

Einflussfaktoren auf zuverlässiges Verhalten am Beispiel
der Teilnahme an psychologischen Studien

Masterarbeit

zur Erlangung des akademischen Grades
Master of Science
der Studienrichtung Psychologie
an der Karl-Franzens-Universität Graz

vorgelegt von

Pierre PACHER, BSc

Begutachter: Em.Univ.-Prof. Dr.rer.nat. Dietrich Albert
Institut für Psychologie

Graz, April 2022

Danksagung

Zuallererst möchte ich mich herzlichst bei Em.Univ.-Prof. Dr.rer.nat. Dietrich Albert für die umfassende Unterstützung, die vielen kritischen Anregungen und spannenden Diskussionen bedanken, die diese Arbeit stetig verbessert haben.

Ich bedanke mich außerdem bei Prof. Dr. Hans-Werner Bierhoff und bei Dr. Lars Satow, deren Tests ich im Zuge der Creative Commons Lizenz nutzen durfte und die mir dadurch eine spannende Forschung ermöglicht haben.

Außerdem möchte ich mich herzlichst bei Frau Dr. Birgit Schöffmann-Petri bedanken, die mich während des Arbeitsprozesses immer wieder durch anregende Diskussionen fachlich unterstützt hat und mir allzeit mit kompetenten Ratschlägen zur Verbesserung dieser Arbeit zur Seite stand.

Ebenso bedanke ich mich bei meiner Freundin, Frau Mag. Hannah Schöffmann, MA für das wiederholte, geduldige und genaue Korrekturlesen sowie die ausgiebige mentale Unterstützung über die letzten Jahre.

Ein großer Dank gebührt auch meiner Familie, ohne deren Unterstützung es mir nicht möglich gewesen wäre, zu studieren und die mich immer wieder aufs Neue motiviert und mir Kraft gegeben haben.

Zum Schluss möchte ich mich auch noch bei all jenen bedanken, die diese Arbeit Korrektur gelesen haben und die mir immer wieder mit Rat zur Seite gestanden sind.

Abstract

An efficient approach is elementary while performing scientific research. The (non-)attendance of participants composes a major issue in efficient research. Participants not attending arranged appointments for testing can lead to resources being unused, wasted and/or a satisfying sample size not being reached. Due to a lack of research being conducted on that issue, two surveys were designed to broadly identify possible factors, which influence attendance rate. The smaller preliminary survey served as basic evaluation of non-attendance of participants in psychological scientific research and to identify possible factors of influence on the non-attendance rate. Based on those findings another more extensive survey was conducted, which directly measured non-attendance, while also asking participants, if they did not attend any scientific research appointment in the past. Participants who did not attend at least one appointment were then compared to those who never before showed such behaviour. Additionally, all participants were asked how their probability of attending any research appointment would change under certain circumstances. Results show that direct contact with potential participants, monetary incentives for participation, interest in the research topic, less time required for participating in the study, reminders of the arranged appointment and free choice of the appointment help to prevent non-attendance of participants. Furthermore, research designs which require more than one appointment, as well as appointments before 8:30 AM or after 6:30 PM and studies which require more than 90 minutes of participation should be avoided. The non-attendance rate presents itself as an adaptable criterion, which is determined by many factors and can still be influenced during the conduction of a study. Concluding this thesis, a recommendation on how researchers can keep non-attendance rates low in scientific research can be found.

Kurzzusammenfassung

Eine effiziente Herangehensweise ist bei der Durchführung von wissenschaftlichen Untersuchungen von elementarer Bedeutung. Dabei spielt der Ausfall von Versuchspersonen eine große Rolle. Wenn Versuchspersonen trotz Anmeldung zu einer Untersuchung nicht zum vereinbarten Termin erscheinen, kann das dazu führen, dass Ressourcen ungenutzt bleiben, verschwendet werden und/oder keine zufriedenstellende Stichprobengröße erreicht werden kann. Da es kaum Studien gibt, die sich dem Ausfall von Versuchspersonen widmen, wurden für diese Arbeit zwei Untersuchungen durchgeführt, die mögliche Einflussfaktoren auf die Ausfallquote weitläufig erfasst haben. Die kleinere Voruntersuchung diente der grundlegenden Einschätzung des Ausmaßes in dem Versuchspersonen bei psychologischen, wissenschaftlichen Untersuchungen ausfallen sowie dem Identifizieren von möglichen Einflussfaktoren auf die Ausfallquote. Bei der darauf aufbauenden, größer angelegten Hauptuntersuchung wurde die Ausfallquote sowohl direkt erhoben als auch eine Befragung über Ausfälle von Versuchspersonen in der Vergangenheit durchgeführt. Versuchspersonen, die bereits bei mindestens einer Untersuchung ausgefallen waren, wurden dann mit jenen verglichen, die noch nie bei einer Untersuchung ausgefallen waren. Zusätzlich wurden alle Versuchspersonen dazu befragt, wie sich unter bestimmten Bedingungen die Wahrscheinlichkeit ändert, dass sie einen zu einer wissenschaftlichen Untersuchung vereinbarten Termin auch tatsächlich wahrnehmen. Die Ergebnisse zeigen, dass der direkte Austausch mit potentiellen Versuchspersonen, monetäre Anreize für die Teilnahme, Interesse am Thema der Untersuchung, eine kürzere Untersuchungsdauer, Erinnerungen an den vereinbarten Termin und eine freie Wahl des Untersuchungstermins helfen, dem Ausfall von Versuchspersonen vorzubeugen. Außerdem sollten Forschungsdesigns, die mehrere Untersuchungstermine benötigen sowie Untersuchungstermine vor 8:30 Uhr bzw. nach 18:30 Uhr und Untersuchungen, die mehr als 90 Minuten dauern, vermieden werden. Die Ausfallquote präsentiert sich als variables Maß, auf das noch während der Durchführung einer Untersuchung Einfluss genommen werden kann und das durch viele Faktoren bestimmt wird. Eine Empfehlung, wie Forschende die Ausfallquote bei wissenschaftlichen Untersuchungen möglichst gering halten können, findet sich am Ende der Arbeit.

Inhaltsverzeichnis

Danksagung.....	II
Abstract	III
Kurzzusammenfassung.....	IV
Inhaltsverzeichnis.....	V
Einleitung	1
Stand der Forschung.....	3
Grundlagen für effiziente Forschung	3
Mögliche Einflussfaktoren auf die Ausfallquote.....	7
Voruntersuchung	23
Ziele der Voruntersuchung.....	23
Methoden der Voruntersuchung.....	24
Aufbau und Ablauf der Untersuchung	24
Stichprobe.....	25
Operationalisierung der Konstrukte	26
Hypothesen der Voruntersuchung	28
Ergebnisse der Voruntersuchung.....	30
Statistische Auswertung	31
Qualitative Auswertung.....	37
Diskussion der Ergebnisse der Voruntersuchung.....	39
Ergänzung: Auffällige Sonderfälle.....	41
Probleme bei der Datenerhebung der Voruntersuchung	44
Implikationen für die Hauptuntersuchung.....	45
Hauptuntersuchung.....	47
Methoden Hauptuntersuchung	47
Aufbau und Ablauf der Untersuchung	48

Stichprobe.....	50
Operationalisierung der Konstrukte	51
Hypothesen der Hauptuntersuchung	54
Ergebnisse Hauptuntersuchung	62
Überprüfung der Homogenität der Stichprobe.....	62
Ausfall von Versuchspersonen.....	63
Auswertung der Teilnahmewahrscheinlichkeiten	63
Testung der Hypothesen.....	64
Gründe für den Ausfall von Versuchspersonen	82
Diskussion	83
Einschränkungen dieser Studie	94
Empfehlung für zukünftige Forschung	98
Quellenverzeichnis	VII
Anhang	XVII
Dokumentation der Literaturrecherche	XVII
Ergänzende Daten zur Hauptuntersuchung	XVIII
Häufigkeitstabelle zur Teilnahmewahrscheinlichkeit	XVIII
Gestaltung der Voruntersuchung.....	XXII
Einleitung und Datenschutzerklärung	XXII
Fragebogen der Voruntersuchung	XXIII
Gestaltung der Hauptuntersuchung	XXVIII
Einleitung und Datenschutzerklärung	XXVIII
Erhebung demografischer Daten bei der Hauptuntersuchung.....	XXIX
Erhebung der Gründe für die Teilnahme und wie Versuchspersonen auf die Untersuchung aufmerksam wurden bei der Hauptuntersuchung.....	XXXI
Fragen zum Ausfall von Versuchspersonen bei der Hauptuntersuchung	XXXIV
Fragebogen zu den Teilnahmewahrscheinlichkeiten bei der Hauptuntersuchung	XXXV

Einleitung

Die Durchführung wissenschaftlicher Untersuchungen gehört zu den Grundlagen der experimentellen Psychologie. Für Forscher*innen ist es wichtig, diese Untersuchungen möglichst effizient durchführen zu können. Um das zu gewährleisten, muss bereits die Planung der Untersuchung darauf abzielen, in der Folge einen reibungslosen Ablauf der Untersuchung zu ermöglichen, eine gute Qualität der Daten sicherzustellen und ausreichend Personen zu einer Teilnahme zu bewegen. Bei Untersuchungen, die es erfordern, dass Personen zu einem vereinbarten Termin an einem bestimmten Ort erscheinen, stellt insbesondere der Ausfall von Versuchspersonen eine Beeinträchtigung von effizienter Forschung dar. Von Forschenden wird oft beklagt, dass nur zu wenigen Untersuchungen alle angemeldeten Proband*innen auch tatsächlich erscheinen; bei manchen Untersuchungen fallen sogar so viele Versuchspersonen aus, dass dies für die Forschenden ein erhebliches Problem darstellt.

Die Entscheidungen, die während der Planung einer wissenschaftlichen Untersuchung getroffen werden, können erheblichen Einfluss auf die Rekrutierung von Versuchspersonen und somit auf eine effiziente Durchführung der Untersuchung nehmen. Zur Rekrutierung von Versuchspersonen für wissenschaftliche Untersuchungen gibt es bereits einige Studien, die sich unter anderem damit beschäftigen, die Effizienz von postalischen Befragungen anhand der Rücksendequote zu verbessern (u.a. Baruch, 1999; Brennan & Hoek, 1992), die Teilnahmebereitschaft von Menschen an Telefonumfragen zu erhöhen (u.a. Maynard & Hollander, 2014; Schnauber & Daschmann, 2008) oder die Abbruchquote von Online-Befragungen zu reduzieren (u.a. Fang, et al., 2012; Liu, et al., 2015). Der Ausfall von Versuchspersonen bei psychologischen Untersuchungen, die das Erscheinen der Versuchspersonen am Versuchsort erfordern, ist jedoch noch kaum erforscht. Fienup und Dixon (2003) zählen zu den wenigen Forscher*innen, die sich mit den Ausfallquoten bei solchen Studien befasst haben. Sie untersuchten, ob eine telefonische Erinnerung an eine bevorstehende Versuchsteilnahme vor Ort Einfluss auf das tatsächliche Erscheinen zum Versuchstermin nimmt. Dabei konnten sie feststellen, dass Versuchspersonen eher zum Untersuchungstermin erscheinen, wenn sie am Tag davor direkt per Telefon an den Termin erinnert werden, als wenn ihnen derselbe Text zur Erinnerung auf den Anrufbeantworter gesprochen wird bzw. sie gar nicht erinnert werden.

Die Studie von Fienup und Dixon zeigt klar, dass Forscher*innen bei der Planung von psychologischen Untersuchungen auf zuverlässiges Verhalten und damit auf die Ausfallquote von Versuchspersonen Einfluss nehmen können.

