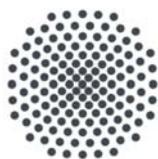


**SISS:**  
**Schriftenreihe des Instituts für Sozialwissenschaften  
der Universität Stuttgart**

**No. 49 / 2020**

***Psychische Gesundheit:  
Eine heikle Thematik in der  
empirischen Umfrageforschung?  
Zur Validierung des Crosswise Mo-  
dells.***

***Anna Götze  
Andreas Wahl***



**Universität Stuttgart**

**Abteilung für Soziologie  
und empirische Sozialforschung**



**SISS:  
Schriftenreihe  
des Instituts für Sozialwissenschaften  
der Universität Stuttgart:  
No. 49 / 2020**

**Psychische Gesundheit: Eine heikle Thematik  
in der empirischen Umfrageforschung? Zur Validierung des  
Crosswise Modells.**

**Anna Götze  
Andreas Wahl**

**ISSN 2199-7780**

**Universität Stuttgart  
Institut für Sozialwissenschaften  
Abt. f. Soziologie u. empirische Sozialforschung  
70174 Stuttgart**

**Anna Götze und Andreas Wahl**

**Psychische Gesundheit: Eine heikle Thematik in der empirischen Umfrageforschung? Zur Validierung des Crosswise Modells.**

SISS – Schriftenreihe des Instituts für Sozialwissenschaften der Universität Stuttgart. No. 49/2020  
(<https://www.sowi.uni-stuttgart.de/dokumente/forschung/siss/SISS49.pdf>)

**Z U S A M M E N F A S S U N G:** Dieses Papier beschäftigt sich mit drei Themen: Erstens soll geklärt werden, ob Fragen zur psychischen Gesundheit in Onlinebefragungen heikel sind, zweitens soll untersucht werden, ob die Erhebung von Merkmalen in diesem Themenbereich durch sozial erwünschtes Antwortverhalten verzerrt wird und drittens soll geprüft werden, ob mittels Crosswise Modell wahrheitsgemäßere Antworten erzielt werden können als mit direkten Fragen. Dazu wurde eine Onlineerhebung mit Studierenden in Deutschland durchgeführt.

Im Ergebnis zeigt sich: (1) dass es sich bei der physischen Gesundheit um kein heikles Thema handelt, wenn dazu Studierende in einem Online-Survey befragt werden, (2) dass bei Onlineumfragen zu diesem Thema nicht mit sozial erwünschtem Antwortverhalten zu rechnen ist, sodass es dabei ausreichend ist, direkte Fragen zu stellen, (3) dass das Crosswise Modell keine Antworten mit erhöhter Reliabilität und erhöhter Validität liefert.

**Anna Götze und Andreas Wahl**

**Mental health: A sensitive subject in empirical survey research? Validating the Crosswise Model.**

SISS – Schriftenreihe des Instituts für Sozialwissenschaften der Universität Stuttgart. No. 49/2020  
(<https://www.sowi.uni-stuttgart.de/dokumente/forschung/siss/SISS49.pdf>)

**A B S T R A C T:** The paper addresses the following three objectives: Firstly, it aims to clarify whether mental health questions in online surveys are sensitive questions; secondly, it examines whether answers to questions in this subject area are distorted by socially desirable response behavior; and thirdly, it examines whether more truthful answers can be obtained using the Crosswise Model. For this purpose, an online survey was conducted with students in Germany.

The results show: (1) that physical health is not a sensitive issue when asking students in an online survey, (2) that online surveys on this topic will not necessarily produce socially desirable answers so that it is sufficient to ask direct questions, (3) that the Crosswise Model does not provide more reliable and therefore more valid answers.

## Inhaltsverzeichnis

Tabellenverzeichnis .....	2
Abbildungsverzeichnis .....	2
1 Einleitung .....	3
2 Analytischer Teil.....	4
2.1 Antwortverzerrung und Soziale Erwünschtheit.....	4
2.2 Das Crosswise Modell .....	6
2.3 Hypothesenableitung .....	9
3 Untersuchungsdesign .....	10
3.1 Datenerhebung und Stichprobenbeschreibung .....	10
3.2 Fragebogendesign .....	11
3.3 Operationalisierung.....	13
3.3.1 Die psychische Gesundheit .....	13
3.3.2 Soziale Erwünschtheit und Kontrollvariablen.....	13
4 Statistische Analyse .....	15
4.1 Prävalenzvergleich.....	15
4.2 Multivariate Analyse.....	17
4.3 Datenqualität.....	20
5 Diskussion und Fazit.....	21
6 Literaturverzeichnis .....	23
7 Anhang.....	26

## Tabellenverzeichnis

Tabelle 1: Prävalenzvergleich zwischen DQ und CM .....	16
Tabelle 2: Logistische Regressionsanalyse .....	18
Tabelle 3: Deskriptive Statistiken zur Datenqualität.....	20

## Abbildungsverzeichnis

Abbildung 1: Crosswise Modell.....	7
------------------------------------	---

## 1 Einleitung

Psychische Erkrankungen haben sich in den letzten Jahren zu einem immer größer werdenden Problem entwickelt. Laut TherapeutInnen sind besonders junge Personen und Studierende häufig von psychischen Erkrankungen betroffen (Benton et al. 2003; Seliger und Brähler 2007). Empirische Studien untermauern diese Aussagen und finden unter Studierenden gar Erkrankungsraten von etwa 20 % (vgl. Grobe und Steinmann 2015; Bailer et al. 2007), wobei besonders für ältere Studierende (über 26 Jahre) die Gefahr einer psychischen Erkrankung besteht (Grobe, Steinmann und Szecsenyi 2018). Die Gründe, die dafür genannt werden, sind vielfältig. Denn neben einer hohen Arbeitsbelastung, die mit der Zeit im Studium immer mehr zunimmt (Stock und Krämer 2001), und hohem, subjektiv wahrgenommenem Stress (Holm-Hadulla 1994) durchlaufen junge Erwachsene in dieser Zeit bedeutende Entwicklungs- und Reifungsprozesse, die oftmals mit einer hohen psychischen Belastung verbunden sind (Holm-Hadulla et al. 2009).

Die Erforschung des Vorkommens von psychischen Erkrankungen kann sich als schwierig erweisen, wenn z. B. keine Krankenkassendaten vorliegen, sondern persönliche Befragungen durchgeführt werden müssen. Fragen zur psychischen Gesundheit können als heikel angesehen werden, weil darüber nicht gerne gesprochen wird und sich Personen nur ungern damit assoziieren werden (DAK-Gesundheit 2013). Oft gestehen Betroffene eine Erkrankung auch nicht ein, weil sie damit eine eventuell verbundene soziale Stigmatisierung befürchten (Corrigan 2004). Zudem fürchten Betroffene häufig Nachteile bei Versicherungen, bei der Arbeitsplatzsuche oder am Arbeitsplatz (DAK-Gesundheit 2013). Aus diesen Gründen kann vermutet werden, dass Befragte bei Fragen nach ihrer psychischen Gesundheit nicht offen antworten. Jedoch können (subjektiv) nicht wahrheitsgemäße Antworten bei einer Analyse von Umfragedaten zu verzerrten Ergebnissen führen. Zwar deutet vieles auf die Sensitivität und Störanfälligkeit von Fragen zum Thema ‚psychische Gesundheit‘ hin; bisher herrscht aber wenig Klarheit darüber, ob mittels Umfragen erhobene Daten in diesem Themenbereich überhaupt zuverlässig sein können, oder ob die Daten durch subjektiv nicht wahrheitsgemäßes Antworten der Befragten generell als unzuverlässig einzustufen sind.

Um dem Problem der (subjektiv) nicht wahrheitsgemäßen Antworten bei heiklen Themen (und damit zu Fragen der psychischen Gesundheit) bereits bei der Erhebung zu entgehen bzw. um deren Anteil zu reduzieren, stehen der Umfrageforschung sogenannte indirekte Fragemethoden zur Verfügung. Dabei wird den Daten meist ein zufälliges Rauschen („random noise“) hinzugefügt, was die Anonymität der Befragten steigern soll (Höglinger et al. 2016). Eine indirekte Fragetechnik, welche sich dieses Rauschen zu Nutze macht, ist das Crosswise Modell (Yu, Tian und Tang 2008). Wie gut diese

Methode allerdings funktioniert, ob sie zur Merkmalerhebung bei heiklen Themen geeignet ist und auch validere Ergebnisse liefert als herkömmliche, direkte Fragen, ist umstritten.

Die vorliegende Studie befasst sich demnach mit zwei Aspekten: Einerseits soll geklärt werden, ob Fragen zur psychischen Gesundheit besonders störanfällig sind, und andererseits soll untersucht werden, ob bei Erhebungen in diesem Themenbereich mittels der indirekten Fragetechnik des Crosswise Modells (subjektiv) wahrheitsgemäßere Antworten erzielt werden können als mit direkten Fragen.

Im folgenden Kapitel werden nun zunächst die Grundlagen für mögliches (subjektiv) unwahrheitsgemäßes Antworten sowie das Crosswise Modell diskutiert. Zudem wird der diesbezügliche aktuelle Forschungsstand vorgestellt und es werden die zu testenden Hypothesen abgeleitet. Im dritten Kapitel wird das Untersuchungsdesign der hier vorgestellten Studie präsentiert. Das vierte Kapitel befasst sich mit der Datenanalyse, während im fünften Kapitel die Ergebnisse und ihre möglichen Implikationen diskutiert werden.

## **2 Analytischer Teil**

### **2.1 Antwortverzerrung und Soziale Erwünschtheit**

Subjektiv falsche Angaben bei Umfragen können zu Antwortverzerrungen führen. In diesem Falle wird mit den zur Antwortmessung benutzten Items nicht das gemessen, was die Items eigentlich messen sollten (Paulhus 1991). Zwei Gründe für Antwortverzerrungen, die in Bezug zum Inhalt einer zu beantwortenden Frage stehen, sind die soziale Erwünschtheit bestimmter Antworten und der Wunsch von Befragten nach sozialer Anerkennung (Gove und Geerken 1977).<sup>1</sup> Diese Art der Antwortverzerrung wird auch als ‚sozialer-Erwünschtheits-Bias‘ bezeichnet (Fisher und Katz 2000). Kennzeichnend dafür ist, dass die Befragten bestimmte Bewertungen, Einstellungen und Verhaltensweisen, die sie als unerwünscht wahrnehmen, eher verschweigen, um soziale Anerkennung zu erhalten – oder andersherum angeben, dass bei ihnen sozial erwünschte Eigenschaften vorliegen, obwohl es diese nicht tun.

Problematisch an dieser Art der Antwortverzerrung ist, dass empirisch in vielen Fällen nicht unterschieden werden kann, ob eine Person:

1. sich nur stark erwartungs- bzw. normkonform verhält;
2. eine (subjektiv) wahre Antwort abgibt, die allerdings u. A. auf ein positiv überhöhtes Selbstbild zurückzuführen ist;

---

<sup>1</sup> Eine Diskussion über verschiedenste Arten von Antwortverzerrungen, die nicht nur aufgrund von sozialer Erwünschtheit oder dem Streben nach sozialer Anerkennung entstehen, findet sich in Krause 2020.

3. bewusst versucht, einen positiven Eindruck zu hinterlassen, indem sie fälschlicherweise positive Dinge über sich selbst berichtet (Tourangeau und Yan 2007).

Während normkonformes Verhalten demnach kein sozial erwünschtes Antwortverhalten darstellt und Antworten, die darauf zurückzuführen sind, als ‚wahr‘ zu verstehen sind, gilt dies für die anderen beiden Fälle nur bedingt. Zwar macht eine Person im zweiten Fall eine richtige Angabe, allerdings kann diese verzerrt sein. Das ist z.B. dann der Fall, wenn einer befragten Person bestimmte Werte, wie Selbstbehauptung oder Selbstverwirklichung, besonders wichtig sind. Kommt es infolgedessen zur Selbsttäuschung und zu einem (positiv) verzerrten Selbstbild, dann wird nur ‚unterbewusst‘ eine falsche Angabe gemacht, weil die Person aufrichtig glaubt, dass die gemachte Angabe wahr sei (Paulhus 1984). Letztendlich liegt nur im letzten Fall eine bewusst falsche Angabe vor, da nur dabei die befragte Person weiß, dass die von ihr genannten Eigenschaften und gemachten Angaben nicht ‚wahheitsgemäß‘ sind. Sie versucht dann also aktiv, den Eindruck, den sie hinterlässt, zu ‚managen‘ bzw. ihn möglichst ‚positiv‘ zu gestalten und gibt dazu bewusst eine ‚unwahre‘ Antwort (Krumpal 2013; Paulhus 2002).

Die Gefahr einer Antwortverzerrung aufgrund des sozialen-Erwünschtheits-Bias ist demzufolge besonders bei der Erhebung von Daten zu heiklen Themen zu erwarten. Denn vor allem dann, wenn eine befragte Person das Gefühl hat, dass ihre eigentlich wahre Antwort auf eine Frage entweder nicht einer sozialen Erwartung/Norm entspricht oder soziale bzw. rechtliche Sanktionen nach sich ziehen könnte, wird sie sich eher dazu entscheiden, eine Frage unwahr zu beantworten.

Ob allerdings ein Effekt der sozialen Erwünschtheit bei Befragungen zu heiklen Themen tatsächlich existiert, ist bislang wenig klar, da bisherige Studien zu verschiedenen Ergebnissen kommen. Van de Mortel (2008) verglich bspw. 31 Studien, die unter anderem den Einfluss der sozialen Erwünschtheit untersuchten. In 43 % der Studien zeigte sich ein Effekt von sozialer Erwünschtheit, während in 45 % der untersuchten Studien ein solcher Effekt nicht zu erkennen war. Signifikante Effekte zeigten sich in Studien zur körperlichen Aktivität, zur häuslichen Gewalt, zu Angst und Depressionen sowie zu Essgewohnheiten. Des Weiteren ließen sich Effekte der sozialen Erwünschtheit bei Themen zur Arbeitsmotivation (Antin und Shaw 2012) und zur politischen Partizipation beobachten (Persson und Solevid 2014; Holbrook und Krosnick 2010) – bei letzterer Studie allerdings nur bei Telefon- und nicht bei Onlinebefragung. Auch bei Befragungen zum Alkoholkonsum wurde ein Effekt der sozialen Erwünschtheit gefunden (Davis, Thake und Vilhena 2010). Generell konnte festgestellt werden, dass der Effekt der sozialen Erwünschtheit geringer ausfällt, wenn Befragungen nicht persönlich, also meist online, durchgeführt werden (Heerwegh 2009). Es zeigte sich also, dass der Effekt oft nur unter bestimmten Bedingungen oder nur in einer bestimmten Form vorliegt.

