amtlichen Statistik zu Studierenden an den Hochschulen des vorangegangenen Semesters (Wintersemester 2015/16) mit der jeweiligen Auswahlwahrscheinlichkeit kombiniert.

[Feldphase] Der Erhebungszeitraum erstreckte sich vom 23. Mai 2016 bis 31. August 2016. Aufgrund des angewendeten Kontaktverfahrens über die Hochschulen (Adressmittlungsverfahren) konnte das DZHW keinen direkten Einfluss auf den genauen Versandzeitpunkt der Einladungs- und Erinnerungs-E-Mails nehmen (s. Abbildung 2).

Abbildung 2: Erinnerungskonzept der 21. Sozialerhebung (2016)



[Rücklaufsteigernde Maßnahmen] Neben den Erinnerungs-E-Mails wurde die 21. Sozialerhebung durch Presse- und Öffentlichkeitsarbeit seitens des Deutschen Studentenwerks und der einzelnen Studierendenwerke beworben. Es wurde mit Pressemitteilungen und Plakatwerbung auf die bevorstehende Befragung hingewiesen. Darüber hinaus stand die Projektwebsite (<http://www.sozialerhebung.de>) als Informationsquelle zur Verfügung. Sie enthielt die Fragebögen, Fotos und Textbausteine für Pressemitteilungen und weiteres Informationsmaterial zur aktuellen Studie sowie zu vorangegangenen Erhebungen.¹⁷

¹⁷ Die Materialien zur 21. Sozialerhebung sind unter <http://www.sozialerhebung.de/sozialerhebung/archiv> (Stand: 01.09.2018) zu finden.

5 Rücklauf

[Rücklauf insgesamt] Die 248 teilnehmenden Hochschulen beinhalten knapp 93 Prozent der Studierenden der Grundgesamtheit. Bei einer Ziehungsquote von 16,67 Prozent (s. Kapitel 3) beträgt die Bruttostichprobe 376.656 Personen¹⁸ (Tabelle 2).

Tabelle 2: Brutto- und Nettostichproben sowie Rücklaufquoten der 21. Sozialerhebung (2016)

	Deutsche und Bildungsinländische Studierende	Bildungsausländische Studierende
Bruttostichprobe ¹⁸	341.651	35.005
Nettostichprobe	55.211	3.586
Rücklaufquote	16,2 %	10,2 %

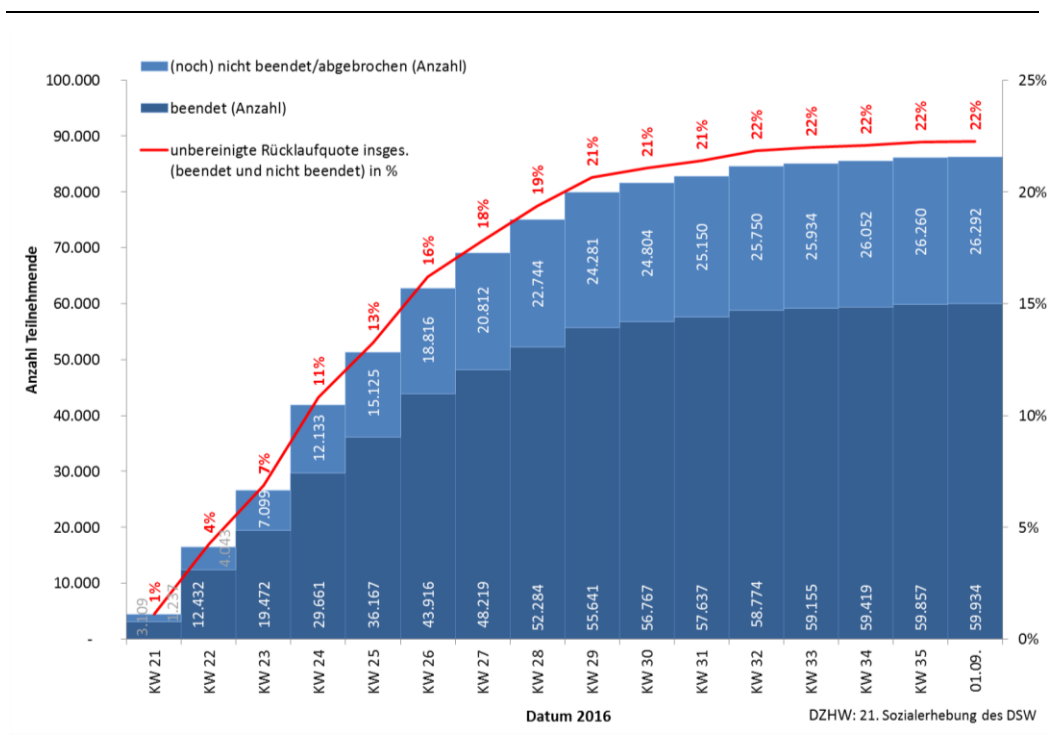
[Rücklauf deutsche und bildungsinländische Studierende] Insgesamt wurden 341.651 deutsche und bildungsinländische Studierende zur 21. Sozialerhebung via E-Mail eingeladen. Davon beteiligten sich 67.007 Personen an der Befragung. Nach umfangreichen Datenprüfungen und -plausibilisierungen (s. Kapitel 6.3) umfasst der Netto-Rücklauf für deutsche und bildungsinländische Studierende 55.211 Fälle, sodass die Rücklaufquote 16,16 Prozent beträgt.

[Rücklauf bildungsausländische Studierende] Im Rahmen der 21. Sozialerhebung wurden 35.005 bildungsausländische Studierende zur Befragung eingeladen. Insgesamt haben sich in der Folge 4.204 bildungsausländische Studierende beteiligt. Nachdem aufwändige Datenprüfungen und -plausibilisierungen durchgeführt wurden (s. Kapitel 6.3), umfasst die Nettostichprobe 3.586 Fälle. Dies entspricht einer Rücklaufquote von 10,24 Prozent.

Abbildung 3 stellt den Rücklauf im Zeitverlauf dar. Wie sich zeigt, hat ein Großteil der Befragten sich während der ersten Hälfte der Feldphase an der Befragung beteiligt, in der auch die Erinnerungs-E-Mails versendet wurden. Insgesamt erstreckte sich die Rücklaufphase über 15 Wochen.

¹⁸ Bei diesen Fällen handelt es sich um eine bereinigte Bruttostichprobe, bei der die nicht zustellbaren E-Mail-Einladungen herausgerechnet wurden. Im Projektbericht wird diese als „Netto-Stichprobe“ bezeichnet (Middendorff et al., 2017a, S. 11).