Ziel der vorliegenden Masterarbeit ist es, mögliche Einflussfaktoren auf die Ausfallquote von Versuchspersonen bei wissenschaftlichen Untersuchungen, welche die persönliche Anwesenheit erfordern, zu finden. Dabei wird immer auch hinterfragt, auf welche Art und Weise die jeweiligen Einflussfaktoren auf die Ausfallquote einwirken. Aufgrund der Bedeutung der Ausfallquote für eine effiziente Forschung und des Mangels an wissenschaftlichen Erkenntnissen zu diesem Themenkomplex bedarf es einer breiten, grundlegenden Aufarbeitung des Problems. Deshalb liegt dieser Arbeit folgende explorative Fragestellung zugrunde:

„Welche Einflussfaktoren auf die Ausfallquote von Versuchspersonen gibt es und wie wirken sie auf zuverlässiges Verhalten von potenziellen Versuchspersonen?“

Um daraus auch Empfehlungen für zukünftige Forschung ableiten zu können, ergibt sich damit zusammenhängend zusätzlich folgende Fragestellung:

*„Wie können Versuchsleiter*innen die Zuverlässigkeit von Versuchspersonen in Bezug auf die Teilnahme an psychologischen Untersuchungen, die die persönliche Teilnahme erfordern, verbessern?“*

Da es, wie bereits erwähnt, noch wenige Erkenntnisse auf diesem Forschungsgebiet gibt, wird in der Folge der Stand der Forschung aus Forschungsfeldern aufgearbeitet, die entweder Berührungspunkte mit dem Ausfall von Versuchspersonen haben oder für die damit einhergehende Forschung relevant sind. Daraus werden erste Hypothesen zur Wirkung von möglichen Einflussfaktoren auf den Ausfall von Versuchspersonen abgeleitet und formuliert. Im Rahmen einer Voruntersuchung wurden Personen zu jeweils einer wissenschaftlichen Untersuchung befragt, die sie an der Karl-Franzens-Universität als Versuchsleiter*in begleitet haben. Die dabei gesammelten Daten sollen abbilden, wie Untersuchungen an dieser Universität geplant werden, inwiefern der Ausfall von Versuchspersonen ein Problem darstellt und was Gründe für diese Ausfälle sein könnten. Anhand der im Kapitel „Stand der Forschung“ besprochenen Erkenntnisse und der Ergebnisse dieser Voruntersuchung wurde eine größere Hauptuntersuchung geplant und durchgeführt. Die Ergebnisse der Hauptuntersuchung und deren Diskussion bilden die Grundlage für die Empfehlung für künftige Forschung am Ende dieser Arbeit, die helfen soll, die Ausfallquote bei zukünftigen Studien zu reduzieren.

Stand der Forschung

Als Basis für die folgenden Ausführungen werden zunächst die Grundlagen effizienter Forschung besprochen. Effiziente Forschung wird in der vorliegenden Arbeit als ein Forschungsprozess definiert, bei dem die vorhandenen Ressourcen wie Zeit, Geld, Forschungspersonal, Örtlichkeiten und Gerätschaften möglichst optimal genutzt sowie eine zufriedenstellende Datenqualität und Stichprobengröße erreicht werden. Aufbauend auf diese Grundlagen werden dann mögliche Einflussfaktoren auf die Ausfallquote anhand verwandter Forschungsgebiete identifiziert und Hypothesen zu deren Wirkung aufgestellt.

Grundlagen für effiziente Forschung

In der experimentellen Psychologie gibt es klare Standards, an die sich Forscher*innen halten sollten, um qualitativ hochwertige Forschungsergebnisse zu erzielen und somit den Grundstein für eine effiziente Forschung zu legen. In der Grundlagenliteratur zur experimentellen Psychologie (z.B. Reiß & Sarris, 2012; Beller, 2016) wird nicht nur die standardisierte Auswertung von Daten ausführlich behandelt, sondern bereits der Planung und Durchführung von psychologischen Experimenten ein erheblicher Stellenwert beigemessen. Wie einleitend angemerkt, können gerade in der experimentellen Psychologie Unregelmäßigkeiten in Planung und Durchführung von Untersuchungen erhebliche Folgen nach sich ziehen.

In der Grundlagenliteratur zur experimentellen Psychologie gibt es ausführliche Beschreibungen davon, wie die Instruktion bei wissenschaftlichen Untersuchungen gestaltet werden sollte, um objektive und standardisierte Ergebnisse zu erhalten (Reiß & Sarris, 2012, S. 110 ff.). Gleiches gilt auch für den Raum, in dem eine Untersuchung durchgeführt wird und die Zeit, zu der diese stattfindet. Beispielsweise können Versuchsaapparaturen in einem Raum so angeordnet sein, dass einige der Versuchspersonen abgelenkt werden und sich nicht so gut konzentrieren können wie andere, wodurch die Vergleichbarkeit der Untersuchungsergebnisse leidet (Reiß & Sarris, 2012, S. 122). Ebenso kann sich ein sehr früher oder sehr später Termin auf die Leistung einzelner Versuchspersonen auswirken (Borella, Ludwig, Dirk, de Ribaupierre, 2011; Lewandowska et al. 2018). Diese und viele weitere Faktoren, wie etwa die Gestaltung der bei der Untersuchung verwendeten Erhebungsmethoden (Hardigan, Popovici & Carvajal, 2016; Sanchez, 1992) oder die Reihenfolge, in der die Testmaterialien dargeboten werden (Faulkner & Cogan, 1990; Galesic & Bosnjak, 2009; Scandell et al., 2003), bilden als Rahmenbedingungen die Grundlage für eine qualitativ hochwertige Forschung.

Die Grundlagenliteratur befasst sich im Zusammenhang mit der Rekrutierung vor allem mit der Zusammensetzung und der optimalen Größe der Stichproben (Reiß & Sarris, 2012; Beller, 2016). Diese sind für die Qualität und die Aussagekraft der Forschungsergebnisse durchaus bedeutsame Faktoren, der Rekrutierung von Versuchspersonen wird jedoch meist weniger Wichtigkeit beigemessen, obwohl diese einen ebenso großen Anteil an der Effizienz der Forschung hat. Es wird also oft besprochen wie eine Stichprobe aussehen soll, jedoch nicht, wie man zu eben dieser Stichprobe gelangt. Dabei können auch die Ausschöpfungsquote und die Ausfallquote die Qualität der Daten und die Stichprobengröße beeinflussen. Die „Ausschöpfungsquote“ bzw. „Rücklaufquote“ (im Englischen „response rate“) bezeichnet dabei das Verhältnis der insgesamt kontaktierten Personen zu jenen Personen, die an einer Untersuchung dann auch tatsächlich teilnehmen, (Koch, Peter & Müller, 2019), während die „Ausfallquote“ jenen Anteil an Personen beschreibt, die sich zwar zu der jeweiligen Untersuchung anmelden, dann jedoch nicht an dieser teilnehmen. Für Forschende bedeuten sowohl eine niedrige Ausschöpfungsquote als auch eine hohe Ausfallquote einen Mehraufwand, weil in diesem Fall die Rekrutierungsmaßnahmen an mehr Personen gerichtet werden müssen, um die gewünschte Stichprobengröße zu erreichen.

Eine hohe Datenqualität, eine hohe Ausschöpfungsquote und eine niedrige Ausfallquote bilden somit zusammen die Basis für effiziente Forschung. Diese Arbeit legt ihren Fokus auf den Forschungsprozess und damit auf diese Kriterien und vernachlässigt dabei andere Aspekte wie beispielsweise die Wirkung (im Englischen: „impact“) die eine Studie nach Veröffentlichung erzielt (Hinrichs-Krapels & Grant, 2016).

Dabei hat die Gestaltung des Rekrutierungsprozesses Einfluss auf die allgemeine Teilnahmebereitschaft und somit auch darauf, welche Versuchspersonen sich zur Teilnahme an einer Untersuchung bereiterklären (Koch, Peter & Müller, 2019). Es gibt unterschiedlichste Gründe, warum Versuchspersonen sich entschließen, an einer Untersuchung teilzunehmen. So entscheidet „[i]m Moment der Kontaktaufnahme [...] ein Zusammenspiel aus überdauernden Persönlichkeitseigenschaften der kontaktierten Person, Situationsmerkmalen sowie der intrinsischen und extrinsischen Motivation der potenziellen Probanden über die Bereitschaft zur Teilnahme.“ (Koch, Peter & Müller, 2019, S. 128). Die entscheidenden Gründe zur Teilnahme an einer Untersuchung können also beispielsweise persönliche Motivationsfaktoren wie Interesse und Hilfsbereitschaft, der Wunsch dem bzw. der Versuchsleiter*in einen Gefallen zu tun, Belohnungen für die Teilnahme (Koch, Peter & Müller, 2019, S. 128) oder schlichtweg die Tatsache, dass die Teilnahme eine Voraussetzung für universitäre Abschlüsse ist, – wie beispielsweise an

der Karl-Franzens-Universität Graz für Psychologiestudierende (Universität Graz Institut für Psychologie, n.d.) – sein.

Einerseits können eine niedrige Teilnahmebereitschaft und somit auch eine niedrige Ausschöpfungsquote dazu führen, dass die gewünschte Stichprobengröße nicht erreicht wird. Andererseits kann dies aber auch dazu führen, dass die Zusammensetzung der Stichprobe eventuell nicht mehr die Population abbildet, die eigentlich hätte untersucht werden sollen, sondern nur noch eine Subpopulation davon. Werden die Untersuchungsergebnisse verzerrt, weil ein Teil der ursprünglich angestrebten Population nicht an der Untersuchung teilnimmt, wird das als „nonresponse bias“ bezeichnet und führt zu einer Reduktion der Qualität der erhobenen Daten (Groves, 2006; Davern 2013). In der Forschung gibt es zwei Ansätze, um die Probleme zu reduzieren, die mit einem „nonresponse bias“ einhergehen. Einerseits können die Auswirkungen des „nonresponse bias“ durch die Erhebung gewisser Parameter besser eingeschätzt werden und verlieren somit einen Teil ihres Einflusses (Groves & Peytcheva, 2008; Wagner 2012; Davern 2013). Andererseits wird versucht, die Ausschöpfungsquote zu optimieren (Singer, 1978; Virtanen, Sirkiä & Jokiranta, 2007; Vicente & Reis, 2010; Pforr et al, 2015), auch wenn die Ausschöpfungsquote allein keinen guten Indikator für den „nonresponse bias“ darstellt (Wright 2015; Sturgis, Williams, Brunton-Smith & Moore, 2017).

Eine gute Ausschöpfungsquote ist oft aber ohnedies nötig, um die gewünschte Stichprobengröße zu erreichen. Laut Tourangeau (2017) befand sich etwa das Geschäft mit den Wahlumfragen in den USA nach der Präsidentschaftswahl 2016 in der Krise aufgrund der in den vorangegangenen 30 Jahren stetig gefallen Ausschöpfungsquote (siehe auch Baruch, 1999) und den damit zusammenhängenden stark gestiegenen Kosten für Wahlumfragen. Für postalische Befragungen, „face-to-face“-Umfragen sowie Telefonumfragen wurden über die Jahre immer mehr Kontaktversuche notwendig, um die gewünschte Stichprobengröße zu erreichen (Tourangeau, 2017). In Anbetracht dieser Entwicklung ist es nicht verwunderlich, dass es eine Vielzahl an Untersuchungen gibt, die sich mit der Teilnahmebereitschaft und damit auch mit der Ausschöpfungsquote von postalischen Befragungen (u.a. Brennan & Hoek, 1992; Groves, Presser & Dipko, 2004; Konradt & Fary, 2006), „face-to-face“-Umfragen (u.a. Singer, Groves & Corning, 1999; Hazel & Clark 2013, Lynn 2014) und telefonischen Befragungen auseinandersetzen (u.a. Schnauber & Daschmann, 2008; Maynard, Freese & Schaeffer, 2010; Maynard & Hollander, 2014). Ebenso gibt es zur Verbesserung der Ausschöpfungsquote bei web-basierenden Umfragen eine Vielzahl an Studien (Jin, 2011; Fang, Wen & Pavur, 2012; Aerny-Perreten, Domínguez-Berjón, Esteban-Vasallo & García-Riolobos, 2015; Liu, Kuriakose, Cohen & Cho,

2015; Keusch, 2015), was auch damit zu tun hat, dass die Ausschöpfungsquoten bei web-basierten Umfragen vergleichsweise relativ niedrig (Lozar Manfreda et al, 2008; Wolfe, Converse, Airen & Bodenhorn, 2009) und die Dropout-Raten relativ hoch sind (Galesic, 2006; Horger, 2010).