Ergebnisse zum Einfluss der sozialen Erwünschtheit bei Fragen zur psychischen Gesundheit liegen dagegen kaum vor. Allein in der Studie von Stead et al. (2010) konnte eine geringe (bivariate) Korrelation zwischen sozialer Erwünschtheit und mentaler Gesundheit nachgewiesen werden.

Die unterschiedlichen Ergebnisse zum Effekt sozialer Erwünschtheit könnten dadurch entstehen, dass dieser in den verschiedenen Studien auf verschiedene Weise gemessen wird. Allerdings könnten auch vermeintlich heikle Themen (wie bspw. die psychische Gesundheit) von befragten Personen nicht als heikel wahrgenommen werden, weshalb sie keinen Grund dafür sehen, ‚unwahr‘ zu antworten.

Um bei heiklen Themen – deren Merkmal eben ist, dass sie Einstellungen und Verhaltensweisen ansprechen, die nicht unbedingt der sozialen Erwartung/Norm entsprechen oder auch rechtlich nicht konform sind – die Gefahr des sozial erwünschten Antwortverhaltens bereits bei der Erhebung zu umgehen, bieten sich indirekte Fragetechniken an. Eine dieser Techniken ist das Crosswise Modell.

## **2.2 Das Crosswise Modell**

Das Crosswise Modell ist eine Variante der 1965 erstmals von Warner vorgestellten Randomized-Response-Technik. Dieses hat gegenüber der Randomized-Response-Technik (und ihren verschiedenen Varianten; siehe dazu Wolter 2012), aber auch gegenüber anderen indirekten Fragetechniken wie dem Triangular Modell (Yu, Tian und Tang 2008) oder dem stochastischen Lügendetektor (Moshagen, Musch und Erdfelder 2012) entscheidende Vorteile: Erstens können Studien nachweisen, dass das Crosswise Modell im Vergleich zu anderen indirekten Fragetechniken am besten verständlich ist (Hoffmann et al. 2017). Zweitens wird, anders als bei anderen indirekten Fragemethoden, kein Zufallsgenerator benötigt, der anzeigt, welche Frage die Befragten beantworten sollen. Das macht die Technik besonders für die Verwendung in Onlinefragebögen oder für selbst auszufüllende Fragebögen interessant, da keine Interviewenden anwesend sein müssen. Drittens kann keine selbst-schützende Antwort beim Crosswise Modell gegeben werden, da keine Antwortkategorie existiert, auf welche die Befragten ausweichen könnten, um sicherzustellen, dass die heikle Eigenschaft keinesfalls vorliegt (Jann 2009; Yu, Tian und Tang 2008). Demzufolge gibt es beim Crosswise Modell keine eindeutige ‚richtige‘ Antwort, womit auch keine Verzerrung zugunsten einer selbst-schützenden Antwort vorliegen kann.

Im Folgenden wird nun die Funktionsweise des Crosswise Modells erläutert. Die Erläuterung bezieht sich auf den tatsächlichen Wortlaut und die Implementation des Crosswise Modells im Fragebogen, mit welchem die Daten für diese Studie erhoben wurden (Abbildung 1, siehe auch Kapitel 3).

Bei der Befragung per Crosswise Modell werden gleichzeitig zwei Fragen gestellt, die jeweils mit ‚Ja‘ oder ‚Nein‘ beantwortet werden können. Bei der einen Frage handelt es sich um die inhaltlich

bedeutsame heikle Frage (2. Frage in Abbildung 1), bei der anderen Frage handelt es sich um eine Frage, die für das inhaltliche Interesse der Befragung keine besondere Bedeutung hat; also um eine thematisch unabhängige Frage (auch ‚unrelated question‘; 1. Frage in Abbildung 1). Bei letzterer ist entscheidend, dass die Prävalenz, bzw. das Vorkommen der ‚Ja‘-Antwort, bekannt ist. Hier wird nach dem Geburtsmonat einer nahestehenden Person gefragt: ‚Hat ihr Vater in den Monaten Januar, Februar oder März Geburtstag?‘. Die Prävalenz für diese unabhängige Frage darf laut Vorgabe für das Crosswise Modell nicht 0,5 betragen. Sie liegt in diesem Fall näherungsweise bei 0,25 (3 aus 12 Monaten =  $0,25^2$ ).

### Abbildung 1: Crosswise Modell

Im Folgenden kommt eine besondere Fragetechnik zum Einsatz. Bitte lesen Sie die Anleitung sorgfältig durch, bevor Sie die Fragen beantworten.

[...] Bitte denken Sie zuerst darüber nach, wie Sie jede Frage einzeln beantworten würden (entweder mit *Ja* oder mit *Nein*). Danach markieren Sie bitte abhängig von Ihren Antworten Option (A) oder Option (B) nach den folgenden Regeln:

- Wenn Ihre Antwort *Nein* zu beiden Fragen oder *Ja* zu beiden Fragen lautet, kreuzen Sie Option (A) an.
- Wenn Ihre Antwort auf eine der Fragen *Ja* und auf die andere *Nein* lautet, kreuzen Sie Option (B) an.

- 1. Frage:** Hat Ihr Vater in den Monaten Januar, Februar oder März Geburtstag?  
**2. Frage:** Haben Sie seit Beginn des Studiums (mind. 1 x) aufgrund von psychischen Beschwerden eine Beratungsstelle (z. B. psychotherapeutische Beratung der Hochschule oder außeruniversitäre Stellen) aufgesucht?

Wie lauten Ihre Antworten auf die beiden Fragen? Bitte kreuzen Sie entsprechendes an.

- (A) *Nein* zu beiden Fragen oder *Ja* zu beiden Fragen  
 (B) *Ja* zu einer der Fragen und *Nein* zu der anderen

Quelle: BLIS 2019.

Das Besondere am Crosswise Modell ist, dass die Antwort auf die erste oder zweite Frage nicht direkt erfragt wird. Stattdessen werden die Befragten gebeten, anzugeben, ob ihre Antworten auf die beiden Fragen gleich sind (Option A: ‚Ja‘ als Antwort auf beide Fragen oder ‚Nein‘ als Antwort auf beide Fragen) oder ob sie sich voneinander unterscheiden (Option B: ‚Ja‘ als Antwort auf eine der Fragen und ‚Nein‘ als Antwort auf die andere Frage). Da die Befragten zwei Fragen gleichzeitig beantworten,

<sup>2</sup> Da es in der Gesamtbevölkerung Deutschlands keine Gleichverteilung der Geburten über das Jahr gibt, handelt es sich bei diesem Wert nicht um die tatsächliche Prävalenz. Zudem handelt es sich bei der zugrundeliegenden Stichprobe auch nicht um eine repräsentative Zufallsstichprobe, sodass der wahre Prävalenzwert unbekannt ist. Weil es allerdings keine Gründe gibt anzunehmen, dass für die Geburtsmonate der Väter keine annähernde Gleichverteilung gegeben ist, stellt der Wert von 0,25 einen annehmbaren Kompromiss dar.

bleibt unbekannt, auf welche Frage die Befragten mit ‚Ja‘ oder ‚Nein‘ geantwortet haben. Damit bleibt auch unbekannt, ob für die jeweils befragte Person das heikle Merkmal zutrifft oder nicht.

Weil nun aber die Prävalenz der unabhängigen Frage bekannt ist, kann die Prävalenz ( $\hat{\pi}$ ) des heiklen Merkmals auf Aggregatebene (bei einer repräsentativen Zufallstischprobe wäre dies das Vorkommen des Merkmals in der Gesamtbevölkerung) mit folgender Formel (aus: Yu, Tian und Tang 2008) errechnet werden, wobei  $\hat{\lambda}$  den Anteil derer angibt, die Option A wählen und  $p$  die Prävalenz der unabhängigen Frage darstellt:

$$\hat{\pi}_{CM} = \frac{\hat{\lambda} + p - 1}{2p - 1}. \quad (2.1)$$

Gleichermaßen lassen sich auch die Varianz dieses Schätzwertes und infolgedessen das dazugehörige Konfidenzintervall berechnen (Formel aus: Yu, Tian und Tang 2008):

$$Var(\hat{\pi}_{CM}) = \frac{\hat{\pi}_{CM} * (1 - \hat{\pi}_{CM})}{n - 1} + \frac{p * (1 - p)}{(n - 1) * (2p - 1)^2}. \quad (2.2)$$

Auf diese Weise ermöglicht es das Crosswise Modell einen Schätzwert für die wahre Prävalenz zu berechnen, ohne dass bekannt sein muss, ob für die jeweils befragte Person das heikle Merkmal zutrifft oder nicht. Damit kann den Befragten zusätzliche Anonymität zugesichert werden. Auch brauchen die Befragten keine (sozialen) Sanktionen zu fürchten, weil eben den Interviewenden oder allen, die mit den Daten in Berührung kommen, unbekannt ist, welche der beiden Fragen nun mit ‚Ja‘ oder ‚Nein‘ beantwortet wurde.

Auch Validierungsstudien können nachweisen, dass das Crosswise Modell bei heiklen Fragen valide(re) Ergebnisse liefert. Hierbei ist anzumerken, dass es, grob betrachtet, zwei Arten von Validierungsstudien gibt. Diese unterscheiden sich danach, ob mit weichen oder mit harten Kriterien gearbeitet wird. Bei Validierungsstudien, die mit harten Kriterien arbeiten, ist bereits vor der eigentlichen Erhebung klar, wie viele befragte Personen in der Stichprobe das zu erhebende Merkmal aufweisen: die Prävalenz ist bekannt. Wenn die bei der Erhebung eingesetzte Fragemethode zuverlässig arbeitet, können die bekannten Prävalenzwerte (mit nur geringen Abweichungen) reproduziert werden. Mittels solcher harten Validierungsstudien kann genau festgestellt werden, wie nahe die geschätzten Prävalenzen an den wahren Prävalenzen liegen und somit auch ermittelt werden, wie gut die Fragemethode funktioniert. Eine Studie, die das Crosswise Modell nach solchen harten Validierungskriterien untersucht, ist die von Hoffmann et al. (2015). Sie weisen nach, „that the CWM [Crosswise Modell; die Verf.] is convincingly capable of obtaining valid prevalence estimates of sensitive attitudes and behaviors. We therefore conclude that the CWM appears to be a very promising indirect questioning

technique that can be used to successfully control for social desirability on surveys of sensitive behavior“ (Hoffmann et al. 2015: 409).

In den meisten Studien sind die wahren Prävalenzen jedoch nicht bekannt. Deshalb werden dort weiche Validierungskriterien eingesetzt. Dabei handelt es sich meist um Annahmen nach dem Prinzip ‚less-is-better‘ oder ‚more-is-better‘ (Tourangeau, Yan 2007). Die more-is-better-Annahme besagt, dass eine höhere Prävalenz einer sozial unerwünschten Eigenschaft oder eines sozial unerwünschten Verhaltens näher am wahren Wert liegt als eine niedrigere. Bei der less-is-better-Annahme ist dies umgekehrt. Um diese Annahmen testen zu können, muss eine zweite Fragemethode eingesetzt werden, die als Vergleichsparameter dient. Prinzipiell können dabei auch zwei indirekte Fragetechniken miteinander verglichen werden. Meist wird als Vergleichsparameter jedoch die direkte Frage benutzt. Bei der more-is-better-Annahme geht man dann davon aus, dass diejenige Fragetechnik, die höhere Prävalenzwerte liefert (bei der less-is-better-Annahme sind es niedrigere Werte), auch zuverlässiger arbeitet und demnach validere Ergebnisse erbringt.

Die meisten Validierungsstudien, die sich mit dem Crosswise Modell auseinandersetzen, nutzen die more-is-better-Annahme. Jann et al. (2012) untersuchen auf diese Weise Plagiate unter Studierenden, Korndörfer et al. (2014) Steuerhinterziehung, Shamsipour et al. (2014) den Konsum von illegalen Drogen und Hoffmann und Musch (2019) Vorurteile gegenüber Frauen. All diese Studien zeigen im Ergebnis, dass das Crosswise Modell eine signifikant höhere Prävalenz erzielt als die direkte Frage (bis auf die letztgenannte Studie, bei der nur bei den befragten Frauen ein signifikanter Unterschied gefunden wird, nicht aber bei den Männern). Letztendlich können also verschiedene Studien nachweisen, dass das Crosswise Modell eine zuverlässige indirekte Fragetechnik ist.