Abbildung 3: Entwicklung des unbereinigten Rücklaufs der 21. Sozialerhebung (2016) im Zeitverlauf (kumulativ, absolut und in %)



6 Datenaufbereitung

Im Folgenden werden die verschiedenen Schritte der Datenübertragung und -aufbereitung beschrieben. Die in den Kapiteln 6.1 bis 6.3 beschriebenen Tätigkeiten wurden bereits durch das Primärforschungsprojekt durchgeführt. Die in den Kapiteln 6.4 bis 6.6 dargestellten Tätigkeiten wurden teils durch das Primärforschungsprojekt, teils durch das FDZ-DZHW vorgenommen. Die im Rahmen der Datenedition vorgenommenen Aufbereitungsprozesse der Gewichtung und Anonymisierung werden in den beiden folgenden Kapiteln 7 und 8 gesondert erläutert.

6.1 Datenübertragung

Die Angaben aus den Online-Befragungen konnten direkt aus der Befragungssoftware als .csv-Datei exportiert und als Stata-Datei (.dta) weiterverarbeitet werden.

6.2 Codierung offener Angaben

Die (halb-)offenen Angaben wurden nur in codierter Form in das SUF/CUF aufgenommen. Im Rahmen der Datenaufbereitung der Daten erfolgte eine Codierung (eines Teils der) (halb-) offenen Angaben. Diese wurden anhand einer Codierliste numerischen Codes zugeordnet. Je nach Variable wurden unterschiedliche Codierlisten verwendet. Es handelt sich um Klassifikationsschlüssel der amtlichen Statistik (z. B. Schlüsselverzeichnis der Studenten- und Prüfungsstatistik oder amtliche Codes für Staaten) oder um bereits in vorherigen Sozialerhebungen eingesetzte Schlüssel. Für wenige Variablen wurden neue Codierlisten auf Basis der in den Daten der 21. Sozialerhebung vorkommenden Nennungen entwickelt. Für einige halboffene Fragen wurden keine neuen Variablen mit numerischen Codierungen erstellt, sondern die Nennungen nur – sofern möglich – den vorhandenen (geschlossenen) Antwortkategorien zugeordnet. Einzelne offene Fragen wurden nicht vercodet. Dies war der Fall, wenn

- sie für die Auswertungen im Rahmen der 21. Sozialerhebung nicht benötigt wurden.
- lediglich relevant war, ob überhaupt eine Nennung erfolgte.
- die Fragen durch geschlossen erfasste Fragen annähernd ersetzt werden konnten.
- eine Vercodung aufgrund mangelnder personeller und zeitlicher Ressourcen nicht möglich war.

Die durch das Primärforschungsprojekt vorgenommenen Codierungsentscheidungen wurden unverändert beibehalten. In Tabelle 3 sind die codierten Merkmale sowie die jeweils verwendete Codierliste dargestellt. Der Datensatz beinhaltet ausschließlich die codierten numerischen Variablen, die offenen Nennungen selbst sind nicht im Datensatz enthalten. Die Ausprägungen der einzelnen Variablen sind im Datensatzreport dokumentiert.

Tabelle 3: Vercodete Merkmale und verwendete Codierlisten in der 21. Sozialerhebung

Merkmal	Pages im Fragebogen ¹	Codierschema	Codierlisten-ID
Studienfach	CST7; CST 8; CST10; CST12; CSV14 B5a; B5b	Destatis Schlüsselverzeichnis für die Studenten- und Prüfungsstatistik (Stand: WS 2015/2016 und SS 2016)	cl-destatis- studienfach- 2016
Hochschule	CST2; CST3; CSV19; CSV20 B10; B11a	Destatis Schlüsselverzeichnis für die Studenten- und Prüfungsstatistik (Stand: WS 2015/2016 und SS 2016)	cl-destatis- hochschule- 2016
Staatsangehörigkeit/ Land Auslandsauf- enthalt	CMI3; CMI6; CMI11; CMI14; CA2 B47c; B48c	Destatis Staats- und Gebietssystema- tik (Stand: 01.01.2012) (mit Anpas- sungen nach dem Verzeichnis der Staatenamen für den amtlichen Gebrauch in der Bundesrepublik Deutschland; Stand: 11.03.2016)	cl-destatis- ausland-2012
Weitere offene Nennungen	CVH 4: Schultyp zum Erwerb der Studienbe- rechtigung CF1: Weitere Finanze- rungsquellen CE3: andere Tätigkeit nicht-berufsbegleitend Studierende CE4: andere Tätigkeit berufsbegleitend Stu- dierende	Zuordnung zu vorgegebenen Katego- rien oder projekteigene Codierung	-

¹ Die Anfangsbuchstaben der Pagebezeichnung ordnen diese der jeweiligen Befragung zu:
C=Deutsche/Bildungsinländer(innen); B=Bildungsausländer(innen)

6.3 Datenprüfung und Datenbereinigung

[Konsistenz- und Plausibilitätsprüfung] Im Anschluss an die Datenübertragung erfolgte eine umfassende Prüfung der Daten. Die jeweiligen Schritte dieser Prüfung wurden syntaxbasiert (ca. 60.000 Syntax-Zeilen) mittels Stata-Do-Files regelgeleitet durchgeführt. Dabei wurden folgende Typen von Prüfungen vorgenommen:

- **Prüfung Exportdatensatz:** Im Rahmen der Datenerhebung wurden zum Zwecke der Feldkontrolle Testzugänge zur Online-Befragung bereitgestellt, mittels derer das Primärforschungsprojekt einen Einblick in den (de facto-) Zeitpunkt und die Güte der von den Hochschulen versendeten Einladungs- und Erinnerungs-E-Mails erhalten hat. Diese Testzugänge wurden nach Abschluss der Datenerhebung aus dem Datensatz vollständig gelöscht. Ferner wurden etwaige im Zuge der Datenübermittlung entstandene Exportinkonsistenzen oder durch das Befragungssystem inadäquat vergebene (Missing-)Werte im Rohdatensatz recodiert.¹⁹

¹⁹ Dies betrifft beispielsweise die Codierung der Kategorie „bitte auswählen“, welche im Online-Fragebogen über ein Drop-Down-Menü erhoben worden ist: Hier wurde der Exportwert „0“ in den Missing-Wert „-9990“ recodiert. Ferner wurde die Missing-Systematik des Primärforschungsprojekts im Datensatz umgesetzt (s. Kapitel 6.6).