Wird als Untersuchungsmethode eine postalische Befragung, eine telefonische Befragung, eine „face-to-face“-Umfrage oder eine web-basierte Umfrage gewählt, ist eine Reduzierung der Ausfallquote also nur bedingt von Bedeutung und sollte vor allem bezüglich eines „nonresponse bias“ im Auge behalten werden. Denn sollten sich Versuchsteilnehmer*innen entschließen (doch) nicht an einer dieser Untersuchungen teilzunehmen, entsteht in der Regel kein bis kaum ein Mehraufwand für die Versuchsleiter*innen, weswegen es vor allem von Bedeutung ist, eine gute Ausschöpfungsquote zu erzielen. Die Effizienz der Untersuchung wird also eher durch einen Verlust an Qualität der Daten durch einen „nonresponse bias“ beeinträchtigt, als durch einen Mehraufwand. Anders gestaltet sich dies jedoch bei Laboruntersuchungen, bei denen die Versuchsleiter*innen sich Zeit nehmen, Räumlichkeiten organisieren und eventuell sogar teure Gerätschaften wie etwa ein fMRT-Scanner reserviert werden müssen. In diesem Fall ist die Ausfallquote sehr wohl ausschlaggebend, um eine effiziente und ressourcenschonende Forschung zu ermöglichen.

Trotz des Mehraufwands, der bei Laboruntersuchungen durch den Ausfall von Versuchspersonen entsteht, gibt es kaum Studien, die sich damit beschäftigen. Im Zuge der Recherche für diese Arbeit wurde nur eine Studie gefunden, die sich direkt mit diesem Problem beschäftigt. Wie bereits erwähnt, haben Fienup und Dixon (2003) festgestellt, dass Versuchspersonen mit einer höheren Wahrscheinlichkeit zu einem Untersuchungstermin erscheinen, wenn sie am Tag davor direkt per Telefon an den Versuch erinnert werden. Weiterführend konnten keine Studien gefunden werden, die sich diesem oder einem ähnlichen Problem annahmen, auf die Studie von Fienup und Dixon aufbauten oder diese zitierten. Die Dokumentation der Literaturrecherche befindet sich im Anhang (siehe „Dokumentation der Literaturrecherche“).

Effizienter Forschung liegen also eine hohe Datenqualität, eine hohe Ausschöpfungsquote und eine niedrige Ausfallquote zugrunde. Während die Datenqualität durch das Einhalten von Forschungsstandards gewährleistet werden kann, ist das Erreichen einer hohen Ausschöpfungsquote abhängig von der gewählten Forschungsmethode und einer entsprechenden Planung der Untersuchung. Dadurch kann sowohl einem „nonresponse bias“ vorgebeugt als auch gewähr-

leistet werden, dass die gewünschte Stichprobengröße erreicht wird. Ob die gewünschte Stichprobengröße erreicht wird, ist aber ebenso abhängig von der Ausfallquote. Welche Einflussfaktoren die Ausfallquote beeinflussen könnten, wird in der Folge näher besprochen.

Mögliche Einflussfaktoren auf die Ausfallquote

Da es nur sehr wenige Forschungsergebnisse gibt, die sich direkt auf die Ausfallquote beziehen, ist es wichtig, Forschung aus verwandten Forschungsgebieten näher zu betrachten. Dabei ist vor allem Forschung zur Steigerung der Effizienz von Untersuchungen von Bedeutung, mit besonderem Augenmerk darauf, mit welchen Methoden die Effizienz gesteigert werden kann. Deshalb werden jene Erkenntnisse zur Ausschöpfungsquote, zu Dropout- bzw. Abbruchquoten und dazu, unter welchen Bedingungen Personen Termine in medizinischen oder psychologischen Praxen eher wahrnehmen, die Parallelen zur Ausfallquote aufweisen, herangezogen.

Solche Parallelen zwischen Ausfallquote und Ausschöpfungsquote zeigen sich unter anderem bei der Motivation der Versuchspersonen an einer Untersuchung teilzunehmen. Durch interessante Aufbereitung und ansprechende Anreize für die Teilnahme kann beispielsweise Einfluss auf die Motivation teilzunehmen genommen werden. Abbruchquoten haben mit der Ausfallquote gemein, dass das initiale Interesse von Personen, an der Untersuchung teilzunehmen, aufrechterhalten bleibt und die mit der Untersuchung einhergehenden Anstrengungen gerechtfertigt erscheinen müssen. Die Gemeinsamkeiten des Ausfalls von Versuchspersonen mit der Arbeit in medizinischen und psychologischen Praxen zeigen sich vor allem darin, dass eine reibungslose Terminkoordinierung für effiziente Arbeit notwendig ist. Dementsprechend ist es von großer Bedeutung, dass Patient*innen ihre Termine einhalten, was durch den Einsatz von verschiedenen Systemen zur Terminbestätigung und durch Erinnerungen an den Termin positiv beeinflusst werden kann.

In der Folge werden anhand der aus der Literatur gesammelten Erkenntnisse mögliche Einflussfaktoren auf die Ausfallquote ausgearbeitet und 14 Forschungshypothesen dazu aufgestellt. Auf Basis dieser Literatur aus den verwandten Forschungsfeldern lassen sich die Einflussfaktoren in fünf Kategorien einteilen. Zwei dieser Kategorien können direkt durch Versuchsleiter*innen gestaltet werden: die Planung, Gestaltung und Durchführung der Untersuchung sowie die Rekrutierung von Versuchspersonen. Drei weitere „Problemfelder“ können vor allem aus Sicht der potentiellen Versuchspersonen über Ausfall oder Teilnahme an einer Untersuchung entscheiden. Diese sind erstens die Persönlichkeitsmerkmale der Versuchsteilnehmer*innen, zweitens motivationale Faktoren, wie das Interesse an der Untersuchung oder eine für die Teilnahme

versprochene Belohnung und drittens Umweltfaktoren, wie das Wetter oder der Zeitpunkt der Testung.

Einflussfaktoren bei der Planung, Gestaltung und Durchführung der Untersuchung

Die Planung, Gestaltung und Durchführung der Untersuchung sind zentrale Elemente der experimentellen Psychologie die, wie oben bereits ausgeführt, auf verschiedenen Ebenen ausschlaggebend für die Qualität wissenschaftlicher Forschung sind.

Eine gute Planung beugt möglichen Problemen vor und ermöglicht effiziente Forschung, indem etwa ausreichend Proband*innen rekrutiert werden, ein „nonresponse bias“ vermieden und die Ausfallquote reduziert wird. Auch im medizinischen und psychotherapeutischen Bereich spielt die Planung von Terminen eine wichtige Rolle, um ein effizientes Arbeiten zu ermöglichen. Deshalb gibt es auch in diesem Bereich Untersuchungen zum Einsatz von Erinnerungen, ähnlich der zuvor erwähnten Studie von Fienup und Dixon (2003), damit vereinbarte Termine tatsächlich wahrgenommen werden. Je nach Kontext konnte ein unterschiedlicher Nutzen von Erinnerungen an Termine nachgewiesen werden.

In einer psychologischen Klinik in Brisbane wurde bei Personen, die eine Psychotherapie in Anspruch nahmen, keine Veränderung der Anwesenheit zu vereinbarten Terminen festgestellt, wenn diese per SMS einen Tag vor dem Termin eine Erinnerung an ihren Termin bekamen. (Clough & Casey, 2014). Im Gegensatz dazu stellten Gullo, Irvine, Feeney & Connor (2018) fest, dass alkoholabhängige Personen, die sich in Behandlung befinden, ihre ambulanten Termine eher wahrnehmen, wenn sie eine Erinnerung per SMS erhalten. Ebenso konnte ein innovativer, an die Patient*innen angepasster Managementansatz eine effiziente Nutzung der Ressourcen in einer ambulanten psychiatrischen Klinik fördern (Boland & Burnett, 2014). Bei diesem Ansatz erhielten Patient*innen zwei Wochen vor dem Termin eine schriftliche Erinnerung an den Termin. Zusätzlich wurde, wenn möglich, eine Entlassung aus der Behandlung vorgeschlagen, nachdem Patient*innen erstmalig nicht zu einem Termin erschienen waren sowie der Umgang mit dem Nichterscheinen von den Fachärzten im Team besprochen. Auch in einer Literaturanalyse von über 60 Studien zu SMS-Erinnerungen im Gesundheitswesen (Kannisto, Koivunen & Välimäki, 2014) konnte ein positiver Effekt zur Unterstützung von Behandlungen durch Erinnerungen festgestellt werden. So berichteten 77% der untersuchten Studien eine Verbesserung von gewünschten Verhaltensweisen. 40% der Studien berichteten eine Verbesserung beim Einhalten des Zeitplans einer Medikation bzw. einer Behandlung und 18% berichteten von einer Verbesserung beim Erscheinen zu Terminen bzw. einer Verringerung des Nicht-Erscheinens. Milward, Lynskey & Strang (2014) halten in einer weiteren Literaturanalyse fest,

dass Vergesslichkeit die häufigste Erklärung für versäumte Termine ist und deshalb Erinnerungen via Text-Nachrichten eine kosteneffiziente Möglichkeit darstellen, um die Anwesenheitsraten im Gesundheitsbereich drastisch zu erhöhen. Entsprechend dieser Untersuchungen ist anzunehmen, dass die Effizienz der Forschung durch Erinnerungen an einen Termin verbessert werden kann. Da dem unerwarteten Ausfall von Versuchspersonen dadurch zumindest teilweise vorgebeugt wird, können so auch zeitliche Ressourcen besser ausgenutzt werden. Daraus folgend lautet die erste Hypothese dieser Arbeit:

Hypothese 1.1.: Erinnerungen an den Untersuchungstermin reduzieren den Ausfall von Versuchspersonen.

Gerade in Kliniken, die mit ambulanten Patient*innen arbeiten, wird außerdem häufig versucht, durch den Einsatz verschiedener Systeme zur Terminfindung die Einhaltung von Terminen zu verbessern. Hawker (2007) berichtet beispielsweise über neun Studien, in denen das Nicht-Erscheinen von Patient*innen reduziert werden konnte, indem man ein „opt-in system“ einführte, ein System, in dem ein erhaltener Termin erst fixiert wird, wenn die Patient*innen den Termin bestätigen. In einer Klinik für Essstörungen wurde außerdem das Nicht-Erscheinen zum ersten Termin reduziert, indem ein „partial booking“ System eingeführt wurde. Bei diesem System erhielten die Patient*innen einen Brief, der sie aufforderte, sich telefonisch bei der Klinik zu melden und einen für sie passenden Termin zu vereinbaren (Jenkins, 2017). Diese Ansätze haben gemeinsam, dass sie von der Patientin bzw. dem Patienten eine Bestätigung zum Termin bzw. die Auswahl eines Termins fordern. Da in der experimentellen Psychologie die Forscher*innen einen Termin mit den Versuchspersonen vereinbaren wollen und nicht umgekehrt, ist es wichtig, entsprechend auf deren Termingestaltung einzugehen. Eine freie Terminauswahl und die damit einhergehende Terminbestätigung könnten, wie im klinischen Kontext, auch in der experimentellen Psychologie zu positiven Effekten bezüglich der Ausfallquote führen.