### **2.3 Hypothesenableitung**

Bereits in der Einleitung zu dem vorliegenden Text wurde erwähnt, dass Daten zur psychischen Gesundheit, sollten sie nicht von Krankenkassen, sondern durch Erhebungen erfasst worden sein, eher wenig zuverlässig sind. Das kann vor allem daran liegen, dass die psychische Gesundheit von den befragten Personen als ein heikles Thema wahrgenommen wird, über das nur sehr ungern Auskunft erteilt wird. Sollte es sich demnach bei der psychischen Gesundheit um ein heikles Thema handeln, dann kann in der empirischen Umfrageforschung damit gerechnet werden, dass mittels indirekter Fragetechniken validere Ergebnisse erzielt werden können als mit direkten Fragen. Deshalb lässt sich folgende Hypothese formulieren:

*Fragen zur psychischen Gesundheit sind heikel; es lassen sich nach der more-is-better-Annahme mit dem Crosswise Modell höhere Prävalenzen beobachten als mit einer direkten Frage (Hypothese 1).*

Weiterhin zeigt sich in den obigen Überlegungen zur sozialen Erwünschtheit, dass das Antwortverhalten nicht immer durch einen Effekt der sozialen Erwünschtheit beeinflusst wird, dass aber bei Befragungen zu heiklen Themen oftmals Antwortverzerrungen aufgrund des sozialen-Erwünschtheits-Bias auftreten. Deshalb ist davon auszugehen, dass, wenn sich die psychische Gesundheit als heikles Thema herausstellen sollte, bisherige Umfragen eher verzerrte Daten geliefert haben und deshalb auch nur wenig valide Ergebnisse erbracht haben. Es kann somit folgende Hypothese formuliert werden:

*Sofern es sich bei der psychischen Gesundheit um eine heikle Thematik handelt, lassen sich bei diesbezüglichen Befragungen im Antwortverhalten auch Effekte der sozialen Erwünschtheit beobachten (Hypothese 2).*

Zum Test der beiden hier formulierten Hypothesen werden die Daten einer eigens durchgeführten Erhebung verwendet. Das Fragebogendesign, die zugrundeliegende Stichprobe und die gewonnenen Daten werden im kommenden Kapitel 3 vorgestellt.

### **3 Untersuchungsdesign**

#### **3.1 Datenerhebung und Stichprobenbeschreibung**

Die Daten wurden per Onlineumfrage mittels der Survey-Software ‚Unipark‘ im Rahmen des Lehrforschungsprojektes ‚Der Alltag von Studierenden: Belastungen und Leistungsdruck im Studium‘ (kurz: BLIS) am Institut für Sozialwissenschaften der Universität Stuttgart erhoben. Ziel des Projektes war die Validierung verschiedener indirekter Fragetechniken bei heiklen Themen; darunter das Crosswise Modell bei Fragen zur psychischen Gesundheit.

Die Befragung selbst wurde zwischen dem 17.06.2019 und dem 14.07.2019 durchgeführt. Zur Teilnahme eingeladen wurden alle online erreichbaren Universitätsstudierenden in Deutschland. Die Einladungen wurden an mehr als 300 verschiedene Mailverteiler, Newsletter und Sekretariate deutscher Universitäten mit der Bitte zur Weiterleitung an ihre Studierenden verschickt; eine Zufallsstichprobe liegt demzufolge nicht vor.<sup>3,4</sup> Nach den ersten beiden Wochen wurde eine Erinnerung an eben diese Verteiler versendet, um nochmals einige TeilnehmerInnen zu akquirieren. Die Teilnahme an der Befragung war freiwillig und anonym<sup>5</sup>.

---

<sup>3</sup> Da es sich um keine Zufallsstichprobe handelt und außerdem die wahren Prävalenzen für die heiklen Merkmale in dieser Stichprobe unbekannt sind, kann diese Studie lediglich eine weiche Validierung des Crosswise Modells vornehmen.

<sup>4</sup> Weiterhin kann nicht ausgeschlossen werden, dass Studierende mehrmals an der Umfrage teilnahmen. Um dies allerdings zu tun, mussten die Studierenden entweder von unterschiedlichen IP-Adressen auf die Umfrage zugreifen, oder die gespeicherten Cookies löschen.

<sup>5</sup> Hypothetisch wäre es möglich, die Identität der befragten Person über die verwendete IP-Adresse zurückzuverfolgen. Das gilt allerdings nur, wenn sichergestellt werden kann, dass nur diese eine Person diese eine IP-Adresse verwendet.

Im Einladungs- und Erinnerungsschreiben fanden die Studierenden einen Link zur Startseite der Umfrage. Bevor sie allerdings den Fragebogen ausfüllen konnten, mussten sie dazu ihr Einverständnis abgeben. In der Einverständniserklärung wurden den Studierenden Informationen zum Forschungsgegenstand, zu den Zielen und dem Nutzen der Umfrage, zum Ablauf und zur Dauer der Befragung sowie Angaben zur Vertraulichkeit und Nutzung der Daten gegeben.

Insgesamt konnte eine Gesamtstichprobe von 1685 befragten Personen erzielt werden. 23 % dieser befragten Personen verweigerten ihr Einverständnis, sodass sie nicht an der Umfrage teilnehmen konnten und weitere 14 % brachen die Befragung im Verlauf ab. Knapp 63 % der befragten Personen beendeten dagegen die Befragung. Das Netto-Sample besteht demzufolge aus 1062 analysierbaren Fällen. Zur Beendigung der Umfrage benötigten die Befragten im Durchschnitt knapp 32 Minuten; das sind 7 Minuten mehr, als zu Beginn der Umfrage angekündigt wurde. Nach listenweisem Fallausschluss<sup>6</sup> der Missing Values (worunter Item-Nonresponses zu verstehen sind) standen für die statistischen Analysen dieser Studie insgesamt 754 Fälle zur Verfügung.<sup>7</sup>

### 3.2 Fragebogendesign

Die Umfrage gliederte sich in fünf Abschnitte. Zunächst wurden Fragen zu den soziodemographischen Merkmalen, zum Studium und zur finanziellen und wohnlichen Situation des/der Befragten gestellt. Der zweite Abschnitt befasste sich mit speziellen Fragen rund um das Studium (Lernverhalten, Anzahl der Prüfungen etc.). Dabei wurden auch Fragen eingebunden, welche die Studierenden womöglich als heikel empfinden konnten, weil damit strafrechtliche Konsequenzen verbunden sein könnten (Thema ‚Plagiate‘). Im dritten Abschnitt wurden schließlich Fragen zur psychischen Gesundheit und damit in Verbindung stehende Themen präsentiert. Im vierten Abschnitt wurden Fragen zum allgemeinen gesellschaftlichen Zusammenleben gestellt. Darunter waren unter anderem Fragen zum Gerechtigkeitsempfinden, zur Risikobereitschaft und weitere Einstellungsfragen. Die Umfrage

---

Allerdings könnten sich die befragten Personen durchaus bewusst sein, dass sie nicht komplett anonym antworten, weshalb die Sicherstellung von deren Anonymität und der Umgang mit den gewonnenen Daten von großer Relevanz ist. Aus diesem Grund wurde den befragten Personen der wissenschaftliche Charakter, die Zusage der Anonymität und der Umgang mit den Daten deutlich dargelegt.

<sup>6</sup> Grundsätzlich hätte eine (multiple) Imputation, anstatt der Ausschluss der fehlenden Werte, zu einer Erhöhung der analysierbaren Fälle geführt. Allerdings arbeitet das Crosswise Modell mit bestimmten Wahrscheinlichkeiten (hier: die Geburtsmonate der Väter), sodass nicht direkt imputiert werden könnte, weil die hinter den Imputationen stehenden logistischen Regressionen keine Anpassungen an diese Wahrscheinlichkeiten zulassen, sodass die Imputation ohne deren Berücksichtigung von statten gehen würde. Es könnte demzufolge zu ‚falsch‘ imputierten Werten kommen. Was die direkte Fragetechnik betrifft, hätte die Möglichkeit der (multiplen) Imputation mittels logistischer Regressionen bestanden. Um der Einheit willen und um zu vermeiden, dass die Daten der beiden Fragetechniken nicht mit unterschiedlichen Methoden aufbereitet werden, die ggf. einen Vergleich der beiden Techniken erschweren, wird darauf verzichtet.

<sup>7</sup> Die Differenz zwischen 1062 und 754 ist nicht nur auf fehlende Werte aufgrund von Item-Nonresponses zurückzuführen. 258 Fälle fehlen aufgrund dessen, weil diese befragten Personen eine andere Version des Fragebogens ausfüllten, indem für die betreffende Frage zur psychischen Gesundheit eine andere indirekte Fragetechnik implementiert war (siehe Kapitel 3.2). Dementsprechend gehen durch den listenweisen Fallausschluss nur 50 Fälle verloren (was ein weiterer Grund für den listenweisen Fallausschluss ist).

endete mit verschiedenen Fragen, die die Qualität der Daten überprüfen und sichern sollten, darunter auch Fragen zur Wahrnehmung und Einschätzung der heiklen Fragen. Insgesamt beantworteten die Teilnehmenden 48 Fragen.

Bevor die eigentliche Umfrage gestartet wurde, wurde der Fragebogen in zwei Pretest-Verfahren geprüft. Beim ersten Pretest wurden ca. 30 willkürlich ausgewählte Personen mit dem ausgedruckten Fragebogen schriftlich befragt. Hierbei interessierte vor allem die Verständlichkeit der Fragen. Aufgrund der Ergebnisse dieses Pretests wurden Veränderungen an den Formulierungen der Fragen, bzw. an der Fragenreihenfolge vorgenommen. Sodann wurde der Fragebogen in das Online-Tool ‚Unipark‘ übertragen. Es folgte die zweite Pretest-Phase, bei welcher Studierende der Universität Stuttgart, die im Sommersemester 2019 das Seminar ‚Standardisierte Erhebungsmethoden‘ besuchten, angewiesen wurden, den Onlinefragebogen mit möglichst verschiedenen Endgeräten zu testen, um die Kompatibilität mit einer Vielzahl an technischen Geräten, unterschiedlichen Betriebssystemen, Browsern etc. sicherzustellen. Während die Anwendung des Tools auf allen Endgeräten funktionierte, mussten einige Veränderungen an den Fragen vorgenommen werden. Nach diesem Pretest ging die Umfrage in die Feldphase.

Der Test der more-is-better-Annahme benötigt, wenn keine wahren Prävalenzen bekannt sind, einen Vergleichsparameter. In den meisten Studien zur Validierung von indirekten Fragetechniken ist dies die direkte Frage nach dem heiklen Merkmal. Dementsprechend waren für die hier vorgestellte Studie zwei verschiedene Fragebogenvarianten notwendig, die sich darin unterschieden, ob die heiklen Fragen zur psychischen Gesundheit mittels indirekter oder direkter Frage gestellt wurden. Zudem waren, weil im Lehrforschungsprojekt BLIS insbesondere die Validierung verschiedener indirekter Fragetechniken bei unterschiedlichen heiklen Thematiken interessierte, noch weitere Fragebogenvarianten notwendig.

Um sicherzustellen, dass die heiklen Themen mit den indirekten Fragetechniken abgefragt werden konnten und um einen Vergleichsparameter durch die direkten Fragen zu besitzen, mussten insgesamt vier verschiedene Fragebogenvarianten eingesetzt werden. Diese unterschieden sich dadurch, dass die Fragen nach den heiklen Merkmalen mittels indirekter oder direkte Abfrage erfolgten. Weiterhin mussten für alle indirekten Fragetechniken genügend Fälle generiert werden (Ziel war es, mindestens 200 Fälle pro indirekter Fragetechnik zu erreichen), weshalb in einer der vier Fragebogenvarianten jeweils zwei verschiedene indirekte Fragetechniken angewendet wurden.

Weil die indirekten Fragetechniken ein gewisses Maß an Konzentration und Verständnis erforderten, wurden diesen Einleitungstexte vorangestellt (siehe Abbildung 1). Die Erklärungen zu den Funkti-

onsweisen der Techniken wurden entweder direkt, leicht angepasst oder aus dem Englischen übersetzt aus anderen Studien, in denen die Verständlichkeit der Texte bereits geprüft worden war, übernommen (Wolter 2012; Höglinger, Jann und Diekmann 2016; Jann, Jerke und Krumpal 2012).

Mittels eines Zufallsverfahren wurden die befragten Personen zu Beginn der Befragung einer der vier Fragebogenvarianten zugeteilt: Von den 754 Befragten, die für die Analyse zur Verfügung standen, antworteten 524 auf die direkten Fragen nach der psychischen Gesundheit, wohingegen 230 Befragte diese Fragen mittels des Crosswise Modells beantworteten.

### **3.3 Operationalisierung**

#### *3.3.1 Die psychische Gesundheit*

Die psychische Gesundheit von Studierenden wurde mit der Frage erfasst, ob eine Person seit dem Beginn ihres Studiums aufgrund von psychischen Beschwerden mindestens einmal eine Beratungsstelle aufgesucht hat. Das konnten psychotherapeutische Beratungsstellen der Hochschule oder außeruniversitäre Beratungsstellen gewesen sein. Die Antwort auf diese Frage konnte entweder mit ‚Ja‘ oder ‚Nein‘ erfolgen (siehe Abbildung 1).

Die Operationalisierung der psychischen Gesundheit mittels dieser Frage wurde gewählt, weil der Besuch einer Beratungsstelle indiziert, dass eine psychische Erkrankung mit einem gewissen Ausmaß an Schwere vorliegt. Auf diese Weise konnten vermutlich mehr Studierende mit einer psychischen Erkrankung erfasst werden, als wenn direkt nach einer tatsächlichen Diagnose gefragt worden wäre.

Der Besuch einer Beratungsstelle sollte demzufolge eine psychische Erkrankung bei Studierenden indizieren,

- da damit bereits ein gewisses Maß an Schwere einer psychischen Erkrankung verbunden ist;
- da mit dieser Frage auch diejenigen Studierenden angesprochen werden können, die sich mit ihrer psychischen Erkrankung noch weiterhin im Studium befinden;
- da die Frage auch diejenigen Studierenden anspricht, die noch ohne Diagnose einer psychischen Erkrankung sind und die ihr Studium noch nicht unterbrochen oder abgebrochen haben.