- *Prüfung von Wertebereichen:* Es wurde geprüft, ob die erfasste Ausprägung einer Variablen in dem für diese Variable definierten Wertebereich lag.
- *Prüfung der Einhaltung der Filterführung:* Es wurde auf der einen Seite geprüft, ob aufgrund der im Fragebogen vorgesehenen Filterführung für die jeweilige befragte Person Angaben zu erwarten gewesen wären (Vollständigkeitsprüfung), und auf der anderen Seite, ob für die jeweilige Person keine Angaben hätten erfolgen dürfen (Filterverstöße). Aufgrund der Möglichkeit des „Zurückgehens“ im Fragebogen ist der relevante Pfad für die Prüfung des Fragebogenverlaufs und der Filterführung der letztgespeicherte Pfad.²⁰
- *Konsistenzprüfung:* Es wurde die Konsistenz der Angaben innerhalb des Fragebogens geprüft. Hierbei wurden insbesondere inhaltliche Angaben in den jeweiligen Themenfeldern überprüft.

[Löschung von Fällen] Fälle, die nicht zur Grundgesamtheit gehörten, wurden aus dem Datensatz gelöscht (s. Kapitel 5).²¹ Überdies sind nur so genannte „gültige Fälle“ im bereitgestellten SUF/CUF enthalten. Als gültig gilt ein Fall, der weniger als 50 Prozent Item-Nonresponse auf den zentralen (189) Variablen des Gesamtfragebogens oder weniger als 50 Prozent Item-Nonresponse im ersten Part des Fragebogens (109 Variablen) aufweist und gleichzeitig auf allen fünf Variablen, die für die Gewichtung verwendet werden (s. Kapitel 7), eine gültige Angabe hat.

6.4 Generierung von Variablen

Neben den Variablen, die die codierten Antworten der Befragten enthalten, beinhaltet der Datensatz der 21. Sozialerhebung auch generierte Variablen. Dabei handelt es sich zum einen um Variablen mit numerischen Codierungen von ursprünglich offenen Nennungen (s. Kapitel 6.2). Zum anderen wurden Variablen aus Datenschutzgründen verändert und es wurden im Forschungsfeld häufiger benötigte Variablen aus den Werten einer oder mehrerer Quellvariablen generiert (z. B. Aggregation der Studienfächer zu Studienbereichen und Fächergruppen oder Ableitung von Hochschulart und Hochschulort aus den Hochschulvariablen). Ein großer Teil der generierten Variablen wurde bereits durch das Primärforschungsprojekt erstellt. Generierte Variablen, die aus Gründen der Anonymisierung bereitgestellt werden, wurden bis auf wenige Ausnahmen durch das FDZ erstellt. Der Variablenname einer generierten Variable ist im Datensatz durch das Suffix „_g#“ gekennzeichnet. Eine Übersicht aller für die 21. Sozialerhebung generierten Variablen sowie eine detaillierte Dokumentation der einzelnen Variablen mit Angabe ihrer jeweiligen Generierungsvorschriften können im MetadatenSuchsystem des FDZ des DZHW (<https://metadata.fdz.dzhw.eu>) heruntergeladen werden.

Generierte Variablen wurden im Datensatz – sofern möglich – hinter der jeweiligen Ausgangsvariable positioniert. Wurde eine Variable aus verschiedenen Ausgangsvariablen generiert, wurde sie hinter jene Variable sortiert, welche ihr thematisch am ehesten entspricht.

²⁰ Die Filterführung ist anhand eines Filterführungsdiagramms im MetadatenSuchsystem (<https://metadata.fdz.dzhw.eu>) dokumentiert.

²¹ Dies betrifft hauptsächlich im Fragebogen enthaltene deutsche oder bildungsinländische Promovierende sowie Studierende in Fernstudiengängen.

Falls eine thematische Zuordnung nicht möglich war, wurde die generierte Variable an das Ende des Datensatzes gestellt.

6.5 Datenstruktur und Dateiformat

Der Datensatz enthält die Befragungsdaten sowie die zusätzlich generierten Variablen. Die Reihenfolge der Variablen orientiert sich an der Reihenfolge der zugehörigen Fragen im Fragebogen. Der Datensatz wird sowohl im Stata- als auch im SPSS-Format bereitgestellt.

6.6 Labels, Variablennamen und Missings

[Variablen- und Wertelabelvergabe] Für Variablen- und Wertelabels wurden Formulierungen des Fragebogens übernommen oder prägnante Kurzformen dieser Formulierungen gewählt. Dabei basieren die Variablenlabels in der Regel auf dem entsprechenden Fragetext. Grundlage für die Wertelabels sind je nach Fragetyp die Texte der Antwortoptionen bzw. eine Kombination der Texte von Frage und Antwortoption.

[Variablenbenennung] Mit Ausnahme der Identifikator- und Gewichtungsvariablen werden die Variablennamen nach einem Präfix-Stamm-Identifizier-Suffix-Schema gebildet. Die Variablennamen liefern Metainformationen zur entsprechenden Variable. Dies erleichtert die automatisierte Verarbeitung.

Das Präfix der Variable enthält das Themenfeld. Im Stamm geht das Thema, dem die Variable zugeordnet ist, aus einem zwei- bis sechststelligen Buchstabenkürzel hervor. Tabelle 4 stellt die verschiedenen Themenbereiche der 21. Sozialerhebung sowie das zugehörige Kürzel für den Präfix und den Stamm des Variablennamens dar. Dem Stamm folgt ein eindeutiger Identifizier der Variable.²² Das anhand eines Unterstrichs vom Identifizier abgetrennte Suffix enthält Zusatzinformationen, wie die Kenntlichmachung von generierten Variablen sowie verschiedenen Datenzugangswegen.

Tabelle 4: Themengebiete im Variablennamen der 21. Sozialerhebung (2016)

Präfix ²³	Themenfeld	Stamm (Thema)	
s	Studienmerkmale und -verlauf	hs	(Hochschule)
		abs	(Abschlüsse)
		fa	(Fächer)
		form	(Studienform)
		beru	(berufsbegleitendes Studium)
		dua	(duales Studium)
		art	(Art des Studiums)
		phas	(Studienphase)
		imma	(Immatrikulation)
		sem	(Semester)
		ma	(Master)
		su	(Studienunterbrechung)

²² Die Variablen der Bildungsausländer(innen)(BA)-Befragung sind dadurch identifizierbar, dass dem Identifizier das Buchstabenkürzel „ba“ direkt angefügt wird.