*Hypothese 1.2.: Die freie Auswahl eines Termins durch die Versuchsteilnehmer*innen reduziert die Ausfallquote.*

Zu hinterfragen ist auch, welche Rolle die gewählten Untersuchungsmethoden, wie Leistungstests, fMRT, EEG, etc., bezüglich der Ausfallquote spielen. Dazu, welchen Einfluss die Wahl der Untersuchungsmethoden auf die Ausfallquote bei Laboruntersuchungen hat, konnten keine Studien gefunden werden (welche Datenbanken und welche Suchbegriffe verwendet wurden, ist im Anhang unter „Dokumentation der Literaturrecherche“ vermerkt). Laut der „Social

Exchange Theory“¹ ist jedoch anzunehmen, dass Personen sich von vornherein nicht zu einer Untersuchung anmelden würden, wenn sie bestimmte Untersuchungsmethoden aufgrund des dabei entstehenden Aufwands vermeiden möchten. Nachdem jene Personen, die sich zu einer Untersuchung anmelden, sich bewusst dazu entscheiden, an der Untersuchung trotz oder sogar wegen der Untersuchungsmethoden teilzunehmen, sollten die Untersuchungsmethoden keinen Einfluss auf die Ausfallquote nehmen, sondern nur auf die Ausschöpfungsquote.

Hypothese 1.3.: Die verwendeten Untersuchungsmethoden haben keinen Einfluss auf den Ausfall von Versuchspersonen.

Anders verhält es sich mit der für die Untersuchung angesetzten Dauer. Dazu gibt es einige Studien, die jedoch zu unterschiedlichen Erkenntnissen kamen. Eine Studie berichtet bei postalischen Umfragen von einer höheren Ausschöpfungsquote bei kurzen Fragebögen als bei deutlich längeren Fragebögen (Roszkowski & Bean, 1990). Weitere Studien konnten sowohl bei postalischen Umfragen (Koloski et al, 2001; Robb, Gattling & Wardle, 2017), als auch bei web-basierten Umfragen (Blumenberg et al, 2019; Han, Albaum, Wiley & Thirkell, 2009) keinen Einfluss der Fragebogenlänge auf die Ausschöpfungsquote feststellen. Dillman, Sinclair und Clark (1993) fanden bei ihrer postalischen Umfrage beim Einsatz von kürzeren Fragebögen in ihrer Stichprobe für eine Subpopulation keinen Effekt der Fragebogenlänge, bei einer anderen Subpopulation jedoch eine verbesserte Ausschöpfungsquote. Aus diesen Ergebnissen schlussfolgerten sie, dass auch die inkonsistenten Forschungsergebnisse vorhergehender Studien eventuell auf die unterschiedlichen befragten Populationen zurückzuführen sein könnten.

Unabhängig von der Ausschöpfungsquote beginnen mehr Personen web-basierte Umfragen, wenn für diese eine kürzere Bearbeitungsdauer angekündigt wird (Crawford, Cooper & Lamias, 2001; Galesic & Bosnjak, 2009; Göritz, 2014). Wird durch eine ungenau oder falsch angegebene Bearbeitungsdauer diese jedoch unterschätzt, so steigt die Anzahl der Personen, die die Umfrage vorzeitig abbrechen (Crawford, Cooper & Lamias, 2001). Bei einem Online-Panel für Umfragen wurde außerdem festgestellt, dass kürzere Umfragen häufiger auch wirklich abgeschlossen werden, nachdem sie bereits begonnen wurden (Göritz, 2014). Durch die kürzere Dauer scheint die Hemmschwelle für Versuchspersonen niedriger zu sein, die Umfrage zu beenden. Die Tendenz zu geringeren Ausschöpfungsquoten und die gehäuften Abbrüche von Ver-

¹Die Social Exchange Theory erklärt das Verhalten von Menschen als Wechselwirkung zwischen dem Aufwand und dem zu erwartenden Vorteil, der aus dem Verhalten hervorgeht (Poon, Albaum & Evangelista, 2003).

suchspersonen bei längeren Untersuchungen zeigen, dass die Untersuchungsdauer durchaus einen Einfluss auf die Teilnahmemotivation der Versuchspersonen haben kann. Dementsprechend könnten Versuchspersonen auch bei Untersuchungen, welche die persönliche Anwesenheit erfordern, bei einer kürzeren Untersuchungsdauer häufiger dazu bereit sein, auch tatsächlich zu einem vereinbarten Untersuchungstermin zu erscheinen.

Hypothese 1.4.: Die Ausfallquote ist bei kurzen Untersuchungen geringer als bei langen Untersuchungen.

Lange Zeit wurde auch über Vertraulichkeitszusagen am Beginn von wissenschaftlichen Untersuchungen diskutiert. Wenngleich diese heutzutage aus Gründen des Datenschutzes obligatorisch sind, gelangten Hippler, Schwarz & Singer (1990, S. 65) zu der Ansicht, „daß Vertraulichkeitszusagen die Bereitschaft an einer Befragung teilzunehmen, negativ beeinflussen können.“ Dabei argumentierten sie, dass diese Zusagen zu der Annahme führen, dass der Inhalt der Untersuchung heikel sei und so die Teilnahmebereitschaft reduziert würde. Trotz der laut dieser Studie geringeren Teilnahmebereitschaft sind diese Zusagen jedoch wichtig, um einen „informed consent“ zu erhalten, da dieser zu einem Sensibilisierungs-Effekt und damit zu qualitativ hochwertigeren Daten führt (Singer, 1978). Da Vertraulichkeitszusagen mittlerweile verpflichtend und auf jeden Fall vor der Untersuchung zu unterschreiben sind, wird keine Hypothese dazu aufgestellt.

Gestaltung des Rekrutierungsprozesses

Auch die Gestaltung des Rekrutierungsprozesses kann Einfluss auf die Ausfallquote haben. „Das Anwerben von Versuchspersonen ist ein Schritt, dem aus verschiedenen Gründen große Relevanz zukommt: Man muss eine zur jeweiligen Untersuchungsanlage passende und ausreichend große Stichprobe finden, die man für eine ernsthafte Mitarbeit im Rahmen des Experiments motiviert.“ (Koch, Peter & Müller, 2019, S.124) Dabei müssen genug interessierte Personen durch den Rekrutierungsprozess erreicht werden, um auch nach dem Ausfall von Versuchspersonen eine ausreichend große Stichprobe zu gewährleisten. Eine Reduktion der Ausfallquote führt dazu, dass initial weniger potentielle Versuchspersonen erreicht werden müssen, um eine entsprechend große Stichprobe zu erhalten, was wiederum die Effizienz im Zuge des Forschungsprozesses erhöht.

Im Wesentlichen können im Rekrutierungsprozess drei Faktoren variiert werden: die Methoden, die zur Rekrutierung eingesetzt werden, die Art wie Einladungen formuliert und gestaltet werden und der Zeitpunkt, zu dem Einladungen versendet oder verteilt werden. Die Wahl der

Rekrutierungsmethoden, wie beispielsweise Aushänge und Rundmails, die für eine Teilnahme werben, oder das direkte Anwerben von Versuchspersonen in Einzelgesprächen bestimmt, welche Subpopulationen angesprochen werden. Gewisse Rekrutierungsmaßnahmen können einerseits nur gewisse Personengruppen erreichen und wirken andererseits nicht auf alle Personengruppen gleich ansprechend (Shih & Fan, 2008). Zusätzlich benötigen die gesetzten Maßnahmen einen unterschiedlichen Zeitraum, um wirksam zu werden. Je nach Maßnahme kann es, wie beispielsweise bei Aushängen, wochenlang dauern, bis die volle Reichweite ausgeschöpft ist, während bei Einladungen via E-Mail bereits nach wenigen Tagen ein Großteil der interessierten, kontaktierten Personen reagiert (Baur & Florian, 2009; SoSci Panel, n.d.). Ob der Einsatz unterschiedlicher bzw. mehrerer Rekrutierungsmaßnahmen auch zu unterschiedlichen Ausfallquoten führen kann, ist unklar und wird deshalb explorativ in der Voruntersuchung erhoben.

Des Weiteren ist es von Bedeutung, wie die gesetzten Rekrutierungsmaßnahmen ausgestaltet werden. Beispielsweise hat der Inhalt der Betreffzeile von E-Mails einen Einfluss darauf, wie viele Personen auf die darin versendete Einladung reagieren (Porter & Whitcomb, 2005). Auch die Personalisierung einer Einladung per E-Mail durch Einfügen des Namens des Empfängers bzw. der Empfängerin kann dazu führen, dass mehr Personen eine ausgesendete Online-Umfrage beginnen und hat somit einen positiven Einfluss auf die Antwortrate (Sauer mann & Roach, 2013). Außerdem kann die Formulierung einer Einladung, die per E-Mail versandt wurde, dazu führen, dass potentielle Versuchspersonen zu einer Teilnahme motiviert oder aber abgeschreckt werden. Koch, Peter & Müller (2019, S.124) schreiben dazu: „Bei der Rekrutierung von Versuchspersonen ist auf einen freundlichen und höflichen Umgang zu achten, Zwang oder Druck dürfen nicht ausgeübt werden.“

Diese Studien beschäftigen sich nur mit der Ausschöpfungsquote und dienen somit nur als Orientierung dafür, welche möglichen Einflussfaktoren es bei der Formulierung und Gestaltung einer Einladung auf die Ausfallquote geben könnte. Wie die optische Gestaltung der Einladungen die Ausfallquote beeinflusst, wurde im Zuge dieser Arbeit nicht überprüft. Der Einfluss der Formulierung von Texten, die zum Anwerben verwendet werden, auf die Ausfallquote wurde jedoch zum Teil in der Hauptuntersuchung explorativ erforscht.

Koch, Peter und Müller (2019, S.129) berichten auch davon, dass der Tag und die Uhrzeit an dem die Einladungen zu einer Untersuchung verschickt werden, einen Einfluss auf die Anzahl der Antworten und Anmeldungen zur Untersuchung haben kann. Wie sich der Zeitpunkt, zu dem die Einladungen ausgesendet werden, auswirkt, ist dabei von der jeweiligen Population

abhängig. So sind Studierende beispielsweise unter der Woche flexibler und leichter zu erreichen als Berufstätige. Im Rekrutierungsprozess sollte also überlegt werden, wie die gewünschte Stichprobe am besten erreicht werden kann, um möglichst viele Personen zu rekrutieren. Eine gelegentlich zitierte Studie von Faught, Whitten & Green (2004) stellt die These auf, dass mittwochvormittags der beste Zeitpunkt für die Aussendung von Einladungen zu einer Umfrage sei. Bei ihrem Experiment versendeten sie per E-Mail Einladungen an „manufacturing firms“ in den USA – nicht aber direkt an Einzelpersonen. Die Einladungen wurden in gleicher Anzahl an jedem Wochentag vormittags und nachmittags versandt. Von den 2516 ausgesendeten E-Mails waren 487 nicht zustellbar, die statistische Verteilung der nicht zustellbaren E-Mails wurde nicht berichtet. Auf die restlichen 2029 E-Mails antworteten 54 Firmen, also 2,66%. Die meisten Antworten, jeweils acht, kamen von den Einladungen, die am Dienstagnachmittag und am Mittwochvormittag gesendet wurden. Am Montagvormittag kamen sechs Antworten und damit die drittmeisten. Ein Argument dafür, warum Mittwochvormittag ein besserer Zeitpunkt sei als Dienstagnachmittag führen die Forscher nicht an. Angesichts der durch die niedrige Rücklaufquote entstandenen geringen Stichprobengröße, der Tatsache, dass Firmen anstatt von Einzelpersonen kontaktiert wurden sowie aufgrund der nicht berichteten Verteilung der nicht-zustellbaren Einladungen, bleibt fraglich, wie diese Ergebnisse zu bewerten sind. Die Forscher lieferten jedoch den Beweis, dass sich die Rücklaufquote bei Einladungen, die nur Mittwochvormittag ausgesandt wurden, auf 4,10% erhöhte (81 Antworten von 1976 zugestellten Einladungen).