#### *3.3.2 Soziale Erwünschtheit und Kontrollvariablen*

Die Tendenz zur sozialen Erwünschtheit beim Antwortverhalten der Befragten wurde mittels der KSE-G-Kurzskala von Kemper et al. (2012) erhoben. Laut Beschreibung der Skala können die verwendeten Items zu zwei Faktoren zusammengefasst werden, welche die Dimensionen ‚Übertreibung positiver Qualitäten‘ (PQ+) und ‚Untertreibung negativer Qualitäten‘ (NQ-) messen. Mittels Hauptkomponentenanalyse konnten diese beiden Dimensionen auch im vorliegenden Datensatz statistisch

nachgewiesen werden. Beide Faktoren konnten die Gesamtvarianz zu knapp 65 % binden und auch die gemessenen Items laden eindeutig auf jeweils eine der beiden Komponenten (siehe Anhang A1<sup>8</sup>). Mit den Ergebnissen der Hauptkomponentenanalyse können die analysierten Items zu zwei additiven Indizes verbunden werden, bei denen die Faktorwerte als Gewichtungen benutzt werden. Diese Indizes können als eigenständige Variablen abgespeichert und für die weiteren Analysen verwendet werden. Sie wurden als SDPQ (Social Desirability: Übertreibung positiver Qualitäten) und SDNQ (Social Desirability: Untertreibung negativer Qualitäten) bezeichnet. Hohe Werte bei SDPQ bedeuten, dass eine befragte Person sozial erwünschte Verhaltensweisen übertreibt, während niedrige Werte bei SDNQ bedeuten, dass die Person sozial negativ konnotierte Verhaltensweisen untertreibt. In beiden Fällen versucht sich die befragte Person also möglichst positiv darzustellen.

Mittels logistischer Regressionsanalyse sollte geprüft werden, ob ein Zusammenhang zwischen der sozialen Erwünschtheit und der psychischen Gesundheit besteht. Lässt sich ein solcher beobachten, wäre das ein Hinweis darauf, dass die Antworten auf die Fragen nach der psychischen Gesundheit durch sozial erwünschtes Antwortverhalten verzerrt werden. Zur Kontrolle dieser möglichen Effekte wurden in die multivariate Analyse weitere Kontrollvariablen aufgenommen. Dazu gehörten: das Geschlecht, das Alter und die Studiendauer der befragten Personen sowie deren (wahrgenommene) Fremdstigmatisierung von psychischen Erkrankungen.

Das Geschlecht wurde als Kontrollvariable verwendet, da sich die Betroffenheit von diagnostizierten psychischen Erkrankungen zwischen Männern und Frauen unterscheidet (Bundespsychotherapeutenkammer 2011). Für das Geschlecht wurde eine Dummyvariable erstellt, bei der die Ausprägung ‚weiblich‘ den Wert 0 zugewiesen bekam und ‚männlich‘ den Wert 1. Als weitere Kontrollvariable wurde das Alter der befragten Studierenden in die Modellschätzung aufgenommen. Zudem wurde die Studiendauer als Kontrollvariable behandelt, da mit ansteigender Studienzeit (besonders gegen Ende oder nach Überschreitung der Regelstudienzeit) auch der Erwartungsdruck ansteigt, das Studium baldmöglichst abschließen zu müssen, was zu psychischen Belastungen führen kann. Ergänzend wurde ein Zusammenhang zwischen einer vorliegenden psychischen Erkrankung und der damit einhergehenden Stigmatisierung durch andere Personen vermutet. Um die Schätzung diesbezüglich kontrollieren zu können, wurde die Stigmatisierung mittels des empfundenen Fremdstigmas operationalisiert. Dazu berichteten die Befragten nicht ihre eigene Einstellung zu psychischen Erkrankungen, sondern schätzten ein, welches Ausmaß die Einstellungen zu psychischen Erkrankungen in der allgemeinen Öffentlichkeit ihrer Meinung nach aufweisen. Das Fremdstigma wurde in der Analyse mittels zweier Items aus der Skala von Kendra et al. (2014) gemessen: Dazu gehört die Auffassung, dass es

---

<sup>8</sup> Die Aufbereitung der Daten und die deskriptiven Analysen erfolgten mit SPSS 25.

besser ist, psychische Probleme vor anderen zu verbergen („PEverbergen“) und andererseits die Erwartung, dass mit psychischen Problemen soziale Vorurteile einhergehen („PEvorurteile“). Beide Variablen sollten die Fremdstigmatisierung erfassen können, weil beide auf die Zieldimension von negativen Konsequenzen durch Andere beim Eingeständnis von psychischen Erkrankungen abzielen und somit die Sensitivität der Thematik ansprechen (Goffman 1974).

## 4 Statistische Analyse

Weil die befragten Personen jeweils einer von zwei unterschiedlichen Gruppen mit jeweils unterschiedlichen Fragebogenvarianten zugeordnet worden waren, wurden die Daten getrennt nach diesen beiden Gruppen analysiert. Bei der einen Gruppe handelte es sich um die Befragten, welche die heikle Frage direkt beantworteten (DQ), bei der anderen Gruppe handelte es sich um die Befragten, welche die heikle Frage mittels Crosswise Modell beantworteten (CM). Im Anhang A2 werden die Kennwerte für die deskriptive Statistik der unabhängigen Modellvariablen sowie der Kontrollvariablen aufgelistet. Zudem findet sich dort ein t-Test bezüglich der metrisch definierten unabhängigen Variablen sowie ein exakter Test nach Fisher mit der Geschlechtsvariablen zum Test auf etwaige Unterschiede zwischen den Befragten der DQ- und der CM-Gruppe. Beide Tests zeigen, dass sich die Gruppen in den unabhängigen Variablen nicht systematisch voneinander unterscheiden, sodass anzunehmen ist, dass die Befragten rein zufällig den jeweiligen Fragebogenvarianten zugeordnet worden waren.

### 4.1 Prävalenzvergleich

Die Prävalenz der direkten Frage, der dazugehörige Standardfehler (SE) und das 95 %ige Konfidenzintervall wurden direkt mittels SPSS berechnet, wohingegen die dementsprechenden Werte für die indirekte Frage mittels CM nach den Formeln in Kapitel 2.2 berechnet wurden. Die Ergebnisse werden in der folgenden Tabelle 1 aufgelistet. Demnach beträgt die Prävalenz von Studierenden, die angeben, seit Beginn des Studiums schon einmal einen Termin in einer psychologischen Beratungsstelle wahrgenommen zu haben, bei der direkten Frage etwas über 21 %. Demnach war also etwa ein Fünftel der Studierenden schon einmal von einer psychischen Erkrankung betroffen gewesen. Dieses Ergebnis entspricht in etwa den gefundenen Erkrankungsraten in Grobe und Steinmann (2015) sowie in Bailer et al. (2007).

In der Gruppe, bei welcher dieses Merkmal mittels indirekter Frage erhoben worden war, beträgt die Prävalenz jedoch nur knapp 13 %. Mithin ist konträr zur more-is-better-Annahme die Prävalenz mittels CM-Messung um einiges geringer als bei der Messung mittels direkter Frage.

*Tabelle 1: Prävalenzvergleich zwischen DQ und CM*

	Prävalenz in % (SE in %)	95 %iges Konfidenzintervall in %		Fallzahl
Direkte Frage (DQ)	21,18 (1,79)	17,68	24,69	524
Crosswise Modell (CM)	12,6 (6,13)	0,59	24,62	230

Quelle: BLIS 2019, eigene Berechnungen.

Einen Hinweis darauf, ob sich diese beiden geschätzten Werte signifikant unterscheiden, und damit, ob die gewählte Fragemethode zu systematisch anderen Ergebnissen führt, liefert ein Vergleich der Konfidenzintervalle. Überlappen sich diese nicht, dann deutet das auf einen signifikanten Unterschied zwischen den Prävalenzen hin. Dabei ist anzumerken, dass das Konfidenzintervall des CM-Prävalenzwertes deutlich breiter als dasjenige des DQ-Wertes ist. Das liegt daran, dass für das CM-Design weniger Fälle vorliegen, ist aber vor allem dadurch bedingt, dass indirekte Fragetechniken aufgrund der Randomisierung mittels Zufallsmechanismen (hier die Frage nach dem Geburtstag) den Antworten keine systematische Varianz hinzufügen (Hoffman und Musch 2016). Selbst bei gleicher Fallzahl wäre demnach das Konfidenzintervall der CM-Prävalenz größer als das der direkten Frage.

Im vorliegenden Fall überschneiden sich die beiden Konfidenzintervalle deutlich. Das bedeutet, dass die beiden Befragungsmethoden keine unterschiedlichen Ergebnisse im Hinblick auf die Prävalenzen generieren und dass die beobachteten Unterschiede zufallsbedingt entstanden sind. Zur Absicherung dieses Hinweises kann zusätzlich ein zweiseitiger t-Test (Formel aus: Jerke und Krumpal 2013) durchgeführt werden:

$$t = \frac{|\hat{\pi}_{DQ} - \hat{\pi}_{CM}|}{\sqrt{SE_{\hat{\pi}_{DQ}}^2 + SE_{\hat{\pi}_{CM}}^2}} \quad (4.1)$$

Sollte der berechnete t-Wert kleiner als -1,96 bzw. größer als 1,96 sein, dann ist die Differenz zwischen den Werten mit 5 %iger Irrtumswahrscheinlichkeit signifikant. Hier ergibt sich ein t-Wert von 1,34. Auch dies zeigt, dass keine signifikanten Unterschiede zwischen den Antworten beider Fragetechniken vorliegen.

Letztendlich wird also die more-is-better-Annahme und damit Hypothese 1, wonach mittels CM ein höherer Prävalenzwert beobachtet werden sollte, zurückgewiesen: Zwischen den Messungen von psychischen Erkrankungen mittels direkter und indirekter Fragetechnik gibt es bei diesen Daten keine Unterschiede. Demzufolge kommen beide Techniken zum selben Ergebnis und etwaige Unterschiede in den Prävalenzen sind zufallsbedingt und nicht systematisch. Dieses Ergebnis deutet darauf hin, dass es sich bei Fragen nach psychischen Erkrankungen um keine heiklen Fragen handelt; zumindest nicht für die befragten Studierenden im Rahmen dieser Onlinebefragung. Um diese Aussage weiter

zu stützen, werden im anschließenden Kapitel die Ergebnisse einer multivariaten logistischen Regressions-schätzung vorgestellt.

## 4.2 Multivariate Analyse

Bevor auf die Ergebnisse der multivariaten Analysen eingegangen wird, ist auf mehrere Punkte hinzuweisen: Zum einen wurde bei der DQ-Analyse die abhängige Variable umcodiert; die Nein-Kategorie erhält den Wert 1. Das erfolgte deshalb, weil es sich bei der Nein-Antwort um die sozial erwünschte Antwort handelt und weil in der Analyse positive Zusammenhänge mit den unabhängigen Variablen der sozialen Erwünschtheit erwartet werden. Je positiver sich eine Person darstellt, umso größer sollte die Wahrscheinlichkeit werden, dass diese Person die Frage nach der psychischen Erkrankung verneint. Zum anderen war für die CM-Analyse aufgrund der eher geringen Fallzahl ( $N = 230$ ), unvollständiger Information und teilweise vollständiger Separation keine stabile Schätzung des Analysemodells mit den Kontrollvariablen möglich. Demzufolge werden hier nur die Ergebnisse der logistischen Regressionsanalyse mit den beiden Variablen der sozialen Erwünschtheit präsentiert. Zudem musste für die logistische Regression der CM-Daten ein Spezialpaket verwendet werden, das die Randomisierungskomponente bei der Schätzung miteinbezieht. Mittels des R-Pakets<sup>9</sup> ‚RReg‘ (Version 0.7.1) war eine solche Schätzung möglich. Weil sich die geschätzten Prävalenzen der DQ- und der CM-Daten nicht voneinander systematisch unterscheiden, sollten auch die Ergebnisse der logistischen Regressionen ähnlich sein. Dass mit den CM-Daten keine komplexeren Analysen möglich sind, kann demzufolge durch die DQ-Analyse aufgefangen werden.

Werden zunächst die Maße des Modellfits analysiert, fällt auf, dass beide Modelle (CM- und DQ-Modell I in Tabelle 2) eine schlechte Anpassung aufweisen: In beiden Fällen verbessert sich die Vorhersage der abhängigen Variablen durch die Hinzunahme der Variablen zur sozialen Erwünschtheit nicht. Auch ist der Likelihood-Ratio-Test (LR-Test) jeweils nicht signifikant. Für das DQ-Modell kann zudem Nagelkerkes- $R^2$  betrachtet werden; hier zeigt sich, dass keine interpretationswürdige Verbesserung des Modellfits gegenüber dem Nullmodell erzielt werden kann, wenn die unabhängigen Variablen hinzugenommen werden.

Sowohl das DQ- als auch das CM-Modell I weisen nicht signifikante Koeffizienten für die Erwünschtheitsdimensionen auf. Gleichzeitig unterscheiden sich die Logits der beiden Modelle nicht signifikant voneinander (geprüft mit Formel 4.1, nicht abgetragen; siehe Urban und Mayerl 2018). Demzufolge liegen keine systematischen Unterschiede vor und ein Moderationseffekt, wonach das

---

<sup>9</sup> R wird in der Version 4.0.0 mittels R-Studio (Version 1.2.5042) verwendet.

CM-Modell den möglichen Einfluss der sozialen Erwünschtheit aufgrund der Randomisierung und der dadurch induzierten, zusätzlichen Anonymität, reduziert, lässt sich nicht nachweisen.

Zwar könnte die Aussagekraft der Signifikanz der Koeffizienten aufgrund der geringen Fallzahl (vor allen Dingen beim CM-Modell) und der damit einhergehenden geringen Teststärke eingeschränkt sein, allerdings zeigt auch die Ausprägung der AMEs, dass die soziale Erwünschtheit nur einen sehr geringen Effekt (wenn überhaupt) auf die abhängige Variable hat (der zudem für die SDPQ-Variable auch in eine nicht erwartete Richtung zeigt).