²³ Die Eineindeutigkeit der Variablen wird nicht durch die Vergabe des Präfix, sondern durch den Gesamt-Variablennamen sichergestellt. Ein Buchstabe kann daher als Präfix für mehrere Themenfelder vergeben werden. Zusätzlich gibt es im Datensatz Variablen zum Thema Studienwahlmotive mit dem Stamm „swm“ und zum Thema Hochschulwahlmotive mit dem Stamm „pwm“ ohne Präfix.

		swe	(Studiengangswechsel)
		hw	(Hochschulwechsel)
v	Vorbildung und Hochschulzugang	sb	(Studienberechtigung)
		ausb	(Ausbildung)
s	Studienperformanz	pernot	(Studiennote)
		qual	(Qualifikation Bildungsausländer; BA)
		int	(soziale Integration)
		saa	(Intention Studienabbruch)
w	Wohnen	ohn	(Wohnortmerkmale)
		gr	(Gründe Wohnform))
f	Studienfinanzierung	ein	(bare und unbare Einnahmen)
		su	(Wohnungssuche; BA)
		baf	(BAföG)
		ausg	(Ausgaben)
		sit	(finanzielle Situation)
e	Erwerbstätigkeit	akt	(akt. Sem.)
		tat	(Art der Tätigkeit)
		oft	(Häufigkeit)
		bez	(Studienbezug)
		gr	(Gründe Erwerbstätigkeit)
		ngr	(Gründe Nicht-Erwerbstätigkeit)
		vb	(Vereinbarkeit Studium; BA)
d	Sozio-demographische Merkmale	emo	(Basisangaben)
		kin	(Kinder)
		elt	(Elternangaben)
		gro	(Bildung Großeltern)
		nat	(Staatsangehörigkeit)
		geb	(Geburtsland)
z	Zeit	lv	(Zeit Lehrveranstaltung)
		ses	(Selbststudium)
		erw	(Erwerbstätigkeit)
		aus	(Ausbildung)
		um	(Zusammenfassend)
a	Auslandserfahrungen	inf	(Information zum Auslandsaufenthalt)
		pro	(Programm)
		fin	(Finanzierung Auslandsaufenthalt)
		mot	(Motive)
		plan	(Auslandsabsicht)
		hin	(Hindernisse)
g	Gesundheit	be	(Beeinträchtigung)
		art	(Art der Beeinträchtigung)
		zu	(Gesundheitszustand)
i	Information- und Beratung	bed	(Beratungsbedarf)
		ina	(Inanspruchnahme)
		hin	(Hindernisse)
p	Psychologische Merkmale	big	(Big-Five)
		sw	(Selbstwirksamkeitserwartung)
		as	(Allgemeines Selbstwertgefühl)
		ss	(Stress scale)
		som	(Psychosomatische Beschwerden)
		aka	(Akademische Passung)
m	Mensa und Ernährung	oft	(Häufigkeit der Mensanutzung)
		haupt	(Hauptmahlzeit)
		ev	(Ernährungsverhalten)
		zuf	(Zufriedenheit)
		hin	(Hinderungsgründe für Mensanutzung)
p	Soziale Prozessmerkmale	frei	(Freizeitaktivitäten)
k	Kompetenzen	spr	(Sprache)

b	Aufenthalt in Deutschland	des	(Schwierigkeiten in Deutschland; BA)
		auf	(Aufenthalt Basisangaben; BA)
		de	(Auswahl und Bewertung Deutschland; BA)
		st	(Studieninformation über Deutschland; BA)
		spr	(Spracherwerb; BA)
		hi	(Hilfestellung; BA)

[Codierung fehlender Werte] Zur Codierung fehlender Werte wurde eine einheitliche Missing-Systematik umgesetzt. Fehlende Angaben werden dabei durch negative Werte codiert. Sie lassen sich sechs verschiedenen Gruppen zuordnen. In den ersten beiden Gruppen wird zwischen fehlenden Werten aufgrund von Nicht-Beantwortung von Fragen seitens der Befragten (Nonresponse) und fehlenden Werten aufgrund der Filterführung bzw. für Befragte nicht relevanten Fragen unterschieden. Die dritte Gruppe (selbstberichtete Missing-Werte) beinhaltet Missingcodierungen, die den Befragten innerhalb der Befragung bei spezifischen Fragen als Antwortoptionen zur Verfügung standen. Die vierte bis sechste Gruppe bilden die durch das Primärforschungsprojekt oder das FDZ im Zuge der Datenaufbereitung vergebenen Missing-Werte. Tabelle 5 stellt die verwendete Missingsystematik dar.

Tabelle 5: Systematik fehlender Werte der 21. Sozialerhebung

Fehlende Werte aufgrund Befragungsverhalten

- 9990 „nicht beantwortet“
- 9993 „vorheriger Befragungsabbruch“

Filterbedingter fehlender Wert

- 9991 „nicht gesehen“ aufgrund Filterführung, Fragebogensplits oder Einblendbedingung“

Selbstberichtete fehlende Werte

- 11 „trifft nicht zu“
- 12 „weiß nicht“
- 13 „keine Angabe“

Fehlende Werte bei generierten Variablen

- 9996 „Zusatzvariable nicht bestimmbar“

Fehlende Werte aufgrund Plausibilisierung

- 9981 „Filter-Plausibilisierung“
- 9982 „Inhaltliche Plausibilisierung“

Fehlende Werte aufgrund Anonymisierung

- 967 „anonymisiert“

7 Gewichtung

Die Gewichtung der Daten dient dem Ausgleich von Verzerrungen der Stichprobe aufgrund des Stichprobendesigns (s. Kapitel 3) und im Vergleich zur definierten Grundgesamtheit. Im Folgenden erfolgt zunächst eine allgemeine Einführung in die Vorgehensweise und eine Darstellung der erstellten Gewichte. Im Anschluss wird die Gewichtungsprozedur im Detail beschrieben.

[Ursachen für die Verzerrungen der Stichproben] Maßgeblich für die Verzerrungen von Stichproben sind zwei Prozesse:

- **Designbedingte Verzerrung:** Disproportionalitäten werden bewusst erzeugt, um in bestimmten relevanten Subgruppen die Fallzahlen zu erhöhen (s. Kapitel 3).
- **Verzerrung durch Nonresponse:** Ausfallprozesse (z.B. Nichtteilnahmen, fehlende Erreichbarkeit, Verlust auf dem Postweg) führen zu einem verringerten Rücklauf und somit zu einer Differenz zwischen Brutto- und Nettostichprobe (s. Kapitel 5). Wenn diese Ausfallsprozesse unsystematisch sind (Missing Completely at Random, MCAR), können sie ignoriert werden.²⁴ Jedoch unterliegen sie zumeist einem systematischen Ausfallprozess (Missing at Random, MAR; Missing Not at Random, MNAR), der einer Modellierung bedarf.²⁵

[Konzeptuelles Vorgehen] Im Zuge einer Gewichtungsprozedur sollten idealerweise zunächst designbedingte Disproportionalitäten ausgeglichen werden. Die hierfür benötigten *Designgewichte* ergeben sich bei zufallsgesteuerten Auswahlverfahren direkt aus dem Stichprobendesign. Im Anschluss sollte eine Adjustierung der Designgewichte mit Hilfe von *Nonresponsegewichten* im Quer- und Längsschnitt erfolgen, die auf der Grundlage von Informationen über Teilnehmer(innen) und Nichtteilnehmer(innen) auf Individualebene erzeugt werden. In einem letzten Schritt können die nonresponse-adjustierten Designgewichte anhand von Merkmalsverteilungen aus der Grundgesamtheit kalibriert werden (*Kalibrierung*).