Auch Sauermann und Roach (2013) haben anhand einer Stichprobe aus Junior Wissenschaftler*innen und Techniker*innen überprüft, an welchen Tagen und zu welcher Uhrzeit sich die Rücklaufquote verändert, konnten dabei aber keinen Zeitpunkt ausmachen, zu dem sich die Rücklaufquote verbessert hätte. Ob diese Ergebnisse von der Stichprobe abhängig sind, bleibt unklar. Welchen Einfluss der Zeitpunkt, zu dem die Einladungen ausgesendet werden, auf die Ausfallquote hat, kann auch im Zuge dieser Arbeit nicht genauer untersucht werden, denn es war nicht möglich, Tag und Zeitpunkt, zu dem die Rekrutierungsmaßnahmen für die Hauptuntersuchung gesetzt wurden, repräsentativ zu variieren.

Im Kapitel „Diskussion der Ergebnisse der Voruntersuchung“ wird auf Seite 44 unter „Ergänzung: allfällige Sonderfälle“ nachträglich Hypothese 2.1. hergeleitet und aufgestellt. Diese Hypothese wird dem Themenblock „Gestaltung des Rekrutierungsprozesses“ zugeordnet.

Ein weiterer Faktor, der über die Teilnahme bzw. den Ausfall von potentiellen Versuchspersonen entscheidet, ist deren Persönlichkeit. Am bedeutendsten dürfte hierbei die Zuverlässigkeit sein. Diese ist im Sinne des Big Five Modells der Persönlichkeit (Digman, 1990; McCrae & John, 1992) – ein Modell, bei dem die menschliche Persönlichkeit in fünf große Faktoren (Offenheit, Extraversion, Neurotizismus, Gewissenhaftigkeit und Verträglichkeit) eingeteilt wird, – der Gewissenhaftigkeit zuzuordnen (Dudley, Orvis, Lebiecki & Cortina, 2006; Hough & Ones, 2001; Roberts et al, 2005; Stewart, 1999). Im Sinne dieses Konstrukts beschreibt die Zuverlässigkeit (im Englischen: dependability) Personen, die verlässlich und verantwortungsbewusst sind, ihren Verpflichtungen nachkommen und Autoritäten respektieren (Olaru et al, 2019). Entsprechend dieser Definition sollten zuverlässigere Personen, da sie ihren Verpflichtungen nachkommen, seltener bei Untersuchungen ausfallen und nach Anmeldung zu einer Untersuchung auch eher zum vereinbarten Termin erscheinen.

Hypothese 3.1.: Zuverlässigere Personen fallen als Versuchspersonen bei wissenschaftlichen Untersuchungen seltener aus.

In mehreren Studien konnten Zusammenhänge zwischen der Zuverlässigkeit und der Gewissenhaftigkeit nachgewiesen werden, weshalb die Zuverlässigkeit auch immer wieder als eine wesentliche Subfacette der Gewissenhaftigkeit gehandhabt wird (Dudley, Orvis, Lebiecki & Cortina, 2006; Hough, 1992; Hough & Ones, 2001; Olaru et al, 2019; Roberts et al, 2005; Stewart, 1999). Die Zuverlässigkeit manifestiert sich hierbei eher auf der interpersonalen Komponente der Gewissenhaftigkeit in Eigenschaften wie dem Verantwortungs- und Pflichtbewusstsein (Roberts et al, 2005). Da das Gesamtkonzept der Gewissenhaftigkeit auch diese Eigenschaften der Zuverlässigkeit umfasst, sollte entsprechend der vorherigen Hypothese also auch die Gewissenhaftigkeit von Versuchspersonen einen Einfluss darauf haben, wie wahrscheinlich diese bei einer Untersuchung ausfallen.

Hypothese 3.2.: Gewissenhaftere Personen fallen als Versuchspersonen bei wissenschaftlichen Untersuchungen seltener aus.

Ein weiteres Persönlichkeitskonstrukt, das ein Einflussfaktor auf die Ausfallquote sein könnte, ist die soziale Verantwortung (im Englischen: social responsibility). Cough, McClosky & Meehl (1952) charakterisieren verantwortungsbewusste Menschen als zuverlässig, vertrauenswürdig, bereit, die Konsequenzen für ihr Handeln zu übernehmen und mit einem Pflichtgefühl

gegenüber Gruppen, denen sie angehören. Altman (2005) beschreibt soziale Verantwortung im Speziellen als ein Gefühl der Anteilnahme am Wohlergehen Anderer in der Gesellschaft, das Personen dazu bewegt, im Sinne Anderer zu handeln, wenn deren Wohlergehen bedroht ist. Laut Bierhoff (2012) „ist soziale Verantwortung dabei einerseits durch die Bereitschaft, sich gegenüber sozialen Adressaten verlässlich zu verhalten und sich dafür auch persönlich zu engagieren, andererseits durch die Verwendung fairer Mittel und allgemein durch die Einhaltung sozialer Spielregeln gekennzeichnet.“ Somit sollten Versuchspersonen, die eine hohe soziale Verantwortung aufweisen, sich gegenüber einer bzw. einem Versuchsleiter*in so weit verantwortlich fühlen, dass sie einen vereinbarten Termin auch tatsächlich wahrnehmen.

Hypothese 3.3.: Versuchspersonen, die eine höhere soziale Verantwortung aufweisen, fallen bei wissenschaftlichen Untersuchungen seltener aus.

Außerdem gibt es Nachweise, dass die Ausschöpfungsquoten bei männlichen Versuchspersonen bei Online-Umfragen niedriger sind als die Ausschöpfungsquoten bei weiblichen Versuchspersonen (Dykema et al., 2012; Laguilles, Williams & Saunders, 2011; Patrick, et al., 2013). Es bleibt jedoch unklar, ob Männer sich nur mit geringerer Wahrscheinlichkeit zu Untersuchungen anmelden oder ob sie auch eine größere Tendenz aufweisen, bei einer Untersuchung auszufallen, wenn sie sich bereits entschlossen haben, an dieser Untersuchung teilzunehmen. Ob das Geschlecht der Versuchspersonen einen Einfluss auf die Ausfallquote hat, wird deshalb explorativ erhoben.

Motivationale Faktoren

Um einem Ausfall von Versuchspersonen vorzubeugen, ist es notwendig, die Versuchspersonen ausreichend zu einer Teilnahme zu motivieren. Auch auf diesem Gebiet konnte keine Forschung gefunden werden, die über direkte Zusammenhänge mit der Ausfallquote berichtet (welche Datenbanken und welche Suchbegriffe verwendet wurden, ist im Anhang unter „Dokumentation der Literaturrecherche“ vermerkt). Allerdings ist es möglich, Lehren aus verwandten Forschungsgebieten zu ziehen und daraus plausible Hypothesen abzuleiten. Die Forschung zu Online- und Telefonumfragen sowie zu postalischen Befragungen und persönlichen Interviews liefert zum Teil recht umfassende Ergebnisse dazu, wie sich verschiedene Anreize (im Englischen: incentives) oder das Interesse der Versuchspersonen auf die Ausschöpfungsquote und Abbruchrate auswirken können. Außerdem werden einige sozialpsychologische Theorien herangezogen, um diese Verhaltensweisen zu erklären.

Mehrere Studien (Han, Albaum, Wiley & Thirkell, 2009; Keusch 2015; Poon, Albaum & Evangelista, 2003) untersuchten, welche sozialpsychologische Theorie am besten die Ausschöpfungsquote bei Umfragen erklären kann. Dazu verglichen sie die „Social Exchange Theory“ sowie die Theorien zu „Self-Perception“, „Cognitive Dissonance“ und „Commitment/Involvement“ – Keusch (2015) zusätzlich auch die Theorien zu „Compliance“, „Leverage-Salience“ und „Planned Behaviour“. Die Social Exchange Theory nimmt dabei eine besondere Rolle ein. In diesem Kontext besagt sie, dass die Motivation zur Teilnahme an einer Untersuchung oder Umfrage von möglichen (nicht notwendigerweise monetären) Gegenleistungen, Belohnungen oder Entschädigungen bestimmt ist. Diese werden dann in Relation zu dem mit einer Teilnahme verbundenen Aufwand gesetzt. Zusätzlich fällt die Glaubwürdigkeit einer Aussendung in Bezug auf die gebotene Gegenleistung, Belohnung oder Entschädigung ins Gewicht (Poon, Albaum & Evangelista, 2003). Der gebotene Anreiz ist dabei laut Poon, Albaum und Evangelista (2003) für viele Versuchspersonen der Hauptgrund für die Teilnahme an Umfragen. Zusätzlich meinen sie, dass auch die Theorie zu „Commitment/Involvement“ einen positiven Einfluss auf die Entscheidung über eine Teilnahme an Umfragen hat, wenn Thema, Sponsor*in oder Forscher*in für die potenziellen Versuchspersonen relevant bzw. wichtig sind.

Han, Albaum, Wiley & Thirkell (2009) ordneten die Antworten zu drei offenen Fragen zu Teilnahmehindernissen bei Online Umfragen, die an einer größeren Untersuchung zum Thema „Datenerhebungstechniken bei Umfragen“ angehängt waren, verschiedenen sozialpsychologischen Theorien zu. Dabei schätzten sie den Einfluss der Social Exchange Theory auf die Ausschöpfungsquote bei Online-Umfragen am größten ein. 43,2% der Teilnehmer*innen gaben an, dass für sie der wichtigste Grund, nicht an einer Untersuchung teilzunehmen, der damit verbundene Aufwand sei, vor allem bezogen auf Zeit und Anstrengung. Im Gegensatz dazu gaben 28,3% der Teilnehmer*innen an, dass für sie der wichtigste Grund, um an einer Untersuchung teilzunehmen, ein versprochener Anreiz sei, während 22,1% das Vertrauen zu den Versuchsleiter*innen bezüglich gebotenen Anreiz, Datensicherheit, Identität, etc. als wichtigsten Grund für oder gegen eine Teilnahme ansahen. Nur 6,4% der Antworten wiesen darauf hin, dass eine der anderen Theorien („Self-Perception“, „Cognitive Dissonance“ und „Commitment/Involvement“) eine Teilnahme entscheidend erklären kann.

In der bislang umfangreichsten Studie dazu analysiert Keusch (2015) wie anhand sozialpsychologischer Theorien das Teilnahmeverhalten an Online-Umfragen erklärt werden kann. Dazu führte er eine Umfrage unter Expert*innen auf dem Gebiet der Durchführung von wissenschaftlichen Untersuchungen durch, also unter an Universitäten angestellten Psycholog*innen, die

zumindest ein Doktoratsstudium abgeschlossen hatten. Diese Expert*innen bewerteten, wie gut verschiedene Theorien den Einfluss verschiedener Faktoren auf die Ausschöpfungsquote erklären können. Die Theorien zur “Self-Perception”, “Commitment/Involvement” sowie die “Leverage-Saliency Theory”, die “Theory of Planned Behavior” und zum Teil auch die „Cognitive Dissonance Theory“ erklären demnach, warum Eigenschaften der Versuchspersonen, wie die Persönlichkeit, persönliches Interesse am Thema, Einstellungen zur Umfragenforschung und vorhergehendes Teilnahmeverhalten, einen Einfluss auf die Entscheidung an einer Online-Umfrage teilzunehmen haben. Außerdem sind die Heuristiken zur „Compliance“ nützlich, um die Auswirkungen des Designs von Online-Untersuchungen auf die Ausschöpfungsquote zu verstehen. Die Expert*innen sahen auch die „Social Exchange Theory“ als hilfreich an, um die Einflüsse von persönlichem Interesse am Thema, Einstellungen zur Umfragenforschung sowie bedingungslos erhaltenen Anreizen und der Länge von Fragebögen auf die Ausschöpfungsquote zu erklären.