Tabelle 2: Logistische Regressionsanalyse

	DQ-Analyse				CM-Analyse	
	Modell I		Modell II		Modell I	
	Logits (SE)	AME in % (95 % KI)	Logits (SE)	AME in % (95 % KI)	Logits (SE)	AME in % (95 % KI)
Konstante	1,32*** (0,11)		6,20*** (0,97)		-2,25*** (0,91)	
SDPQ	-0,10 (0,11)	-1,70 (-5,24; 1,85)	-0,18 (0,12)	-2,66 (-6,12; 0,82)	-0,66 (0,68)	-2,25 (-8,25; 3,75)
SDNQ	0,05 (0,11)	0,90 (-2,59; 4,39)	0,07 (0,12)	0,98 (-2,43; 4,39)	0,71 (0,59)	3,68 (-2,50; 9,86)
Geschlecht			0,81** (0,28)	11,06** (4,43; 17,69)		
Alter			-0,11*** (0,03)	-1,67*** (-2,63; -0,71)		
Studienjahr <sup>10</sup>			-0,18* (0,07)	-2,65* (-4,78; -0,52)		
PEverbergen			-0,26* (0,10)	-3,87* (-6,93; -0,82)		
PEvorurteile			-0,27* (0,13)	-4,08* (-8,01; -0,14)		
Fallzahl		524		524		230
LL-Nullmodell		-270,58		-270,58		-142,95
LL-Modell I bzw. II		-269,96		-247,21		-142,47
LR-Test signifikant?		Nein		Ja		Nein
McFadden-R <sup>2</sup>		0,002		0,085		-- <sup>11</sup>
Nagelkerke-R <sup>2</sup>		0,004		0,133		--

Note: Abhängige Variable DQ: 0: Ja, 1 = Nein; Abhängige Variable CM: 0: Nein zu beiden Fragen oder Ja zu beiden Fragen, 1 = Ja zu einer der Fragen und Nein zu der anderen; SDPQ/SDNQ: Soziale Erwünschtheit, standardisiert; Geschlecht: 0 = weiblich, 1 = männlich; Alter: in Jahren; Studienjahr: in Jahren; PEverbergen/PEvorurteile: Skala von 1 (keine Zustimmung) bis 5 (vollkommene Zustimmung); \*p<0,05; \*\*p<0,01; \*\*\*p<0,001.

Quelle: BLIS 2019, eigene Berechnungen.

<sup>10</sup> Weil psychische Erkrankungen oft auch zu längeren Studienzeiten führen, wurde eine weitere Analyse ohne die Variable ‚Studiendauer‘ gerechnet, um auszuschließen, dass mittels dieser möglicherweise zu viel an Varianz kontrolliert wird. Diese Analyse (nicht abgetragen) zeigt keine anderen Ergebnisse; auch geht die bessere Modellanpassung in Modell II nicht alleine auf die Studiendauer zurück.

<sup>11</sup> Die Modellanpassung erfolgt mit dem RReg-Paket über den LR-Test; ein R<sup>2</sup> als Anpassungsmaß wird nicht ausgegeben.

Es besteht die Möglichkeit, dass ein Effekt der sozialen Erwünschtheit durch Störfaktoren verdeckt wird. Darum wird das erste DQ-Modell um Kontrollvariablen erweitert. Im Gegensatz zu DQ-Modell I ist die Modellanpassung nun etwas besser. Nagelkerkes- $R^2$  beträgt 0,133, was zwar noch gering ist, aber eine Verbesserung gegenüber dem Nullmodell darstellt, und auch der LR-Test ist für das DQ-Modell II im Vergleich zum Nullmodell signifikant. Demzufolge kann die abhängige Variable mit dieser Modellschätzung besser vorhergesagt werden als mit dem Nullmodell. Allerdings tragen auch dabei die eigentlich interessierenden Erwünschtheitsvariablen nichts zur Erklärungskraft bei: Sie bleiben beide nicht signifikant und auch die AMEs sind im Vergleich zu den Kontrollvariablen weniger stark ausgeprägt. Von systematischen Effekten der sozialen Erwünschtheit ist also auch in diesem Fall nicht zu sprechen.

Im Gegensatz dazu sind die Effekte der Kontrollvariablen fast alle signifikant. Das Geschlecht hat einen positiven Effekt, was bedeutet, dass Männer weniger von psychischen Erkrankungen betroffen sind. Dieser Effekt ist im Vergleich zu den anderen Kontrolleffekten auch am stärksten ausgeprägt: Im Vergleich zu Frauen und unter sonst gleichen Bedingungen weisen Männer durchschnittlich eine um 11 % niedrigere Wahrscheinlichkeit auf, von einer psychischen Erkrankung betroffen zu sein, oder zumindest dafür, auf entsprechende Hilfsangebote schon einmal zugegriffen zu haben. Zudem ist es auch wahrscheinlicher, von einer psychischen Erkrankung betroffen zu sein, wenn Studierende älter sind oder sich in einem fortgeschrittenen Stadium des Studiums befindet. Auch die Effekte der beiden Stigma-Variablen sind signifikant: Je mehr eine Person das Verbergen einer psychischen Erkrankung befürwortet, bzw. je mehr sie der Aussage zustimmt, dass in der allgemeinen Öffentlichkeit bedeutsame Vorurteile gegenüber psychischen Erkrankungen bestehen, desto eher wird sie auch von einer psychischen Erkrankung betroffen sein.

Die Effekte der Kontrollvariablen stimmen größtenteils mit denjenigen überein, die in der Forschungsliteratur berichtet werden: Bei Frauen gibt es häufiger eine Betroffenheit von psychischen Erkrankungen als bei Männern (Seedat et al. 2009; Bundespsychotherapeutenkammer 2011). Auch das Alter und das Studienjahr haben, wie erwartet, einen positiven Effekt auf die Betroffenheit von psychischen Erkrankungen (Mackenzie et al. 2012; Grobe, Steinmann und Szecsenyi 2018). Etwas unerwartet ist dabei die Effektrichtung der Stigma-Variablen. Anzunehmen wäre gewesen, dass Personen, die auf diesen beiden Variablen hohe Werte aufweisen, auch eher dazu neigen, die Betroffenheit von einer psychischen Erkrankung nicht anzugeben. Der ermittelte negative Zusammenhang bedeutet aber, dass Personen die der Meinung sind, dass es besser sei, psychische Erkrankungen zu verbergen, bzw. dass damit Vorurteile einhergingen, auch eher angeben, von psychischen Erkrankun-

gen betroffen zu sein. Womöglich ist die hier spezifizierte Kausalrichtung jedoch fehlerhaft. Alternativ zur vorliegenden Modellspezifikation wäre vielleicht eher anzunehmen, dass Personen, die von psychischen Erkrankungen betroffen sind, häufiger die Erfahrung machen, dass sie durch ihre Erkrankung benachteiligt oder diskriminiert werden (Ahmedani 2011; Thornicroft et al. 2016), was gesunden Personen eher nicht passieren kann oder auch eher nicht auffällt.

Die Hypothese 2, wonach Effekte der sozialen Erwünschtheit vorliegen, kann mittels dieser Ergebnisse nicht beibehalten werden. Letztendlich lässt die vorliegende Modellanalyse nur folgenden Schluss zu: Psychische Erkrankungen stellen (zumindest für die hier analysierte Stichprobe und die hier benutzte Erhebungsmethodik) keine heikle Thematik dar und Berichte über das Vorliegen von psychischen Erkrankungen werden nicht durch ein sozial erwünschtes Antwortverhalten beeinflusst.

### 4.3 Datenqualität

Weil sich das BLIS-Projekt zum Ziel gesetzt hatte, verschiedene indirekte Fragetechniken zu untersuchen, wurden am Ende des Fragebogens einige Fragen gestellt, die evtl. einen Rückschluss auf die Qualität der mittels Befragung ermittelten Daten zulassen. Dabei interessierte vor allen Dingen, wie konzentriert und ehrlich die Fragen beantwortet wurden, ob die befragten Personen der Zusicherung von Anonymität und Schutz der Persönlichkeit vertrauten und ob sie glaubten, die Befragungsmethoden (insbesondere die komplexeren indirekten Fragemethoden) korrekt befolgt zu haben. Die Antworten wurden mit einer 5er-Skala erhoben. Die Anzahl und die Anteile der befragten Personen, die den genannten Frage-Items zugestimmt haben, werden in Tabelle 3 aufgeführt (siehe Anhang A3).

Tabelle 3: Deskriptive Statistiken zur Datenqualität

	DQ-Daten			CM-Daten		
	Anzahl	Anteil in %	N	Anzahl	Anteil in %	N
Konzentration	435	83,02	524	190	82,61	230
Ehrlichkeit	523	99,81		229	99,57	
Vertrauen	416	79,39		182	79,13	
Befolgung	512	97,71		226	98,26	

Quelle: BLIS 2019, eigene Berechnungen.

Mehr als 82 % der befragten Personen haben demnach die Fragen eher oder sehr konzentriert beantwortet. Etwas mehr als drei Viertel hatten eher ein starkes oder sehr starkes Vertrauen in die zugesicherte Anonymität der Studie. Mehr als 97 % der Befragten haben die Anweisungen im Fragebogen eher oder ganz bestimmt korrekt befolgt. Alle diese Verteilungen lassen darauf schließen, dass die befragten Personen die gestellten Fragen gewissenhaft beantwortet und auch verstanden haben. Und da nahezu alle befragten Personen ehrlich geantwortet haben wollen, ist anzunehmen, dass die ge-

messenen Prävalenzen im DQ- und im CM-Datensatz den ‚wahren‘ Werten der Befragten wohl relativ nahe kommen.<sup>12</sup> Folglich kann davon ausgegangen werden, dass der fehlende Unterschied zwischen den DQ- und CM-Prävalenzen darauf zurückzuführen ist, dass die befragten Personen die Fragen zur psychischen Gesundheit im Kontext der hier durchgeführten Erhebung nicht als heikel wahrgenommen haben und sich demzufolge nicht dazu veranlasst sahen, unehrlich zu antworten.

## 5 Diskussion und Fazit

In den letzten Jahren ist ein deutlicher Anstieg von Untersuchungen zur empirischen Erforschung von heiklen Themen und von indirekten Fragetechniken zu beobachten gewesen. Ein Thema, das dabei bislang wenig Aufmerksamkeit erhielt, ist die Wahrnehmung psychischer Gesundheit durch befragte Personen. Die vorliegende Studie versucht, die psychische Gesundheit als heikles Thema der Umfrageforschung zu untersuchen. Es sollten dabei zwei Fragen beantwortet werden: Zum einen, ob die psychische Gesundheit als ein heikles Thema zu behandeln ist, und zum anderen, ob mit indirekten Fragetechniken höhere Prävalenzen psychischer Gesundheit geschätzt werden.

In dieser Studie konnte nachgewiesen werden, dass das Crosswise Modell keine höheren Prävalenzen schätzt als die direkte Frage. Eine mögliche Begründung dafür ist, dass die befragten Personen der Anonymität des Crosswise Modells nicht trauen, also glauben, dass mit dem Crosswise Modell das individuelle Vorliegen einer psychischen Erkrankung errechnet werden kann und dass nicht nur die Prävalenz auf Aggregatebene geschätzt werden kann, sondern auch die jeweilige Betroffenheit erfasst werden kann, wie es bei der direkten Frage der Fall ist und dass das Crosswise Modell dazu dient, eben diese Möglichkeit ‚zu vertuschen‘.

Dagegen spricht jedoch: Erstens, dass die errechneten Prävalenzen der direkten Befragung den bereits vorgestellten Krankenkassendaten ähnlich sind und bei rund 20 % liegen. Zweitens, dass die Befragten ein hohes Vertrauen in die Befragung angegeben haben. Drittens, dass die Anweisungen zum Crosswise Modell aus anderen Studien übernommen werden konnten und dort bereits sichergestellt wurde, dass die befragten Personen die jeweiligen Anweisungen verstehen (was die Befragten auch hier angegeben haben; Tabelle 3). Und viertens, dass mit Blick auf die bisherige Forschung zum Crosswise Modell der fehlende Unterschied zwischen den Prävalenzen eher ungewöhnlich ist. Denn in den meisten Untersuchungen können damit höhere Prävalenzen als mit der direkten Frage erzielt werden, wobei dabei aber meist andere heikle Themen erforscht wurden.

---

<sup>12</sup> Eine solche Interpretation ist mit Vorsicht zu tätigen, denn die befragten Personen könnten bei der Frage nach der Ehrlichkeit der Antworten auch gelogen haben. Letztendlich ist es an dieser Stelle nicht möglich zu prüfen, ob tatsächlich ehrliche Antworten vorliegen. Allerdings deutet der fehlende Unterschied zwischen den DQ- und den CM-Prävalenzen an, dass dies zumindest für die psychischen Erkrankungen der Fall ist.

Dass es in dieser Studie keinen Unterschied zwischen den Befragungsmethoden gibt, ist daher auf das untersuchte heikle Thema zurückzuführen. Offenbar ist für Studierende in Deutschland eine psychische Erkrankung weniger heikel als angenommen. Ein Ergebnis, das für diese Sichtweise spricht, ist das der multivariaten Analysen. In diesen Analysen konnte kein Effekt der sozialen Erwünschtheit gefunden werden. Die Idee, dass Personen, die eine hohe Normkonformität angeben, auch versuchen, psychische Erkrankungen zu verschweigen, bewahrheitete sich also nicht. So haben die meisten Befragten keine Bedenken, offen über ihre psychischen Erkrankungen zu sprechen, solange die Befragung im Rahmen von anonymen Onlinebefragungen stattfindet. Entgegen der Annahme zeigt sich also, dass die psychische Gesundheit kein Tabuthema darstellt und dass damit von befragten Studierenden offen umgegangen wird. Letztendlich lässt dies nur den folgenden Schluss zu: Die psychische Gesundheit stellt für die empirische Umfrageforschung mittels Onlineerhebung unter Studierenden keine heikle Thematik dar. Ob dies aber auch für die Gesamtbevölkerung oder für Umfragen, die nicht mittels Onlineerhebung durchgeführt werden, gilt, ist mit dieser Studie nicht bestimmbar; dies muss gesondert geprüft werden.