Aufgrund des Stichprobendesigns (s. Kapitel 3) der 21. Sozialerhebung wird in einem ersten Schritt ein Designgewicht gebildet, um die ungleichen Inklusionswahrscheinlichkeiten auszugleichen. Da auf individueller Ebene keine Informationen zu Nichtteilnehmer(inne)n vorliegen, kann keine Nonresponse-Adjustierung des Designgewichts auf Individualebene erfolgen. Das Designgewicht wird somit in einem zweiten Schritt anhand einer Merkmalsverteilung der Grundgesamtheit kalibriert. Da hier Informationen über Teilnehmer(innen) und Nichtteilnehmer(innen) auf aggregierter Ebene vorliegen, erfolgt hier zugleich eine Nonresponse-Adjustierung.

[Gewichtungsverfahren] Innerhalb der Sozialerhebung wird eine Design- und Redressmentgewichtung durchgeführt, wobei im Rahmen der 21. Sozialerhebung folgende Gewichte verwendet werden:

²⁴ Insofern die Einbußen an statistischer Teststärke durch die Verringerung der Stichprobe als irrelevant erachtet werden.

²⁵ Grundlegend zu den unterschiedlichen Formen von Ausfallprozessen s. Rubin (1976).

Tabelle 6: Bereitgestellte Gewichte zur 21. Sozialerhebung (2016)

Variablenname	Beschreibung
wgt01_d	Design-Gewicht Bundesebene
wgt02_d	Design- u. Redressement-Gewicht Bundesebene
wgt03_d	Design-Gewicht Länderebene*
wgt04_d	Design- u. Redressement-Gewicht Länderebene*

* nur im Datensatz der Deutschen und Bildungsinländer(innen) enthalten

Durch die Gewichtung der Einzelfälle werden Abweichungen zwischen realisierter Stichprobe und Grundgesamtheit laut amtlicher Statistik für die Merkmale Geschlecht, Hochschulart (Universität, Fachhochschule), Bundesland der Hochschule, Fächergruppe der amtlichen Statistik und für das Alter (fünf Altersgruppen: bis 19 Jahre, 20-22 Jahre, 23-25 Jahre, 26-30 Jahre, 31 Jahre und älter) korrigiert. Die realisierte Stichprobe ist damit aussagefähig für Studierende im Bundesgebiet sowie grundsätzlich auch auf der Ebene der Länder.

Schritt 1 – Design-Gewichtung

Studierende einiger Hochschulen sind aufgrund des Stichprobendesigns überrepräsentiert. Folglich weisen diese Individuen eine höhere Wahrscheinlichkeit auf, in die Stichprobe zu gelangen. Dies wird durch die Verwendung der inversen Auswahlwahrscheinlichkeit (π_i) korrigiert. Das Design-Gewicht wgt_d_i ist somit:

$$wgt_d_i = \frac{1}{\pi_i}$$

Elemente, die mit höherer Wahrscheinlichkeit als andere in die Stichprobe eingehen, erhalten somit ein niedrigeres Gewicht und *vice versa*.

Schritt 2 – Kalibrierung der Designgewichte

Wie beschrieben, ist eine Nonresponse-Adjustierung der Designgewichte auf Individualebene nicht möglich. Es liegen jedoch folgende demographischen Merkmale aus der Grundgesamtheit²⁶ vor, die zur Kalibrierung der Gewichte herangezogen werden können: Region (Ost/West), Geschlecht, Fächergruppe, Hochschultyp (Universität/Fachhochschule), Alter und Personen mit deutscher Staatsangehörigkeit versus Bildungsinländer(innen).²⁷ Die Kalibrierung erfolgt sowohl auf Bundesebene als auch gesondert auf Länderebene.

Da die Merkmalsträger in der Grundgesamtheit ebenfalls Informationen über Nichtteilnehmer(innen) enthalten, erfolgt durch die Verwendung der Redressement-Gewichte zusätzlich eine Art Nonresponse-Adjustierung im Hinblick auf die verwendeten Merkmale.

²⁶ Alle Informationen, die zur Kalibrierung der Designgewichte verwendet wurden, leiten sich aus Daten des Wintersemesters 2014/2015 des Statistischen Bundesamtes ab, da die aktuelle Statistik zum Zeitpunkt der Gewichtung noch nicht vorlag.

²⁷ Die Gewichtung wurde entlang folgender Ausprägungen durchgeführt: Geschlecht: weiblich versus männlich; Region: Ost (Berlin, Brandenburg, Mecklenburg-Vorpommern, Sachsen, Sachsen-Anhalt, Thüringen) vs. West (Baden-Württemberg, Bayern, Bremen, Hamburg, Hessen, Niedersachsen, Nordrhein-Westfalen, Rheinland-Pfalz, Schleswig-Holstein, Saarland); Hochschultyp: Universität (inklusive Pädagogische Hochschulen, Theologische Hochschulen, Kunst- und Musikhochschulen) versus Fachhochschule; Fächergruppe: entsprechend Schlüsselverzeichnis und der Sozialerhebung (WiSe 2015/16).

Weiterhin werden im Rahmen dieser Redressement-Gewichtung, die mittels Zellgewichtung der Daten an der Amtsstatistik umgesetzt wird, diejenigen Zellen, welche (vorab definierte) Ausreißerwerte in den Gewichten ($<0,2$ und >6) enthalten, mittels Zellverschmelzung nivelliert.

8 Anonymisierung

[Datenschutzrechtlicher Rahmen] Für personenbezogene Daten²⁸, 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 2017.²⁹ 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). Das heißt, für die Weitergabe von Daten aus wissenschaftlichen Forschungsprojekten an Dritte sind die Daten derart zu anonymisieren, dass kein Bezug zur Person mehr hergestellt werden kann.

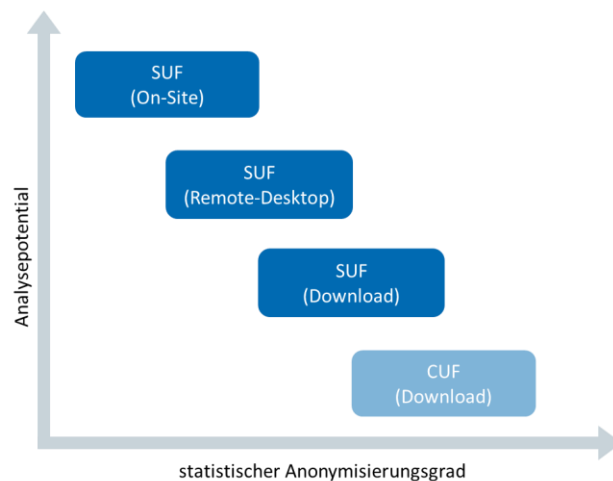
[Datenzugang, Anonymisierungsgrad und Analysepotential] Das FDZ des DZHW stellt für die 21. Sozialerhebung ein SUF für die wissenschaftliche Sekundärnutzung und ein CUF für Lehr- und Übungszwecke zur Verfügung. Die Anonymität der Befragten wird dabei über eine Kombination aus statistischen Maßnahmen und technischen Zugriffsbeschränkungen sichergestellt. Je stärker der Datenzugang technisch kontrolliert wird, desto geringer ist das Risiko einer De-Anonymisierung der Daten, desto weniger müssen die Daten mittels statistischer Maßnahmen um Informationen reduziert werden und desto größer bleibt ihr Analysepotential.