Entsprechend der „Social Exchange Theory“ sollten Versuchsleiter*innen also den Aufwand für die Versuchspersonen gering halten, während sie möglichst große Anreize für eine Teilnahme schaffen und dabei so vertrauensvoll wie möglich wirken sollten. Je nach Versuchsdesign ist eine Änderung des Aufwands nur schwer möglich, es gilt jedoch abzuwiegen, welche erhobenen Informationen unbedingt notwendig sind und welche eine Untersuchung nur unnötig in die Länge ziehen. Galesic (2006) berichtet passend dazu, dass Versuchspersonen, die eine Online-Untersuchung als belastend empfinden, diese eher abbrechen. Han, Albaum, Wiley & Thirkell (2009) stellten außerdem fest, dass der am häufigsten genannte Grund, nicht an einer Untersuchung teilzunehmen, der damit verbundene Aufwand ist. Daher ist zu vermuten, dass aufwendigere Untersuchungen nicht nur eine geringere Ausschöpfungs- und eine größere Abbruchquote, sondern auch eine größere Ausfallquote nach sich ziehen.

Hypothese 4.1.: Aufwendige Untersuchungen weisen eine höhere Ausfallquote auf.

Zusätzlich zur Reduzierung des Aufwands sollte ein möglichst großer Anreiz für Versuchspersonen geschaffen werden, um an der Untersuchung teilzunehmen. Es gibt mehrere verschiedene Arten von Anreizen, um die Motivation von Versuchspersonen zu erhöhen. Vor allem der positive Einfluss monetärer Anreize auf die Ausschöpfungsquote in Form einer Aufwandsentschädigung ist gut belegt (Cosgrove, 2018; Erwin & Wheelright, 2002; Furse & Stewart, 1984; Mercer, Caporaso, Cantor & Townsend, 2015; Patrick et al., 2013). Nachdem Aufwandsentschädigungen zusätzliche finanzielle Ressourcen darstellen, wird deren Effizienz aber immer

wieder hinterfragt. Denn um den finanziellen Mehraufwand zu rechtfertigen, müsste dieser Mehraufwand die Stichprobe erheblich vergrößern. In der Forschung gibt es unterschiedliche Meinungen dazu, wie effizient Aufwandsentschädigungen sind (Cosgrove, 2018, Erwin & Wheelright, 2002; Mercer, Caporaso, Cantor & Townsend, 2015; Patrick et al., 2013). Unabhängig von der effizienten Nutzung der finanziellen Ressourcen sollten monetäre Anreize aber nicht nur die Ausschöpfungsquote verbessern, sondern auch die Ausfallquote senken, da monetäre Anreize nur durch ein Erscheinen zum Termin in Anspruch genommen werden können.

Hypothese 4.2.: Monetäre Anreize für die Teilnahme reduzieren die Ausfallquote.

Auch der Einsatz von Preisverlosungen als Anreiz zur Teilnahme beeinflusst die Ausschöpfungsquote bei Online-Untersuchungen positiv (Laguilles, Williams & Saunders, 2011; Sauer mann & Roach, 2013). Ebenso kann in Kliniken zu Substanzmissbrauch der Einsatz von Preisverlosungen dazu führen, dass mehr Personen ihre Therapietermine wahrnehmen (Milward, Lynskey & Strang, 2014). In beiden Fällen sind die Auswirkungen am stärksten, wenn die Möglichkeit auf ein paar wenige, aber hohe Gewinne besteht (Milward, Lynskey & Strang, 2014; Sauer mann & Roach, 2013). Die Möglichkeit an einer Preisverlosung teilzunehmen sollte diesen Ergebnissen entsprechend auch dazu führen, dass bei Untersuchungen weniger Personen ausfallen.

Hypothese 4.3.: Die Möglichkeit der Teilnahme an einer Preisverlosung reduziert die Ausfallquote.

Eine weitere Möglichkeit, um Anreize für die Teilnahme zu schaffen, ist es, Studierende dazu zu verpflichten, sich im Laufe ihres Studiums für eine gewisse Mindestanzahl an Stunden als Versuchsperson zur Verfügung zu stellen. Beispielsweise müssen an der Karl-Franzens-Universität Graz Psychologiestudierende nachweisen, dass sie bis zum Beginn des Bachelorseminars sogenannte „Untersuchungsscheine“ im Ausmaß von 25 Stunden gesammelt haben (Universität Graz Institut für Psychologie, n.d.). Koch, Peter und Müller (2019, S.124) warnen jedoch: „Solche Einschränkungen persönlicher Freiheiten können unter Umständen Reaktanz der Versuchspersonen nach sich ziehen und sogar bewirken, dass diese das Experiment absichtlich zu stören versuchen.“ Diese Reaktanz hätte zwar Einfluss auf die Qualität der Untersuchungsergebnisse, die Verpflichtung sollte aber prinzipiell dazu führen, dass weniger Versuchspersonen ausfallen, sofern sie einen Untersuchungsschein oder ähnliches benötigen.

Hypothese 4.4.: Personen, die dazu verpflichtet sind, sich als Versuchsperson zur Verfügung zu stellen, fallen seltener aus.

All diese Anreize motivieren die Versuchspersonen extrinsisch. Versuchspersonen können für die Teilnahme an einer Untersuchung jedoch auch intrinsisch motiviert sein. In diesem Fall stellen sich Versuchspersonen aus eigenem Interesse oder auch zur Unterhaltung zur Verfügung oder möchten einen Beitrag zur Wissenschaft leisten. (Koch, Peter & Müller, 2019, S.129). Dementsprechend nehmen Personen eher an einer Umfrage teil, wenn sie das Thema dieser Umfrage interessanter finden (Martin, 1994) und brechen diese dann auch seltener ab (Galesic, 2006). Auch die gründliche Bearbeitung der Umfrage hängt vom Interesse am Thema ab, wie genau, ist jedoch umstritten. Während bei der Untersuchung von Martin (1994) die Gruppe, die mehr Interesse am Thema hatte, weniger Items übersprang, stellten Tourangeau, Groves, Kennedy & Yan (2009) bei ihren Versuchspersonen fest, dass bei jenen mit größerem Interesse mehr Antworten zu einzelnen Items fehlten. Dass Versuchspersonen mit Interesse am Thema eher dazu bereit sind, an Umfragen teilzunehmen und diese nicht abbrechen, lässt jedoch vermuten, dass Personen mit Interesse am Thema auch nach Anmeldung zu einer wissenschaftlichen Untersuchung mit höherer Wahrscheinlichkeit zu dieser erscheinen. Die Ausfallquote sollte somit sinken.

Hypothese 4.5.: Versuchspersonen mit Interesse am Thema der Untersuchung fallen seltener aus als Versuchspersonen ohne Interesse am Thema.

Eine Studie zu Online-Panels zeigt auch eine erhöhte Bereitschaft Online-Umfragen zu beginnen, wenn Versuchspersonen sich beim Online-Panel zur Unterhaltung anmeldeten (Keusch, Batinic & Mayerhofer, 2014). In derselben Studie wird aber auch festgehalten, dass Personen, die sich aufgrund der versprochenen Bezahlung beim Online-Panel angemeldet haben, insgesamt mehr Umfragen beginnen und weniger abbrechen als jene Personen, die sich aus intrinsisch motivierten Gründen angemeldet haben, wie die Verbesserung von Produkten und Dienstleistungen oder zur Unterhaltung. Da die Anmeldung zu einem Online-Panel schon mit der Intention passiert, an mehreren Umfragen teilzunehmen, sind diese Ergebnisse jedoch schwer mit einzelnen unabhängigen Untersuchungen vergleichbar. Denn für einzelne Untersuchungen werden oft auch Versuchspersonen kontaktiert, die nicht von vornherein Interesse daran gezeigt haben, sich an wissenschaftlicher Forschung zu beteiligen. Welchen Einfluss die intrinsische Motivation auf die Teilnahme an Untersuchungen zur Unterhaltung oder um einen Beitrag zur

Wissenschaft zu leisten hat, bleibt somit offen und wird in der vorliegenden Arbeit nicht weiter erforscht.

Im Kapitel „Diskussion der Ergebnisse der Voruntersuchung“ wird auf Seite 42 ff unter „Ergänzung: allfällige Sonderfälle“ nachträglich Hypothese 4.6. hergeleitet und aufgestellt. Diese Hypothese wird ebenfalls dem Themenblock „Motivationale Faktoren“ zugeordnet.

Umweltfaktoren

Die Teilnahme von Versuchspersonen an einer Untersuchung kann auch durch Umweltfaktoren beeinflusst werden. Wie bei den Persönlichkeitsmerkmalen der Versuchspersonen können Versuchsleiter*innen auch hier zum Teil keinen oder nur wenig Einfluss auf die Auswirkungen nehmen. Zu wissen, welche möglichen Einflussfaktoren es gibt, kann durch gewissenhafte Planung der Untersuchung dennoch dabei helfen, die Ausfallquote gering und damit die Forschung effizient zu halten.²

Der Zeitraum, in dem eine Untersuchung durchgeführt wird, ist ein solcher Faktor, der die Teilnahme an Untersuchungen beeinflussen kann. Es ist damit zu rechnen, dass zu gewissen Zeiten im Jahr, wie zum Beispiel in Ferienzeiten, nicht die volle Reichweite der gesetzten Rekrutierungsmaßnahmen ausgeschöpft werden kann (SoSci Panel, n.d.). Insofern ist es empfehlenswert, eine Untersuchung so zu planen, dass während der Durchführung der Untersuchung möglichst viele potentielle Versuchspersonen erreicht werden und die Möglichkeit haben, an der Untersuchung teilzunehmen. Dies betrifft jedoch vor allem die Ausschöpfungsquote, nicht aber die Ausfallquote.

Ebenso spielt der Zeitpunkt der Untersuchung am jeweiligen Tag eine Rolle für die Qualität der Untersuchungsergebnisse. Die Leistung einiger Versuchsteilnehmer*innen sinkt am Abend (Borella, Ludwig, Dirk & de Ribaupierre, 2011; Hasher, Chung, May & Foong 2002; Mayhorn, Fisk & Whittle, 2002) und kann zu einem „response bias“ führen (Lewandowska et al, 2018). Ob der Zeitpunkt der Untersuchungstermine auch auf die Ausfallquote Einfluss nehmen könnte, bleibt allerdings unklar und wird deshalb explorativ erhoben.

² Einer dieser Umweltfaktoren sind kulturelle Unterschiede zwischen verschiedenen Gebieten oder Ländern. Es gibt groß angelegte internationale Studien, die versuchen, Versuchsteilnehmer*innen in verschiedenen Ländern zu rekrutieren, um die Daten dann zu vergleichen. Dabei werden unterschiedliche Ausschöpfungsquoten für die Daten verschiedener Staaten berichtet (Burholt et al., 2007; de Heer, 1999), auch wenn auf eine einheitliche Vorgehensweise geachtet wurde (Tolonen et al, 2015). Da diese Arbeit nicht darauf ausgelegt ist, Daten international zu erheben, kann in der Untersuchung nicht auf kulturelle Unterschiede eingegangen werden. Es ist aber empfehlenswert, bei der Rekrutierung auf die Besonderheiten der jeweiligen Population zu achten und darauf einzugehen.