Eine andere Erklärung für die unerwarteten Ergebnisse könnte aber auch mit dem hier benutzten Forschungsdesign zusammenhängen. Es ist zum Beispiel möglich, dass das Aufsuchen einer Beratungsstelle kein guter Proxy für das Vorliegen einer psychischen Erkrankung ist, denn Beratungsstellen an deutschen Universitäten bieten meistens nicht nur Hilfe zu Fragen der psychischen Gesundheit, sondern auch zu anderen persönlichen Problemen wie z. B. zur Organisation des Studiums, zum Zeitmanagement oder zu Stresserfahrungen. Diese Probleme hängen nicht unbedingt mit einer vorliegenden psychischen Erkrankung zusammen und könnten Fragen zur selbstberichteten psychischen Gesundheit etwas von deren ‚Sensitivität‘ genommen haben. Weiterhin könnte auch die hier benutzte Operationalisierung der sozialen Erwünschtheit problematisch sein. Es könnte bei der Messung der sozialen Erwünschtheit zu verzerrten Angaben gekommen sein, die in Richtung einer sozial erwünschten Antwort vom ‚wahren‘ Wert abweichen. Zudem muss bedacht werden, dass keine Daten (Frageantworten) einer Zufallsstichprobe vorliegen. Es könnte der Fall sein, dass nur Personen, die von vornherein offen sind, über heikle Themen zu sprechen, sich entschieden haben, an der Befragung teilzunehmen.

Sollte in Zukunft eine ähnliche Studie durchgeführt werden um zu prüfen, ob die vorliegenden Ergebnisse unabhängig vom Forschungsdesign gültig sind, wäre zu empfehlen, dass die Selbstberichterstattung über das Vorliegen einer psychischen Erkrankung anders operationalisiert wird, dass zur Operationalisierung der sozialen Erwünschtheit eine andere Skala verwendet wird und dass die zu analysierenden Daten aus einer Zufallsstichprobe stammen sollten.

## 6 Literaturverzeichnis

- Ahmedani, Brian K. 2011. Mental health stigma: Society, individuals, and the profession. *Journal of Social Work Values and Ethics* 8 (2): 4-1.
- Antin, Judd und Aaron Shaw. 2012. Social desirability bias and self-reports of motivation: a study of amazon mechanical turk in the US and India. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*.
- Bailer, Josef, Daniela Schwarz, Michael Witthöft, Cornelia Stübinger und Fred Rist. 2008. Prävalenz psychischer Syndrome bei Studierenden einer deutschen Universität. *PPmP. Psychotherapie, Psychosomatik, medizinische Psychologie* 58 (11): 423-429.
- Benton, Sherry A., John M. Robertson, Wen-Chih Tseng, Fred B. Newton und Stephen L. Benton. 2003. Changes in counseling center client problems across 13 years. *Professional Psychology: Research and Practice* 34 (1): 66.
- Bundespsychotherapeutenkammer. 2011. BPtK-Studie zur Arbeitsunfähigkeit. [https://www.bptk.de/wp-content/uploads/2019/01/20110622\\_BPtK-Studie\\_Arbeitsunfa%CC%88higkeit\\_2010.pdf](https://www.bptk.de/wp-content/uploads/2019/01/20110622_BPtK-Studie_Arbeitsunfa%CC%88higkeit_2010.pdf). Zugegriffen: 02. August 2020.
- Corrigan, Patrick. 2004. How stigma interferes with mental health care. *American Psychologist* 59 (7): 614.
- DAK-Gesundheit. 2013. DAK-Gesundheitsreport 2013. <https://dak.de/dak/download/vollstaendiger-bundesweiter-gesundheitsreport-2013-2120160.pdf>. Zugegriffen: 02. August 2020.
- Davis, Christopher G., Jennifer Thake und Natalie Vilhena. 2010. Social desirability biases in self-reported alcohol consumption and harms. *Addictive Behaviors* 35 (4): 302-311.
- Fisher, Robert J. und James E. Katz. 2000. Social-desirability bias and the validity of self-reported values. *Psychology & Marketing* 17 (2): 105-120.
- Goffman, Erving. 1974. *Stigma: Notes on the management of spoiled identity*. New York: J. Aronson.
- Gove, Walter R. und Michael R. Geerken. 1977. Response bias in surveys of mental health: An empirical investigation. *American Journal of Sociology* 82 (6): 1289-1317.
- Grobe, Thomas und Susanne Steinmann. 2015. Gesundheitsreport 2015: Gesundheit von Studierenden. <https://www.tk.de/resource/blob/2034304/47b32c03c8f94de08485134741a5b340/gesundheitsreport-2015-data.pdf>. Zugegriffen: 02. August 2020.
- Grobe, Thomas, Susanne Steinmann und Joachim Szecsenyi. 2018. BARMER-Arztreport 2018. <https://www.barmer.de/blob/144368/08f7b513fdb6f06703c6e9765ee9375f/data/dl-barmer-arztreport-2018.pdf>. Zugegriffen: 02. August 2020.
- Heerwegh, Dirk. 2009. Mode differences between face-to-face and web surveys: an experimental investigation of data quality and social desirability effects. *International Journal of Public Opinion Research* 21 (1): 111-121.
- Hoffmann, Adrian, Berenike Waubert De Puiseau, Alexander F. Schmidt und Jochen Musch. 2017. On the comprehensibility and perceived privacy protection of indirect questioning techniques. *Behavior Research Methods* 49 (4): 1470-1483.
- Hoffmann, Adrian, Birk Diedenhofen, Bruno Verschuere und Jochen Musch. 2015. A strong validation of the crosswise model using experimentally-induced cheating behavior. *Experimental Psychology* 62 (6): 403-414.
- Hoffmann, Adrian und Jochen Musch. 2019. Prejudice against women leaders: Insights from an indirect questioning approach. *Sex Roles* 80 (11): 681-692.

- Höglinger, Marc, Ben Jann und Andreas Diekmann. 2016. Sensitive Questions in Online Surveys: An Experimental Evaluation of Different Implementations of the Randomized Response Technique and the Crosswise Model. *Survey Research Methods* 10 (3): 171-187.
- Holbrook, Allyson L. und Jon A. Krosnick. 2010. Social desirability bias in voter turnout reports: Tests using the item count technique. *Public Opinion Quarterly* 74 (1): 37-67.
- Holm-Hadulla, R. 1994. Psychotherapeutische Beratung und Behandlung von Studierenden im Rahmen einer psychotherapeutischen Beratungsstelle. *PPmP. Psychotherapie, Psychosomatik, medizinische Psychologie* 44 (1): 15-21.
- Holm-Hadulla, Rainer M., Frank-Hagen Hofmann, Michael Sperth und Joachim Funke. 2009. Psychische Beschwerden und Störungen von Studierenden. *Psychotherapeut* 54 (5): 346-356.
- Jann, Ben, Julia Jerke und Ivar Krumpal. 2012. Asking sensitive questions using the crosswise model: an experimental survey measuring plagiarism. *Public Opinion Quarterly* 76 (1): 32-49.
- Jerke, Julia und Ivar Krumpal. 2013. Plagiate in studentischen Arbeiten. *Methoden, Daten, Analysen* 7 (3): 347-368.
- Kemper, Christoph J., Constanze Beierlein, Doreen Bensch, Anastassiya Kovaleva und Beatrice Rammstedt. 2012. Eine Kurzsкала zur Erfassung des Gamma-Faktors sozial erwünschten Antwortverhaltens: Die Kurzsкала Soziale Erwünschtheit-Gamma (KSE-G). *Mannheim: GESIS*.
- Kendra, Matthew S., Jonathan J. Mohr und Jeffrey W. Pollard. 2014. The stigma of having psychological problems: Relations with engagement, working alliance, and depression in psychotherapy. *Psychotherapy* 51 (4): 563-573.
- Korndörfer, Martin, Ivar Krumpal und Stefan C. Schmukle. 2014. Measuring and explaining tax evasion: Improving self-reports using the crosswise model. *Journal of Economic Psychology* 45:18-32.
- Krause, Thomas. 2020. Machine Learning basierte Response Style Identifikation: eine simulationsstatistische Pilotstudie. *Schriftenreihe des Instituts für Sozialwissenschaften (SISS)* 47/2020.
- Krumpal, Ivar. 2013. Determinants of social desirability bias in sensitive surveys: a literature review. *Quality & Quantity* 47 (4): 2025-2047.
- Mackenzie, Corey S., Kristin Reynolds, John Cairney, David L. Streiner und Jitender Sareen. 2012. Disorder-specific mental health service use for mood and anxiety disorders: Associations with age, sex, and psychiatric comorbidity. *Depression and anxiety* 29 (3): 234-242.
- Moshagen, Morten, Jochen Musch und Edgar Erdfelder. 2012. A stochastic lie detector. *Behavior Research Methods* 44 (1): 222-231.
- Paulhus, Delroy L. 1984. Two-component models of socially desirable responding. *Journal of Personality and Social Psychology* 46 (3): 598-609.
- Paulhus, Delroy L. 1991. Measurement and control of response bias. In *Measures of Personality and Social Psychological Attitudes*, Hrsg. John P. Robinson, Phillip R. Shaver und Lawrence S. Wrightsman, 17-59. San Diego: Academic Press.
- Paulhus, Delroy L. 2002. Socially desirable responding: The evolution of a construct. In *The role of constructs in psychological and educational measurement*, Hrsg. Henry I. Braun, Jackson N. Douglas und David E. Wiley, 49-69. Mahwah: Erlbaum.
- Persson, Mikael und Maria Solevid. 2014. Measuring political participation - Testing social desirability bias in a web-survey experiment. *International Journal of Public Opinion Research* 26 (1): 98-112.

- Seedat, Soraya, Kate Margaret Scott, Matthias C. Angermeyer, Patricia Berglund, Evelyn J. Bromet, Traolach S. Brugha, Koen Demyttenaere, Giovanni De Girolamo, Josep Maria Haro und Robert Jin. 2009. Cross-national associations between gender and mental disorders in the World Health Organization World Mental Health Surveys. *Archives of General Psychiatry* 66 (7): 785-795.
- Seliger, Kerstin und Elmar Brähler. 2007. Psychische Gesundheit von Studierenden der Medizin. *Psychotherapeut* 52 (4): 280-286.
- Shamsipour, Mansour, Masoud Yunesian, Akbar Fotouhi, Ben Jann, Afarin Rahimi-Movaghar, Fariba Asghari und Ali Asghar Akhlaghi. 2014. Estimating the prevalence of illicit drug use among students using the crosswise model. *Substance Use & Misuse* 49 (10): 1303-1310.
- Stead, Rebecca, Matthew J Shanahan und Richard WJ Neufeld. 2010. "I'll go to therapy, eventually": Procrastination, stress and mental health. *Personality and Individual Differences* 49 (3): 175-180.
- Stock, Christiane und Alexander Krämer. 2001. Die Gesundheit von Studierenden im Studienverlauf. *Das Gesundheitswesen* 63 (Sup. 1): 56-59.
- Thornicroft, Graham, Nisha Mehta, Sarah Clement, Sara Evans-Lacko, Mary Doherty, Diana Rose, Mirja Koschorke, Rahul Shidhaye, Claire O'Reilly und Claire Henderson. 2016. Evidence for effective interventions to reduce mental-health-related stigma and discrimination. *The Lancet* 387 (10023): 1123-1132.
- Tourangeau, Roger und Ting Yan. 2007. Sensitive questions in surveys. *Psychological Bulletin* 133 (5): 859.
- Urban, Dieter und Jochen Mayerl. 2018. *Angewandte Regressionsanalyse: Theorie, Technik und Praxis*. Wiesbaden: Springer.
- Van de Mortel, Thea F. 2008. Faking it: social desirability response bias in self-report research. *Australian Journal of Advanced Nursing* 25 (4): 40-48.
- Warner, Stanley L. 1965. Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association* 60 (309): 63-69.
- Wolter, Felix. 2012. *Heikle Fragen in Interviews: Eine Validierung der Randomized Response-Technik*. Wiesbaden: Springer.
- Yu, Jun-Wu, Guo-Liang Tian und Man-Lai Tang. 2008. Two new models for survey sampling with sensitive characteristic: design and analysis. *Metrika* 67 (3): 251-263.

## 7 Anhang

### 7.1 Anhang A1: Ergebnisse der Faktorenanalyse

Tabelle 4: KMO- und Bartlett-Test der Faktorenanalyse

KMO- und Bartlett-Test	
Maß der Stichprobeneignung nach Kaiser-Meyer-Olkin.	0,528
	Ungefähres Chi-Quadrat
	149,551
Bartlett-Test auf Sphärizität	df
	6
	Signifikanz nach Bartlett
	0,000

Quelle: BLIS 2019, eigene Berechnungen.

Tabelle 5: Erklärte Gesamtvarianz der Faktorenanalyse

Komponente	Anfängliche Eigenwerte			Summen von quadrierten Faktorladungen für Extraktion		
	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %
1	1,445	36,113	36,113	1,445	36,113	36,113
2	1,146	28,661	64,774	1,146	28,661	64,774
3	0,744	18,601	83,374			
4	0,665	16,626	100,000			

Extraktionsmethode: Hauptkomponentenanalyse.

Quelle: BLIS 2019, eigene Berechnungen.