Während das CUF nach einer Registrierung direkt durch das FDZ des DZHW übermittelt wird, wird das SUF über drei verschiedene Zugangswege angeboten: Download, Remote-Desktop und On-Site (für weiterführende Informationen s. Kapitel II). Für jeden Zugangsweg wird eine andere SUF-Variante bereitgestellt, die unterschiedlich stark anonymisiert worden ist und entsprechend weniger oder mehr Informationen umfasst. Abbildung 4 gibt einen Überblick über den jeweiligen Grad der statistischen Anonymisierung und dem damit verbundenen Analysepotential. Im Folgenden werden die durchgeführten statistischen Anonymisierungsmaßnahmen in Abhängigkeit von Datenprodukt (SUF/CUF) und Zugangsweg erläutert.

²⁸ „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 DSGVO, S. 1).

²⁹ 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.

Abbildung 4: Datenzugangsweg, statistischer Anonymisierungsgrad und Analysepotential der Daten der 21. Sozialerhebung (2016)



[Statistische Anonymisierungsmaßnahmen] Im Rahmen der Anonymisierung sind zunächst alle Informationen, mit denen sich Personen oder Institutionen direkt identifizieren lassen, zu löschen. Von diesen sogenannten *direkten Identifikatoren*, wie Namen, Adressen oder E-Mail-Adressen, wurde im Rahmen der 21. Sozialerhebung nur die letztgenannte erfasst, und auch nur dann, wenn die Befragten ihre Bereitschaft erklärten, am HISBUS-Panel bzw. an dem Gewinnspiel teilzunehmen. Die E-Mail-Adressen wurden getrennt von der Hauptbefragung in einer Add-on Befragung erfasst und gespeichert. Sie sind nicht mit dem Datensatz der 21. Sozialerhebung verknüpfbar und somit weder im CUF noch in den verschiedenen SUF-Varianten enthalten. Des Weiteren wurde, um einen Rückbezug auf die Originaldaten zu verhindern, die Original-Identifikationsnummer aus dem Datensatz entfernt und durch eine neue, zufällig vergebene Identifikationsnummer ersetzt.

Anschließend wurden die *Quasi-Identifikatoren* bestimmt, also Informationen, die in Kombination oder durch die Anspielung externer Informationen geeignet sind, eine Person indirekt zu identifizieren.³⁰ Für die 21. Sozialerhebung wurden beispielsweise folgende Quasi-Identifikatoren identifiziert: Name sowie Art und Ort der Hochschule, Studienfach, Abschlussart, Alter und Staatsangehörigkeit. Um eine eindeutige Zuordnung der Daten der 21. Sozialerhebung zu unterbinden, wurden diese Schlüsselmerkmale – je nach Datenprodukt bzw. Zugangsweg – aggregiert oder gelöscht (s. Tabelle 7). Beispielsweise ist das Merkmal „Wohnort bei Erwerb der Studienberechtigung (PLZ)“ im SUF für die On-Site-Nutzung uneingeschränkt verfügbar. Im Remote-Desktop-SUF und im Download-SUF hingegen wird das Merkmal zu NUTS-2-Regionen aggregiert. Im CUF wird die Information gelöscht.

Darüber hinaus empfehlen Ebel und Meyermann, 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

³⁰ Dabei ist darauf hinzuweisen, dass die Identifikation einer Person bereits durch die Stichprobenauswahl erschwert wird, da eine Ungewissheit darüber besteht, ob eine befragte Person eine einzigartige Merkmalskombination in der Population aufweist.

führen könnten“ (Ebel & Meyermann, 2015, S. 5). Die offenen Angaben wurden größtenteils bereits im Rahmen der Datenaufbereitung durch das Primärforschungsprojekt vercodet und werden in dieser Form im CUF sowie in allen SUF-Varianten zur Verfügung gestellt. Teilweise wurden jedoch – in Abhängigkeit von der Sensibilität der enthaltenen Informationen und vom Zugangsweg – die vom Primärforschungsprojekt vorgenommenen Codierungen zusätzlich aggregiert. Nicht codierte offene Angaben (s. Kapitel 6.2) wurden im CUF und in allen SUF-Varianten gelöscht.

Zuletzt wurde geprüft, ob in den Daten *sensible Informationen*, z. B. zur Gesundheit, sexuellen Orientierung oder zu politischen Einstellungen, enthalten waren. Diese eignen sich zwar nicht notwendig zur Re-Identifikation von Individuen oder Institutionen, jedoch können die Informationen im Falle einer De-Anonymisierung nutzbringend sein (Koberg, 2016, S. 694) und sind daher besonders schützenswert (Art. 9 DSGVO, Erwägungsgrund 51 DSGVO). In der 21. Sozialerhebung wurden Gesundheitsinformationen erhoben, für die bei den Befragten kein zusätzliches Einverständnis für die Sekundärnutzung eingeholt wurde. Daher wurden diese Antworten im CUF und allen SUF-Varianten gelöscht.

Zur Realisierung der Anonymität der Befragten im CUF wurde zusätzlich zu den vergleichsweise restriktiven statistischen Anonymisierungsmaßnahmen des Download-SUF eine 50-Prozent-Substichprobe der Download-SUF-Daten gezogen. Die nachfolgende Tabelle 7 stellt in Kurzform die durchgeführten statistischen Anonymisierungsmaßnahmen je nach Datenform bzw. Zugangsweg dar.