Bei der Behandlung von Hepatitis C in Kliniken kommt es oft vor, dass Patient*innen ihre Termine nicht wahrnehmen und angeben, diese vergessen zu haben, oder dass die Termine für sie keine Priorität hatten. Poll, Allmark & Tod (2017) legen mit ihrer Forschung nahe, dass solche Angaben zu hinterfragen sind, da in Wahrheit oft andere Gründe dahinterstecken. In Interviews mit den Personen, die Termine nicht wahrgenommen hatten, identifizierten sie die tatsächlich hinter der Absage stehenden Gründe. Diese beinhalteten unter anderem die Kosten der Anreise, eine beschwerliche Anreise und den Zeitpunkt des Termins. Auch wenn die Zielgruppe dieser Untersuchung sehr spezifisch ist und vermutlich mit mehr psychosozialen Problemen zu kämpfen hat (Butt, McGuinness, Buller-Taylor, Mitchell, 2013; Poll, Allmark & Tod, 2017) als die Personen aus Stichproben üblicher psychologischer Untersuchungen, so lässt sich daraus doch ableiten, dass der benötigte Aufwand, um zu einem Termin zu gelangen, in die Entscheidung diesen auch wahrzunehmen, miteinfließt. Dies entspricht auch der oben bereits erläuterten „Social Exchange Theory“. Der Aufwand, der von einer Person betrieben werden muss, um zu einer Untersuchung zu gelangen, könnte also darauf Einfluss haben, ob diese Person den vereinbarten Termin auch tatsächlich wahrnimmt.

Hypothese 5.1.: Personen, die eine aufwendigere Anreise haben, fallen bei wissenschaftlichen Untersuchungen eher aus.

Von den Hepatitis C Patient*innen wurde auch der Zeitpunkt des Termins als Grund, nicht in der Klinik zu erscheinen, angegeben, weil sie beispielsweise mehrere Termine hatten und sich entscheiden mussten, welche sie priorisieren (Poll, Allmark & Tod, 2017). Auch wenn keine Terminüberschneidung vorliegt, kann die Geschäftigkeit – das Ausmaß in dem Personen objektiv bzw. subjektiv empfinden, durch verschiedene Beschäftigungen ausgelastet zu sein – (im Englischen: busyness) von Personen über die Teilnahme an Umfragen entscheidend sein. Darauf hat aber vor allem die subjektiv gefühlte Geschäftigkeit einen Einfluss, während objektive Maßstäbe keinen Schluss über den Zusammenhang zwischen Geschäftigkeit und der Teilnehmerate an Umfragen zulassen (Vercruyssen, Roose, Carton & Van De Putte, 2014). Zum Zeitpunkt, zu dem sich Versuchspersonen entscheiden müssen, ob sie auch tatsächlich zu einem Untersuchungstermin erscheinen, könnte die wahrgenommene Geschäftigkeit und das damit verbundene Stresslevel ebenso ein entscheidender Faktor sein. Dementsprechend lässt sich folgende Hypothese ableiten:

Hypothese 5.2.: Versuchspersonen, die eine hohe Geschäftigkeit empfinden, fallen bei wissenschaftlichen Untersuchungen eher aus.

Im Kapitel „Diskussion der Ergebnisse der Voruntersuchung“ wird auf Seite 43 unter „Ergänzung: allfällige Sonderfälle“ nachträglich Hypothese 5.3. hergeleitet und aufgestellt. Diese Hypothese wird ebenfalls dem Themenblock „Umweltfaktoren“ zugeordnet.

Voruntersuchung

Um die bereits aufgestellten Hypothesen zu ergänzen, wurde eine Voruntersuchung durchgeführt und explorative Daten gesammelt. Dabei wurde eine kurze Online-Umfrage entworfen, die an Psycholog*innen gerichtet war, die bereits mindestens einmal eine wissenschaftliche Laboruntersuchung an der Karl-Franzens-Universität Graz als Versuchsleiter*in begleitet hatten. Eine Einladung zu dieser Umfrage wurde über soziale Medien speziell an Psychologiestudierende der Karl-Franzens-Universität Graz gesendet. Zusätzlich wurden alle Personen per Mail kontaktiert, die im vorangegangenen Jahr eine Einladung zu einer Laboruntersuchung über „student.umfrage“, einem Mail-Verteiler der Karl-Franzens-Universität Graz, versendet hatten.

Ziele der Voruntersuchung

In der Voruntersuchung wurden drei Ziele explorativ untersucht, wobei zu jedem der drei Ziele jeweils eine Fragestellung formuliert wurde.

Erstens sollte festgestellt werden, wie groß der Anteil jener Personen bei wissenschaftlichen Untersuchungen ist, die nach Anmeldung nicht zu einem vereinbarten Untersuchungstermin erscheinen. Die sich daraus ergebende durchschnittliche Ausfallquote sollte dann als Basis dafür dienen, die nötige Stichprobengröße der Hauptuntersuchung zu bestimmen.

1. Fragestellung der Voruntersuchung: Wie hoch ist die durchschnittliche Ausfallquote bei Untersuchungen an der Karl-Franzens-Universität Graz?

Zweitens sollte erfasst werden, ob und wie stark Versuchsleiter*innen an der Karl-Franzens-Universität Graz, den Ausfall von Versuchspersonen als Problem wahrnehmen

2. Fragestellung der Voruntersuchung: Wird der Ausfalls von Versuchspersonen an der Karl-Franzens-Universität Graz als Problem wahrgenommen?

Das dritte Ziel der Voruntersuchung war es, einen Überblick darüber zu erlangen, wie Versuchspersonen von den Versuchsleiter*innen an der Karl-Franzens-Universität Graz rekrutiert und welche Untersuchungsmethoden eingesetzt werden. Außerdem sollte durch die Voruntersuchung ein umfassender Einblick in die Gestaltung des Forschungsprozesses bei psychologischen Untersuchungen an der Karl-Franzens-Universität Graz gewonnen werden.

3. Fragestellung der Voruntersuchung: Wie werden Untersuchungen an der Karl-Franzens-Universität Graz geplant und durchgeführt?

Methoden der Voruntersuchung

In der Folge werden Aufbau und Ablauf der Voruntersuchung, die Stichprobe und die Operationalisierung der Konstrukte näher beschrieben. Die Herleitung der Hypothesen findet sich oben im Unterkapitel „Mögliche Einflussfaktoren auf die Ausfallquote“.

Aufbau und Ablauf der Untersuchung

Für die Voruntersuchung wurde ein Online-Fragebogen ausgearbeitet und mithilfe der Web-Applikation LimeSurvey umgesetzt. Dieser Fragebogen umfasste 17 Fragen (siehe „Gestaltung der Voruntersuchung“ im Anhang) und wurde an ehemalige Versuchsleiter*innen per Mail ausgesendet. Die Fragen wurden in die vier Themenblöcke Thema (z.B.: „Bitte geben Sie zunächst das Thema der Untersuchung an, auf die Sie sich beim Beantworten der Fragen beziehen werden“), Rekrutierung (z.B.: „Welche Methoden haben Sie zur Rekrutierung von Versuchspersonen verwendet?“), Daten zur Stichprobe (z.B.: „Welche Versuchsdauer [in Minuten] haben Sie für Ihre Untersuchung bei der Rekrutierung angegeben?“ bzw. „Wie viele Versuchspersonen haben sich ursprünglich zu Ihrer Untersuchung angemeldet?“) und Besonderheiten der Untersuchung (z.B.: „Gab es bei Ihrer Untersuchung Besonderheiten [bezüglich Rekrutierung, Ablauf, etc.] durch die sich Ihre Untersuchung von anderen Untersuchungen unterscheidet?“) aufgeteilt. Dadurch sollte nachvollziehbar werden, wie Versuchsleiter*innen den Rekrutierungsprozess von der ersten Kontaktaufnahme bis zur Versuchsteilnahme von Proband*innen planen und gestalten. Die Daten zur Stichprobe und damit zur Ausfallquote wurden erhoben, um abschätzen zu können, wie groß die Stichprobe bei der Hauptuntersuchung mindestens sein muss, damit die Veränderung der Ausfallquote unter verschiedenen Bedingungen in der Hauptuntersuchung quantitativ messbar wird.

Voraussetzungen für die Teilnahme an dieser Voruntersuchung waren, dass die teilnehmenden Personen zuvor zumindest eine Untersuchung als Versuchsleiter*in begleitet hatten, dass die Versuchspersonen zu einem vorab vereinbarten Termin erscheinen mussten sowie, dass die Untersuchung an der Karl-Franzens-Universität Graz durchgeführt wurde.

Der Fragebogen begann mit dem Hinweis darauf, dass die Antworten „nur auf eine einzige bestimmte, von Ihnen [Anmerkung: den Befragten] durchgeführte Untersuchung, zu der die Versuchspersonen zu einem vorab vereinbarten Termin erscheinen mussten“, zu beziehen waren. Um ausschließen zu können, dass Untersuchungen mehrfach in die Ergebnisse einfließen, wurden zuerst die Themen der jeweiligen Untersuchungen erhoben sowie in welchem Rahmen

– Lehrveranstaltung, Abschlussarbeit oder Forschungstätigkeit – die Untersuchung durchgeführt wurde.

Danach wurden Multiple-Choice-Fragen dazu gestellt, wie bei der Rekrutierung von Versuchspersonen vorgegangen wurde. Zusätzlich zu den vorgegebenen Antwortkategorien gab es jeweils auch ein Textfeld, um Antworten zu ermöglichen, die nicht den vorgegebenen Kategorien entsprachen. Außerdem wurde erhoben, wie die Untersuchung beworben wurde, also welche Versuchsdauer und welche zu erwartenden Untersuchungsmethoden angegeben wurden und ob eine Belohnung als Anreiz für die Untersuchungsteilnahme versprochen wurde. Auch Teilnahme Kriterien wurden abgefragt, so es denn welche gab, um Besonderheiten der jeweils herangezogenen Stichproben aufzudecken.

Im nächsten Schritt wurde erhoben, wie viele Versuchspersonen sich zur Untersuchung ursprünglich angemeldet hatten, wie viele trotz Zusage den Termin nicht eingehalten hatten und wie viele Versuchspersonen tatsächlich am Versuch teilgenommen hatten. Außerdem war anzugeben, ob die angegebenen Zahlen exakt oder geschätzt waren. Dadurch sollte einerseits gewährleistet werden, dass Personen den Fragebogen nicht direkt abbrechen, sollten sie sich nicht an die exakten Zahlen erinnern und andererseits sollten die angegebenen Zahlen beim Auftauchen von Widersprüchen transparenter werden. Zusätzlich war von den Versuchsleiter*innen anzugeben, wie sehr sie den Ausfall von Versuchspersonen bei ihrer Untersuchung als Problem wahrgenommen hatten und ob sie die angemeldeten Versuchspersonen vor der Untersuchung telefonisch, per SMS, per E-Mail oder gar nicht an den vereinbarten Termin erinnert hatten.

Zum Abschluss folgten vier offene Fragen zu den Gründen des Ausfalls von Versuchspersonen. Hier sollte die Möglichkeit gegeben werden, auf Unterschiede oder Besonderheiten beim Aufbau und der Planung der Untersuchungen hinzuweisen und interessante Gedanken und Ideen zum Thema zu äußern.

Stichprobe

Die Einladung zur Teilnahme an der Voruntersuchung wurde per Mail an über 300 Psychologiestudierende verschickt. Außerdem wurde der Link zur Umfrage auf Facebook in einer Gruppe für die Psychologiestudierenden der Karl-Franzens-Universität Graz geteilt.