Tabelle 6: Mustermatrix der Faktorenanalyse

	Komponente	
	1	2
Es ist schon mal vorgekommen, dass ich jemanden ausgenutzt habe.	-0,083	0,769
Auch wenn ich selbst gestresst bin, behandle ich andere immer freundlich und zuvorkommend.	0,804	-0,050
Im Streit bleibe ich stets sachlich und objektiv.	0,823	0,046
Ich habe schon mal Müll einfach in die Landschaft oder auf die Straße geworfen.	0,077	0,811

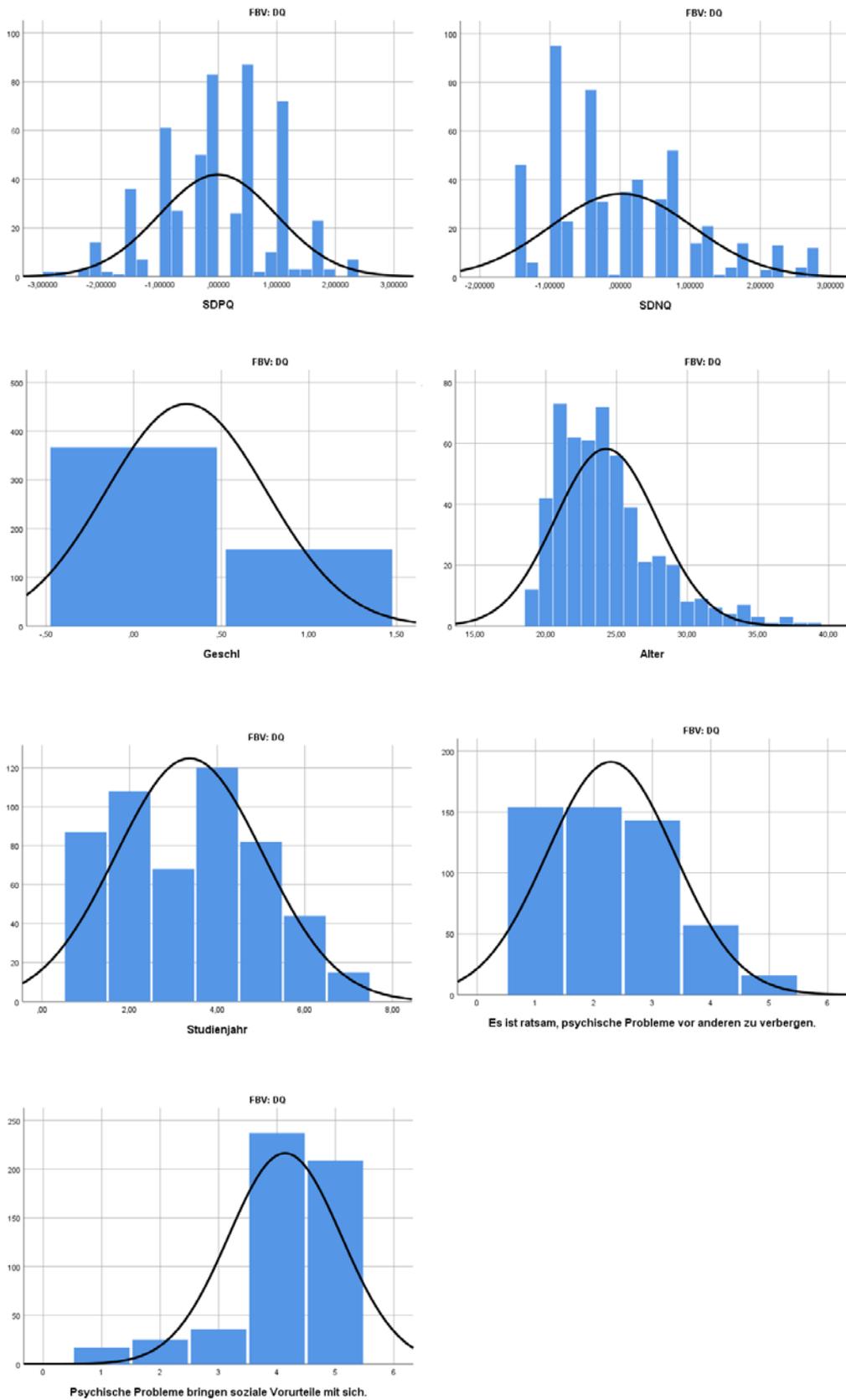
Extraktionsmethode: Hauptkomponentenanalyse.

Rotationsmethode: Oblimin mit Kaiser-Normalisierung.

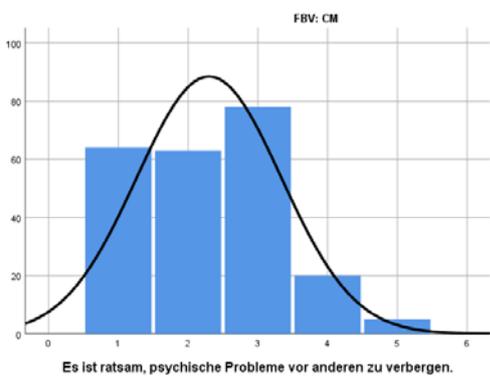
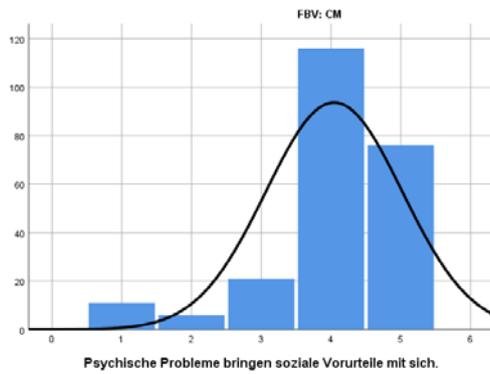
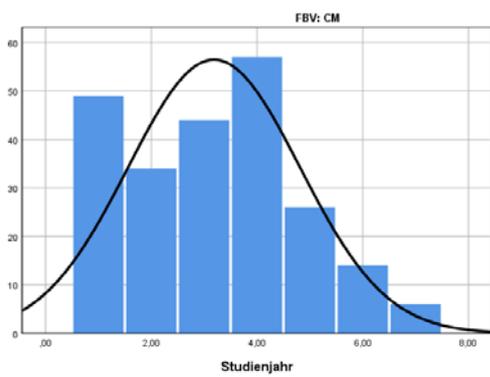
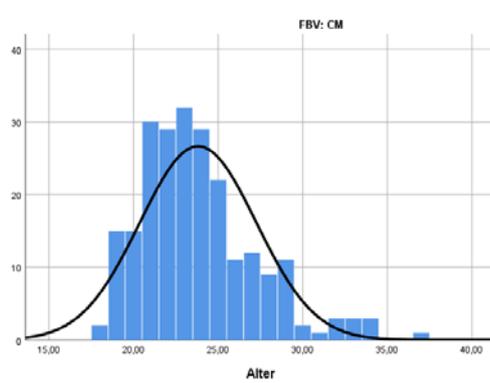
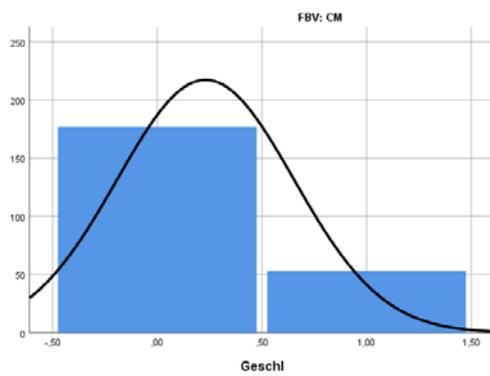
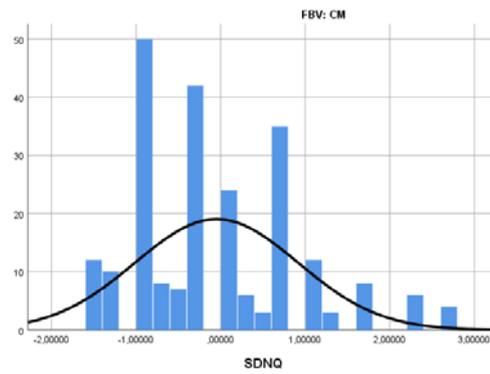
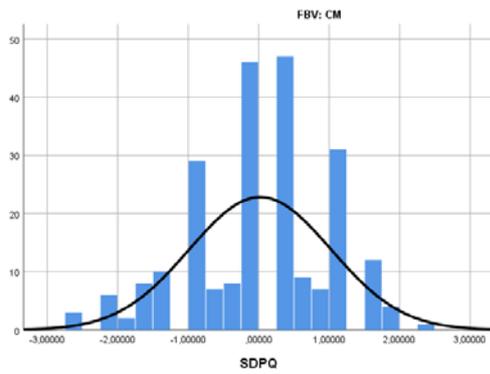
Quelle: BLIS 2019, eigene Berechnungen.

## 7.2 Anhang A2: Deskriptive Ergebnisse

### 7.1.1 Histogramme Crosswise-Modell (DQ)



7.1.2 Histogramme Crosswise-Modell (CM)



## 7.1.3 Deskriptive Statistiken

Tabelle 7: Deskriptive Statistik

FBV		N	Minimum	Maximum	Mittelwert	SD
DQ	SDPQ	524	-2,856	2,372	-0,009	0,999
	SDNQ	524	-1,485	2,774	0,021	1,016
	Geschl	524	0,000	1,000	0,300	0,459
	Alter	524	19,000	39,000	24,258	3,591
	Studienjahr	524	1,000	7,000	3,370	1,675
	PEverbergen	524	1	5	2,29	1,094
	PEvorurteile	524	1	5	4,14	0,966
	Gültige Werte (Listenweise)	524				
CM	SDPQ	230	-2,746	2,372	0,021	1,004
	SDNQ	230	-1,483	2,776	-0,048	0,963
	Geschl	230	0,000	1,000	0,230	0,422
	Alter	230	18,000	37,000	23,813	3,443
	Studienjahr	230	1,000	7,000	3,187	1,625
	PEverbergen	230	1	5	2,30	1,037
	PEvorurteile	230	1	5	4,04	0,979
	Gültige Werte (Listenweise)	230				

Note: FBV: Fragebogenversion; SD: Standardabweichung; SDPQ/SDNQ: Soziale Erwünschtheit, standardisiert; Alter: in Jahren; Studienjahr: in Jahren; PEverbergen/PEvorurteile: Skala von 1 (keine Zustimmung) bis 5 (vollkommene Zustimmung).

Quelle: BLIS 2019, eigene Berechnungen.

Tabelle 8: T-Test: Gruppenstatistiken

	FBV	N	Mittelwert	SD	SEM
SDPQ	DQ	524	-0,009	0,999	0,044
	CM	230	0,021	1,004	0,066
SDNQ	DQ	524	0,021	1,016	0,044
	CM	230	-0,048	0,963	0,064
Alter	DQ	524	24,258	3,591	0,157
	CM	230	23,813	3,443	0,227
Studienjahr	DQ	524	3,370	1,675	0,073
	CM	230	3,187	1,625	0,107
PEverbergen	DQ	524	2,288	1,094	0,048
	CM	230	2,300	1,037	0,068
PEvorurteile	DQ	524	4,137	0,966	0,042
	CM	230	4,043	0,979	0,065

Note: FBV: Fragebogenversion; SD: Standardabweichung; SEM: Standardfehler des Mittelwertes; SDPQ/SDNQ: Soziale Erwünschtheit, standardisiert; Alter: in Jahren; Studienjahr: in Jahren; PEverbergen/PEvorurteile: Skala von 1 (keine Zustimmung) bis 5 (vollkommene Zustimmung).

Quelle: BLIS 2019, eigene Berechnungen.

Tabelle 9: Test bei unabhängigen Stichproben

	Varianzen sind	Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit				
		F	Signifikanz	T	df	Sig. (2-seitig)	95% KI der Differenz	
							Untere	Obere
SDPQ	gleich	0,049	0,824	-0,378	752,000	0,706	-0,185	0,125
	nicht gleich			-0,377	435,042	0,706	-0,186	0,126
SDNQ	gleich	0,605	0,437	0,877	752,000	0,381	-0,086	0,225
	nicht gleich			0,895	459,211	0,371	-0,083	0,222
Alter	gleich	0,222	0,638	1,585	752,000	0,113	-0,106	0,995
	nicht gleich			1,611	454,494	0,108	-0,098	0,987
Studienjahr	gleich	1,909	0,167	1,396	752,000	0,163	-0,075	0,441
	nicht gleich			1,412	449,613	0,159	-0,072	0,438
PEverbergen	gleich	0,749	0,387	-0,139	752,000	0,890	-0,179	0,155
	nicht gleich			-0,142	459,192	0,887	-0,176	0,152
PEvorurteile	gleich	1,024	0,312	1,224	752,000	0,221	-0,057	0,245
	nicht gleich			1,218	431,957	0,224	-0,058	0,246

Note: FBV: Fragebogenversion; SDPQ/SDNQ: Soziale Erwünschtheit, standardisiert; Alter: in Jahren; Studienjahr: in Jahren; PEverbergen/PEvorurteile: Skala von 1 (keine Zustimmung) bis 5 (vollkommene Zustimmung).

Quelle: BLIS 2019, eigene Berechnungen.

Tabelle 10: Kreuztabelle Geschlecht und FBV

Anzahl der Fälle		Kreuztabelle		Gesamt
		FBV		
Geschlecht	0	DQ	CM	
		1	157	53
Gesamt		524	230	754
Exakter Test nach Fisher				0,053 (2-seitig)
				0,030 (1-seitig)

Note: FBV: Fragebogenversion; Geschlecht: 0 = weiblich, 1 = männlich.

Quelle: BLIS 2019, eigene Berechnungen.

### 7.3 Anhang A3: Deskriptive Ergebnisse der Datenqualitäts-Fragen

Tabelle 11: Häufigkeiten Konzentration

		<b>"Wie hoch war Ihre Konzentration während der Befragung?"</b>		
FBV		Absolut	Relativ	Relativ (kumuliert)
DQ	Sehr gering	1	0,2	0,2
	Eher gering	14	2,7	2,9
	Teils/Teils	74	14,1	17,0
	Eher hoch	340	64,9	81,9
	Sehr hoch	95	18,1	100,0
	Gesamt	524	100,0	
CM	Sehr gering	2	0,9	0,9
	Eher gering	8	3,5	4,3
	Teils/Teils	30	13,0	17,4
	Eher hoch	148	64,3	81,7
	Sehr hoch	42	18,3	100,0
	Gesamt	230	100,0	

Note: FBV: Fragebogenversion; DQ: Direkte Befragung; CM: Crosswise-Model-Befragung

Quelle: BLIS 2019, eigene Berechnungen.