Tabelle 7: Maßnahmen der statistischen Anonymisierung der Daten der 21. Sozialerhebung (2016) nach Zugangsweg³¹

Merkmal	On-Site-SUF	Remote-Desktop-SUF	Download-SUF	Download-CUF (Substichprobe)
Direkte Identifikatoren	Löschung / Vergabe einer zufälligen ID	Löschung / Vergabe einer zufälligen ID	Löschung / Vergabe einer zufälligen ID	Löschung / Vergabe einer zufälligen ID
Studienfach	Freigabe	Aggregation: Studienbereiche nach Schlüsselverzeichnis der Studenten- und Prüfungsstatistik WiSe 2015/2016 und SoSe 2016 von Destatis	Aggregation: Fächergruppen nach Schlüsselverzeichnis der Studenten- und Prüfungsstatistik WiSe 2015/2016 und SoSe 2016 von Destatis	Aggregation: Fächergruppen nach Schlüsselverzeichnis der Studenten- und Prüfungsstatistik WiSe 2015/2016 und SoSe 2016 von Destatis
Abschlussart	Freigabe	Freigabe	Aggregation: Zusammenfassung von "kirchliche Prüfung" und "anderer Abschluss"	Aggregation: Zusammenfassung von "kirchliche Prüfung" und "anderer Abschluss"
Hochschule	Aggregation: Hochschulart nach Schlüsselverzeichnis der Studenten- und Prüfungsstatistik WiSe 2015/2016 und SoSe 2016 von Destatis	Aggregation: Hochschulart nach Schlüsselverzeichnis der Studenten- und Prüfungsstatistik WiSe 2015/2016 und SoSe 2016 von Destatis	Aggregation: Hochschulart (Uni/FH)	Aggregation: Hochschulart (Uni/FH)
	Aggregation: Hochschulort (Baden-Württemberg, Bayern, Hessen und NRW einzeln ausgewiesen; ansonsten Zusammenfassung zu fünf Bundesländer-Gruppen	Aggregation: Hochschulort (Baden-Württemberg, Bayern, Hessen und NRW einzeln ausgewiesen; ansonsten Zusammenfassung zu fünf Bundesländer-Gruppen	Aggregation: Hochschulort (alte/neue Bundesländer)	Aggregation: Hochschulort (alte/neue Bundesländer)
Alter (in Jahren)	Freigabe	Freigabe	Bottom-/Topcodierung: Zusammenfassung von < 19; 19-30 einzeln ausgewiesen, Zusammenfassung von > 30	Bottom-/Topcodierung: Zusammenfassung von < 19; 19-30 einzeln ausgewiesen, Zusammenfassung von > 30
Geschlecht	Freigabe	Freigabe	Aggregation: männlich/weiblich/Missing	Aggregation: männlich/weiblich/Missing

³¹ Detaillierte Informationen zu den anonymisierten Variablen sind dem Metadatensuchsystem (<https://metadata.fdz.dzhw.eu>) zu entnehmen.

Staatsangehörigkeit/Geburtsland (Ausland) der Befragten und der Eltern der Befragten	Freigabe	Aggregation: NEPS-Länderklassifikation (https://www.neps-da.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/6-0-0/SC5_6-0-0_Anonymisation.pdf Seite 33-34)	Aggregation: Weltregionen (Kontinente; bei Europa zusätzlich Unterscheidung EU/nicht EU)	Aggregation: Weltregionen (Kontinente; bei Europa zusätzlich Unterscheidung EU/nicht EU)
Ort der Studienberechtigung (Bundesland)*	Freigabe	Freigabe	Freigabe	Aggregation: alte/neue Bundesländer
Wohnort bei Erwerb der Studienberechtigung (PLZ)*	Freigabe	Aggregation: NUTS-2	Aggregation: NUTS-2	Löschung
Erlangung der deutschen Staatsangehörigkeit der Befragten und der Eltern der Befragten*	Freigabe	Freigabe	Löschung	Löschung
Sprache im Elternhaus*	Freigabe	Freigabe	Löschung	Löschung
Stipendienförderung*	Freigabe der Studienstiftung des deutschen Volkes, sonst Löschung	Löschung	Löschung	Löschung
Förderform BAföG*	Freigabe	Freigabe	Löschung	Löschung
Gründe für Studienunterbrechung*	Aggregation: Zusammenfassung der Gründe „akute gesundheitliche Probleme“, „chronische Krankheit/Behinderung“, „Pflege Angehöriger“ und „andere familiäre Gründe“ zu „private Gründe“	Aggregation: Zusammenfassung der Gründe „akute gesundheitliche Probleme“, „chronische Krankheit/Behinderung“, „Pflege Angehöriger“ und „andere familiäre Gründe“ zu „private Gründe“	Aggregation: Zusammenfassung der Gründe „akute gesundheitliche Probleme“, „chronische Krankheit/Behinderung“, „Pflege Angehöriger“ und „andere familiäre Gründe“ zu „private Gründe“	Aggregation: Zusammenfassung der Gründe „akute gesundheitliche Probleme“, „chronische Krankheit/Behinderung“, „Pflege Angehöriger“ und „andere familiäre Gründe“ zu „private Gründe“
gesundheitliche Beeinträchtigung bzw. Studienbeeinträchtigung (ja/nein)*	Freigabe	Freigabe	Löschung	Löschung
Grad der Studienbeeinträchtigung*	Freigabe der Gesamteinschätzung, sonst Löschung	Freigabe der Gesamteinschätzung, sonst Löschung	Löschung	Löschung
Gesundheitszustand*	Freigabe	Freigabe	Löschung	Löschung
Migrationsstatus**	Löschung	Löschung	Löschung	Löschung

(Sonstige) gesundheitliche/private Informationen	Löschung	Löschung	Löschung	Löschung
(Sonstige) offene Angaben	Vercodung / Löschung	Vercodung / Löschung	Vercodung / Löschung	Vercodung / Löschung
* betrifft nur Datensatz der Deutschen/Bildungsinländer(innen)				
** betrifft nur Datensatz der Bildungsausländer(innen)				

9 Literaturverzeichnis

- Apolinarski, B., & Brandt, T. (2018). *Ausländische Studierende in Deutschland 2016. Ergebnisse der Befragung bildungsausländischer Studierender im Rahmen der 21. Sozialerhebung des Deutschen Studentenwerks durchgeführt vom Deutschen Zentrum für Hochschul- und Wissenschaftsforschung*. Berlin: Bundesministerium für Bildung und Forschung (BMBF).
- Bornkessel, P. (Hrsg.). (2018). *Erfolg im Studium. Konzeptionen, Befunde und Desiderate*. Bielefeld: wbv.
- Bornkessel, P. (2018). Einleitung. In P. Bornkessel (Hrsg.), *Erfolg im Studium. Konzeptionen, Befunde und Desiderate* (S. 7–29). Bielefeld: wbv.
- Bornkessel, P., Heißenberg, S., & Becker, K. (i. E.). Studentische Heterogenität im Spiegel hochschulischer Homogenitätsorientierung – zur Komposition Studierender verschiedener Bildungswege und Studienberechtigungen und deren Implikationen für das Studium. In C. Driesen & A. Ittel (Hrsg.), *Erfolg im Studium. Konzeptionen, Befunde und Desiderate* (S. 108–174). Bielefeld: wbv.
- Dahm, G., Becker, K., & Bornkessel, P. (2018). Determinanten des Studienerfolgs nicht-traditioneller Studierender – zur Bedeutung der sozialen und akademischen Integration, der Lebensumstände und des Studienkontextes für die Studienabbruchneigung beruflich qualifizierter Studierender ohne Abitur. In P. Bornkessel (Hrsg.), *Erfolg im Studium. Konzeptionen, Befunde und Desiderate* (S. 108–174). Bielefeld: wbv.
- Daniel, A., Sarclotti, A., & Vietgen, S. (2017). *20. Sozialerhebung. Daten- und Methodenbericht zur Studierendenbefragung 2012*. Version 1.0.0. Hannover: FDZ-DZHW.
- 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. Verfügbar unter: <http://www.forschungsdaten-bildung.de/fdb-informiert>
- Hochfellner, D., Müller, D., Schmucker, A., & Roß, E. (2012). *FDZ-Methodenreport. Datenschutz am Forschungsdatenzentrum* (Nr. 06). Nürnberg: Institut für Arbeitsmarkt- und Berufsforschung (IAB).
- Koberg, T. (2016). 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
- Kultusministerkonferenz (KMK) (2009). Hochschulzugang für beruflich qualifizierte Bewerber ohne schulische Hochschulzugangsberechtigung (Beschluss der Kultusministerkonferenz vom 06.03.2009). Verfügbar unter: https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2009/2009_03_06-Hochschulzugang-beruflich-qualifizierte-Bewerber.pdf
- 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.
- Middendorff, E. (2016). *Die Sozialerhebungen des Deutschen Studentenwerks 1951 – 2016. Ein historischer Überblick über Akteure, Wellen und Themen*. Hannover: DZHW.
- Middendorff, E., Apolinarski, B., Becker, K., Bornkessel, P., Brandt, T., Heißenberg, S., Naumann, H., & Poskowsky, J. (2016). *21. Sozialerhebung – Wissenschaftliche Vorbereitung, Durchführung und Dokumentation der Ergebnisse. Bericht zur Konzeption des Teilprojektes und zur Vorbereitung der Feldphase im Rahmen des Verbundprojektes*. Hannover: DZHW.
- Middendorff, E., Apolinarski, B., Becker, K., Bornkessel, P., Brandt, T., Heißenberg, S., & Poskowsky, J. (2017a). *Die wirtschaftliche und soziale Lage der Studierenden in Deutschland 2016. 21. Sozialerhebung des Deutschen Studentenwerks – durchgeführt vom Deutschen Zentrum für Hochschul- und Wissenschaftsforschung*. Berlin: Bundesministerium für Bildung und Forschung (BMBF).
- Middendorff, E., Apolinarski, B., Becker, K., Bornkessel, P., Brandt, T., Heißenberg, S., & Poskowsky, J. (2017b). *Die wirtschaftliche und soziale Lage der Studierenden in Deutschland 2016. Zusammenfassung zur 21. Sozialerhebung des Deutschen Studentenwerks – durchgeführt vom Deutschen Zentrum für Hochschul- und Wissenschaftsforschung*. Berlin: Bundesministerium für Bildung und Forschung (BMBF).
- Middendorff, E., Apolinarski, B., Becker, K., Bornkessel, P., Brandt, T., Heißenberg, S., & Poskowsky, J. (2017c). *The Economic and Social Situation of Students in Germany 2016. Summary of the 21st Social Survey of Deutsches Studentenwerk, conducted by the German Centre for Higher Education Research and Science Studies*. Federal Ministry of Education and Research (BMBF), Berlin.
- Rubin, D. B. (1976). Inference and missing data. *Biometrika*, 63(2), 581–592.

- Schirmer, H. (2017). *Die wirtschaftliche und soziale Lage der Studierenden in Nordrhein-Westfalen 2016. Regionalauswertung der 21. Sozialerhebung des DSW durchgeführt vom DZHW für die Arbeitsgemeinschaft Studierendenwerke NRW*. Bochum: Arbeitsgemeinschaft Studierendenwerke Nordrhein-Westfalen.
- Schirmer, H. (2018). *Die wirtschaftliche und soziale Lage der Studierenden in Potsdam 2016. Regionalauswertung der 21. Sozialerhebung durchgeführt vom DZHW für das Studentenwerk Potsdam*. Potsdam: Studentenwerk Potsdam.
- Schirmer, H. (2018). *So leben Studierende in Hamburg. Zur wirtschaftlichen und sozialen Lage der Studierenden in Hamburg 2016. Online-Befragung an Hamburger Hochschulen*. Hamburg: Studierendenwerk Hamburg.
- Schirmer, H. (2018). *Die wirtschaftliche und soziale Lage der Studierenden in Hannover 2016. Regionalauswertung der 21. Sozialerhebung des DSW durchgeführt vom DZHW für das Studentenwerk Hannover*. Hannover: Studentenwerk Hannover.
- Schirmer, H. (2018). *Die wirtschaftliche und soziale Lage der Studierenden in Berlin 2016. Regionalauswertung der 21. Sozialerhebung durchgeführt vom DZHW für das StudierendenWERK BERLIN*. Berlin: StudierendenWERK BERLIN.
- Schirmer, H. (i. E.). *Die wirtschaftliche und soziale Lage der Studierenden in Köln 2016. Regionalauswertung der 21. Sozialerhebung des DSW durchgeführt vom DZHW für das Kölner Studierendenwerk*. Köln: Kölner Studierendenwerk.
- Schirmer, H., & Bröker, L.-M. (i. E.). *Die wirtschaftliche und soziale Lage der Studierenden in Sachsen 2016. Regionalauswertung der 21. Sozialerhebung durchgeführt vom DZHW für die Studentenwerke in Sachsen*. Dresden: Studentenwerk Dresden.
- Schirmer, H. (i. E.). *Die wirtschaftliche und soziale Lage der Studierenden in Baden-Württemberg 2016. Regionalauswertung der 21. Sozialerhebung durchgeführt vom DZHW für die Studierendenwerke in Baden-Württemberg*. Karlsruhe: Studierendenwerk Karlsruhe.
- Schirmer, H. (i. E.). *Studieren in München. Die wirtschaftliche und soziale Lage der Studierenden in München 2016. Regionalauswertung der 21. Sozialerhebung durchgeführt vom DZHW für das Studentenwerk München*. München: Studentenwerk München.
- Weber, A., Daniel, A., Becker, K., & Bornkessel, P. (2018). Proximale Prädiktoren objektiver wie subjektiver Studienerfolgskriterien. In P. Bornkessel (Hrsg.), *Erfolg im Studium. Konzeptionen, Befunde und Desiderate* (S. 59–107). Bielefeld: wbv.