Tabelle 12: Häufigkeiten Ehrlichkeit

		<b>"Mit welchem Ausmaß an Ehrlichkeit haben Sie unsere Fragen beantwortet?"</b>		
FBV		Absolut	Relativ	Relativ (kumuliert)
DQ	Überhaupt nicht ehrlich	0	0	0
	Eher nicht ehrlich	0	0	0
	Teils/Teils	1	0,2	0,2
	Eher ehrlich	101	19,3	19,5
	Sehr ehrlich	422	80,5	100,0
	Gesamt	524	100,0	
CM	Überhaupt nicht ehrlich	1	0,4	0,4
	Eher nicht ehrlich	0	0,4	0,4
	Teils/Teils	0	0,4	0,4
	Eher ehrlich	47	20,4	20,9
	Sehr ehrlich	182	79,1	100,0
	Gesamt	230	100,0	

Note: FBV: Fragebogenversion; DQ: Direkte Befragung; CM: Crosswise-Model-Befragung

Quelle: BLIS 2019, eigene Berechnungen.

Tabelle 13: Häufigkeiten Vertrauen

**"Wie stark vertrauen Sie unseren Maßnahmen zur Anonymität und zum Persönlichkeitsschutz der Teilnehmenden in dieser Umfrage?"**

FBV		Absolut	Relativ	Relativ (kumuliert)
DQ	Überhaupt nicht	6	1,1	1,1
	Eher nicht	27	5,2	6,3
	Teils/Teils	75	14,3	20,6
	Eher stark	205	39,1	59,7
	Sehr stark	211	40,3	100,0
	Gesamt	524	100,0	
CM	Überhaupt nicht	7	3,0	3,0
	Eher nicht	12	5,2	8,3
	Teils/Teils	29	12,6	20,9
	Eher stark	96	41,7	62,6
	Sehr stark	86	37,4	100,0
	Gesamt	230	100,0	

Note: FBV: Fragebogenversion; DQ: Direkte Befragung; CM: Crosswise-Model-Befragung  
 Quelle: BLIS 2019, eigene Berechnungen.

Tabelle 14: Häufigkeiten Befolgung

**"Denken Sie, dass Sie die besonderen Befragungsmethoden jeweils korrekt befolgt haben?"**

FBV		Absolut	Relativ	Relativ (kumuliert)
DQ	Überhaupt nicht	0	0,0	0,0
	Eher nicht	0	0,0	0,0
	Teils/Teils	12	2,3	2,3
	Eher ja	167	31,9	34,2
	Ja, ganz bestimmt	345	65,8	100,0
	Gesamt	524	100,0	
CM	Überhaupt nicht	0	0,0	0,0
	Eher nicht	0	0,0	0,0
	Teils/Teils	4	1,7	1,7
	Eher ja	67	29,1	30,9
	Ja, ganz bestimmt	159	69,1	100,0
	Gesamt	230	100,0	

Note: FBV: Fragebogenversion; DQ: Direkte Befragung; CM: Crosswise-Model-Befragung  
 Quelle: BLIS 2019, eigene Berechnungen.

**SISS:****Schriftenreihe des Instituts für Sozialwissenschaften der Universität Stuttgart**<http://www.uni-stuttgart.de/soz/institut/forschung/siss.html>

- No. 49/2020 Psychische Gesundheit: Eine heikle Thematik in der empirischen Umfrageforschung? Zur Validierung des Crosswise Modells.  
Anna Götze, Andreas Wahl
- No. 48/2020 Analyse und Schätzung von Mehrgruppen-Strukturgleichungsmodellen mittels SPSS und EQS. Eine praxisnahe Anleitung.  
Andreas Wahl, Dieter Urban
- No. 47/2020 Machine Learning basierte Response Style Identifikation: eine simulations-statistische Pilotstudie.  
Thomas Krause
- No. 46/2019 Informationskriterien zur latenten Klassenidentifikation für Multiple Indicator Growth Mixture Models.  
Thomas Krause
- No. 45/2017 Die Angst vor dem Fremden. Ein Vignettenexperiment zur Bedrohungswahrnehmung. Ergebnisse eines studentischen Projektseminars.  
Alexander Geisler, Alexandra Möck, Rebecca Nell, Svenja Wiechmann, Thomas Krause
- No. 44/2016 Regressionsanalyse bei fehlenden Variablenwerten (missing values): Imputation oder Nicht-Imputation? Eine Anleitung für die Regressionspraxis mit SPSS.  
Dieter Urban, Jochen Mayerl, Andreas Wahl
- No. 43/2014 Meta-Analyse mit Mehrebenenmodellen unter Verwendung von HLM.  
Joachim Fiebig, Dieter Urban
- No. 42/2013 Panelanalyse mit Mehrebenenmodellen. Eine anwendungsorientierte Einführung.  
Thomas Krause, Dieter Urban
- No. 41/2012 Meta-Analyse und "publication bias". Eine empirisch-statistische Fallstudie.  
Joachim Fiebig, Gerrit Rüdebusch, Dieter Urban
- No. 40/2011 Pädosexuelle Viktimisierung und pädosexuelle Straffälligkeit. Ein „harter“ empirischer Hypothesentest.  
Dieter Urban, Joachim Fiebig
- No. 39/2010 Binär-logistische Regressionsanalyse.  
Grundlagen und Anwendung für Sozialwissenschaftler.  
Jochen Mayerl, Dieter Urban
- No. 38/2010 Das Working-Poor-Problem in Deutschland.  
Empirische Analysen zu den Ursachen von Armut trotz Erwerbstätigkeit.  
Leonie Hellmuth, Dieter Urban
- No. 37/2010 Der Bystander-Effekt in alltäglichen Hilfsituationen: Ein nicht-reaktives Feldexperiment.  
Katrin Alle, Jochen Mayerl
- No. 36/2008 Berufseintritt und Berufssituation von Soziologieabsolventen der Universität Stuttgart.  
Deskriptive Ergebnisse einer Absolventenbefragung aus dem Jahr 2007.  
Jochen Mayerl, Dieter Urban
- No. 35/2007 Antwortlatenzzeiten in TRA-Modellen. Zur statistischen Erklärung von (Geld)-Spendenverhalten.  
Dieter Urban, Jochen Mayerl

**SISS:****Schriftenreihe des Instituts für Sozialwissenschaften der Universität Stuttgart**<http://www.uni-stuttgart.de/soz/institut/forschung/siss.html>

- No. 34/2006 Zur soziologischen Erklärung individuellen Geldspendens. Eine Anwendung und Erweiterung der Theory of Reasoned Action unter Verwendung von Antwortlatenzzeiten in einem Mediator-Moderator-Design.  
Jochen Mayerl
- No. 33/2006 Ist Bildung gleich Bildung? Der Einfluss von Schulbildung auf ausländerablehnende Einstellungen in verschiedenen Alterskohorten.  
Klaus Hadwiger
- No. 32/2005 Analyzing cognitive processes in CATI-Surveys with response latencies: An empirical evaluation of the consequences of using different baseline speed measures.  
Jochen Mayerl, Piet Sellke, Dieter Urban
- No. 31/2005 Persönliche Netzwerke in der Systemtheorie.  
Jan A. Fuhse
- No. 30/2004 Erhöht ein Opfer-Täter-Zyklus das Risiko, Sexualstraftaten als pädosexuelle Straftaten zu begehen? Ergebnisse einer ereignisanalytischen Pilotstudie.  
Dieter Urban, Heiko Lindhorst
- No. 29/2003 Können Nonattitudes durch die Messung von Antwortreaktionszeiten ermittelt werden? Eine empirische Analyse computergestützter Telefoninterviews.  
Jochen Mayerl
- No. 28/2003 Systeme, Netzwerke, Identitäten. Die Konstitution sozialer Grenzziehungen am Beispiel amerikanischer Straßengangs.  
Jan A. Fuhse
- No. 27/2002 Probleme bei der Messung individueller Veränderungsraten. 13 empirisch und methodisch induzierte Effekte, die es schwierig machen, Veränderungen von generalisierten Bewertungen zu ermitteln.  
Dieter Urban
- No. 26/2002 Vertrauen und Risikoakzeptanz. Zur Relevanz von Vertrauen bei der Bewertung neuer Technologien.  
Martin Slaby, Dieter Urban
- No. 25/2002 Risikoakzeptanz als individuelle Entscheidung. Zur Integration der Risikoanalyse in die nutzentheoretische Entscheidungs- und Einstellungsforschung.  
Martin Slaby, Dieter Urban
- No. 24/2001 Ist Glück ein affektiver Sozialindikator subjektiven Wohlbefindens? Dimensionen des subjektiven Wohlbefindens und die Differenz zwischen Glück und Zufriedenheit.  
Jochen Mayerl
- No. 23/2001 Religiosität und Profession. Longitudinale Analysen zur Entwicklung des religiösen Selbstbildes bei Erzieherinnen.  
Heiko Lindhorst
- No. 22/2001 Differentielle Technikakzeptanz, oder: Nicht immer führt die Ablehnung einer Technik auch zur Ablehnung ihrer Anwendungen. Eine nutzentheoretische und modell-statistische Analyse.  
Martin Slaby, Dieter Urban
- No. 21/2001 Unser „wir“ - ein systemtheoretisches Modell von Gruppenidentitäten.  
Jan A. Fuhse

**SISS:****Schriftenreihe des Instituts für Sozialwissenschaften der Universität Stuttgart**<http://www.uni-stuttgart.de/soz/institut/forschung/siss.html>

- No. 20/2000 Längsschnittanalysen mit latenten Wachstumskurvenmodellen in der politischen Sozialisationsforschung.  
Dieter Urban
- No. 19/2000 Die Zufriedenheit von Stuttgarter Studierenden mit ihrer Lebens- und Wohnsituation. Erste deskriptive Ergebnisse einer sozialwissenschaftlichen Studie zu allgemeinen und bereichsspezifischen Zufriedenheiten der Studierenden des Campus Vaihingen und des Campus Hohenheim.  
Projektgruppe Campus: Slaby, M.; Grund, R.; Mayerl, J.; Noak, T.; Payk, B.; Sellke, P.; Urban, D.; Zudrell, I.
- No. 18/1999 Role Models and Trust in Socio-Political Institutions: A Case Study in Eastern Germany, 1992-96.  
Joachim Singelmann, Toby A. Ten Ayck, Dieter Urban
- No. 17/1998 Zur Interaktion zwischen Befragten und Erhebungsinstrument. Eine Untersuchung zur Konstanz des Meinungsurteils von Befragten im Interviewverlauf.  
Martin Slaby
- No. 16/1998 Technikeinstellungen: gibt es die überhaupt? Ergebnisse einer Längsschnittanalyse von Bewertungen der Gentechnik.  
Dieter Urban
- No. 15/1998 Bewertende Einstellungen zur Gentechnik: ihre Form, ihre Inhalte und ihre Dynamik. Kurzbericht zu Ergebnissen des Forschungsprojektes „Einstellungen zur Gentechnik“.  
Dieter Urban, Uwe Pfenning, Joachim Allhoff
- No. 14/1997 Politische Sozialisation im Transformationsprozeß: Die Entwicklung demokratiebezogener Einstellungen von ostdeutschen Jugendlichen und deren Eltern zwischen 1992 und 1996.  
Barbara Schmidt, Dieter Urban, Joachim Singelmann
- No. 13/1997 Die soziale Vererbung von Ausländer“feindlichkeit“. Eine empirische Längsschnittanalyse der intra- und intergenerativen Transmission von sozialen Einstellungen.  
Dieter Urban und Joachim Singelmann
- No. 12/1997 Gentechnik: „Fluch oder Segen“ versus „Fluch und Segen“. Bilanzierende und differenzierende Bewertungen der Gentechnik in der öffentlichen Meinung.  
Dieter Urban und Uwe Pfenning
- No. 11/1996 Einstellungsmessung oder Einstellungsgenerierung? Die Bedeutung der informationellen Basis bei Befragten für die empirische Rekonstruktion von Einstellungen zu gentechnischen Anwendungen.  
Martin Slaby
- No. 10/1996 Jugend und Politik im Transformationsprozeß. Eine Fallstudie zur Stabilität und Veränderung von politischen Einstellungen bei ostdeutschen Jugendlichen zwischen 1992 und 1995.  
Dieter Urban, Joachim Singelmann, Helmut Schröder
- No. 09/1995 Social Indicators in a Nonmetropolitan County: Testing the Representativeness of a Regional Nonrandom Survey in Eastern Germany.  
Dieter Urban, Joachim Singelmann

**SISS:****Schriftenreihe des Instituts für Sozialwissenschaften der Universität Stuttgart**<http://www.uni-stuttgart.de/soz/institut/forschung/siss.html>

- No. 08/1995 Handbuch zur empirischen Erhebung von Einstellungen/Kognitionen zur Bio- und Gentechnologie (inklusive Diskette)  
(zweite, überarbeitete und erweiterte Auflage)  
Uwe Pfenning, Dieter Urban, Volker Weiss
- No. 07/1995 Latent Attitude Structures Directing the Perception of New Technologies.  
An Application of SEM-Methodology to the Construction of Attitude Measurement Models  
Related to Technologies of Prenatal Genetic Engineering and Testing.  
Dieter Urban
- No. 06/1995 Regionale Strukturprobleme.  
Sammelband der Beiträge zum Symposium vom 13. und 14. Oktober 1994.  
Frank C. Englmann (Hrsg.)
- No. 05/1995 Die Bedeutung Neuronaler Netze in der Ökonomie.  
Hermann Schnabl
- No. 04/1994 Rechtstatsachen zum Unternehmens- und Gesellschaftsrecht aus den neuen  
Bundesländern.  
Udo Kornblum
- No. 03/1994 Neue württembergische Rechtstatsachen zum Unternehmens- und  
Gesellschaftsrecht.  
Udo Kornblum
- No. 02/1994 Report on the German Machine Tool Industry.  
Frank C. Englmann, Christian Heyd, Daniel Köstler, Peter Paustian  
with the assistance of Susanne Baur and Peter Bergmann
- No. 01/1994 "Vertrauen" - soziologisch betrachtet. Ein Beitrag zur Analyse binärer  
Interaktionssysteme.  
Peter Antfang, Dieter Urban