An der Voruntersuchung nahmen 30 Personen teil, die den Fragebogen vollständig ausfüllten. Es wurden keine demographischen Daten zu den Teilnehmer*innen an der Umfrage erhoben,

da die Umfrage sich ausschließlich auf die Durchführung einer wissenschaftlichen Untersuchung bezog. Aufgrund widersprüchlicher Angaben oder Angaben zu den Untersuchungen, die einen schlüssigen Vergleich der erhobenen Daten nicht mehr zuließen, wurden die Angaben zu acht Untersuchungen von der quantitativen Auswertung der Voruntersuchung ausgeschlossen.³ In der Stichprobe verblieben also die Daten zu 22 Untersuchungen, die zur statistischen Auswertung herangezogen wurden.

Von den 22 Untersuchungen, von denen berichtet wurde, waren elf Untersuchungen zu Masterarbeiten, sieben Untersuchungen zu Bachelorarbeiten, außerdem wurden zwei Untersuchungen im Zuge von universitären Forschungsprojekten und jeweils eine Untersuchung im Zuge einer Doktorarbeit bzw. im Zuge eines Seminars an der Universität durchgeführt.

Operationalisierung der Konstrukte

Alle Daten wurden mit LimeSurvey erhoben und dann zur weiteren Verarbeitung mit SPSS konvertiert. Zur Analyse der erhobenen Daten wurde SPSS 26 (Stand: 22. März 2021) verwendet. Im Zuge der Voruntersuchungen werden nur deskriptive Statistiken und sowohl Pearson- als auch Spearman-Korrelationen berichtet. Auf die Anwendung anderer statistischer Methoden wurde aufgrund der kleinen Stichprobengröße verzichtet.

Aufgrund des inferenzstatistischen Verfahrens zur Berechnung von Korrelationen erreichen Korrelationskoeffizienten bei kleinen Stichprobengrößen erst bei relativ gesehen größeren Koeffizienten das Signifikanzniveau. Da die Voruntersuchung explorativen Charakter hatte, wurden deshalb auch Korrelationskoeffizienten die ein Signifikanzniveau von $p < .10$ aufwiesen zur Analyse herangezogen. So sollte vermieden werden, mögliche für die Hauptuntersuchung interessante Erkenntnisse vorzeitig auszuschließen, da in diesem Fall ein β -Fehler (das fälschliche Beibehalten der Null-Hypothese) schwerer wiegt als ein α -Fehler (das fälschliche Verwerfen der Null-Hypothese). Diese Ergebnisse werden erst mit der statistischen Auswertung

³ Durch die Angabe des jeweiligen Themas der berichteten Studien wurde ersichtlich, dass zwei der Umfrageteilnehmer*innen an derselben Studie arbeiteten. Deshalb wurde jener der beiden zugehörigen Datensätze von der statistischen Datenauswertung ausgeschlossen, bei dem die Teilnehmerzahl der Untersuchung nur geschätzt wurde. Von den 29 verbliebenen Untersuchungen wurden wiederum vier ausgeschlossen, weil die Angaben zur Ausfallquote nicht nachvollziehbar waren sowie eine weitere, weil es einen Widerspruch bei den Angaben zu den Teilnehmer*innenzahlen gab. Bei zwei weiteren Untersuchungen wurden besondere Maßnahmen zur Rekrutierung eingesetzt, weshalb die zu diesen Untersuchungen zugehörigen Daten nicht mehr mit den anderen Daten der Stichprobe vergleichbar waren. Zusätzlich wurden die Daten zu einer der berichteten Untersuchungen bei der statistischen Auswertung aus der Stichprobe ausgeschlossen, weil sie die Versuchspersonen per Voruntersuchung vorselektierte, aber aus den Zahlen nicht klar ersichtlich war, ob die aussortierten Versuchspersonen zu den Ausgefallenen gezählt wurden oder nicht.

der Hauptuntersuchung verifiziert bzw. falsifiziert und können nur als Tendenzen angesehen werden.

Analyse der Ausfallquote

Um die Ausfallquote berechnen zu können, wurde jeweils die Zahl der zur Untersuchung angemeldeten Personen („Wie viele Versuchspersonen haben sich ursprünglich zu ihrer Untersuchung angemeldet?“), die Anzahl der trotz Anmeldung nicht zur Untersuchung erschienen Personen („Wie viele Personen haben zugesagt an ihrer Untersuchung teilzunehmen, haben den vereinbarten Termin jedoch nicht eingehalten“) und die tatsächliche Anzahl an Teilnehmer*innen an der Untersuchung erhoben („Wie viele Personen haben schlussendlich an ihrer Untersuchung teilgenommen?“). Da einige der Versuchsleiter*innen die zu ihrer Untersuchung angemeldeten Personen aufgrund von Teilnahme Kriterien vorselektierten und sich einige ihrer Versuchspersonen wieder abgemeldet haben dürften, wurde die Anzahl der Anmeldungen nicht zur Berechnung der Ausfallquote herangezogen. Dementsprechend wurde mit Hilfe der anderen beiden Kennwerte die Ausfallquote zu den jeweiligen Untersuchungen anhand folgender Formel berechnet:

$$\text{Ausfallquote} = \frac{\text{ausgefallene Versuchspersonen}}{\text{ausgefallene Versuchspersonen} + \text{erschienene Versuchspersonen}}$$

Analyse der Problemwahrnehmung

Zusätzlich zur Ausfallquote wurde auch erhoben, wie stark der Ausfall von Versuchspersonen von den Versuchsleiter*innen als Problem wahrgenommen wurde („Haben Sie den Ausfall von Versuchspersonen bei ihrer Untersuchung als Problem wahrgenommen?“). Diese Erhebung ist insofern interessant, weil eine hohe Ausfallquote nicht notwendigerweise dazu führen muss, dass diese als großes Problem wahrgenommen wird. Das ist beispielsweise dann der Fall, wenn die Untersuchung keine aufwendige Vorbereitung vor einzelnen Terminen benötigt bzw. keine teuren oder aufwendigen Methoden verwendet werden und die benötigte Stichprobengröße auch trotz ausfallenden Versuchspersonen noch leicht erreicht werden kann. Insofern ist eine Analyse davon, wie stark der Ausfall von Versuchspersonen als Problem wahrgenommen wird, hilfreich, um Besonderheiten von Untersuchungen aufdecken zu können. Denn solche Besonderheiten können dazu führen, dass der Ausfall von Versuchspersonen besonders schwer wiegt oder aber kaum von Bedeutung ist. Die Problemwahrnehmung wurde auf einer vierteiligen Likert-Skala („großes Problem“, „eher ein Problem“, „eher kein Problem“ und „kein Problem“) erfasst.

Hypothesen der Voruntersuchung

Da die Voruntersuchung explorativen Charakter hatte, wurden dazu keine gerichteten Hypothesen aufgestellt. In der Folge werden deshalb die im Kapitel „Ziele der Voruntersuchung“ definierten Fragestellungen erneut angeführt und erörtert, welche Variablen und Methoden zur Beantwortung der jeweiligen Fragestellung erhoben und verwendet wurden.

1. Fragestellung der Voruntersuchung: Wie hoch ist die durchschnittliche Ausfallquote bei Untersuchungen an der Karl-Franzens-Universität Graz?

Zur Erhebung der durchschnittlichen Ausfallquote wurden die Teilnehmer*innen der Umfrage gefragt, wie viele Personen sich zur Untersuchung angemeldet haben („Wie viele Versuchspersonen haben sich ursprünglich zu ihrer Untersuchung angemeldet?“), wie viele Personen ausgefallen sind („Wie viele Personen haben zugesagt an ihrer Untersuchung teilzunehmen, haben den vereinbarten Termin jedoch nicht eingehalten“) und wie viele Personen tatsächlich an der Untersuchung teilgenommen haben („Wie viele Personen haben schlussendlich an ihrer Untersuchung teilgenommen?“). Mit Hilfe dieser Daten wurde wie oben beschrieben, die Ausfallquote berechnet.

2. Fragestellung der Voruntersuchung: Wird der Ausfall von Versuchspersonen an der Karl-Franzens-Universität Graz als Problem wahrgenommen?

Es wurde erhoben, ob der Ausfall von Versuchspersonen als Problem wahrgenommen wurde („Haben Sie den Ausfall von Versuchspersonen bei ihrer Untersuchung als Problem wahrgenommen?“; Antwortkategorien: „großes Problem“, „eher ein Problem“, „eher kein Problem“ und „kein Problem“) und dann überprüft, ob es eine Korrelation mit den berichteten Ausfallquoten gibt.

3. Fragestellung der Voruntersuchung: Wie werden Untersuchungen an der Karl-Franzens-Universität Graz geplant und durchgeführt?

Um einen Überblick darüber zu bekommen, wie Untersuchungen an der Karl-Franzens-Universität Graz geplant und durchgeführt werden, wurde erhoben, welche Rekrutierungsmaßnahmen gesetzt wurden („Welche Methoden haben Sie zur Rekrutierung von Versuchspersonen verwendet?“ Antwortkategorien: Postings in sozialen Medien [Facebook, Twitter, etc.], Aushang an der Universität, Aushang außerhalb der Universität, Aussendung über einen E-Mail-Verteiler der Universität [z.B. student.umfrage], Aussendung über einen nichtuniversitären E-Mail-Verteiler, Kontaktaufnahme mit zuvor unbekanntem Personen von einer privaten E-Mail-

Adresse, telefonische Kontaktaufnahme mit zuvor unbekanntem Personen, persönliche Kontaktaufnahme mit zuvor unbekanntem Personen durch direktes Ansprechen, Vorsprechen in Lehrveranstaltungen zum Anwerben von Versuchspersonen, Rekrutierung von Verwandten, Freunden oder Bekannten, Sonstiges [mit Textfeld]), welche Versuchsdauer die Versuchsleiter*innen für die Untersuchung angegeben haben („Welche Versuchsdauer [in Minuten] haben Sie für Ihre Untersuchung bei der Rekrutierung angegeben?“), welche Untersuchungsmethoden angekündigt wurden („Welche zu erwartenden Untersuchungsmethoden haben Sie bei der Rekrutierung angegeben?“; Antwortkategorien: Persönlichkeitsfragebogen, andere Fragebögen, Intelligenztests, andere Leistungstests, physiologische Messungen [wie Blutdruck, Puls, etc.], fMRT, EEG, aktive Übungen [zu kognitiven Prozessen], aktive Teilnahme [Musik, Sport, Videos, etc.], apparativer Versuchsaufbau [z.B. Messung der Blickbewegungen], Computertestung, Gruppentestung, Einzeltestung, Sonstiges [mit Textfeld]), ob eine Belohnung bzw. ein Anreiz zur Teilnahme angeboten wurde („Haben Sie den Versuchspersonen bei der Rekrutierung eine Belohnung angeboten? [Wenn ja, welche?]“; Antwortkategorien: Versuchsscheine, Geld, Schokolade bzw. andere Süßigkeiten, Teilnahme an einer Verlosung unter den VersuchsteilnehmerInnen, Keine, Sonstiges [mit Textfeld]), ob und wie die Versuchspersonen vor der Untersuchung an den Termin erinnert wurden („Haben Sie die angemeldeten Personen vor der Untersuchung noch einmal an den Termin erinnert?“; Antwortkategorien: Nein, per Mail, per SMS, per Anruf, Sonstiges [mit Textfeld]), ob Versuchspersonen, die trotz fixer Zusage nicht an der Untersuchung teilgenommen haben, Gründe dafür angegeben haben und welche das waren („Haben Versuchspersonen, die trotz fixer Zusage an Ihrer Untersuchung teilzunehmen, nicht teilgenommen haben, Gründe für Ihre Abwesenheit angegeben? [Wenn ja, welche Gründe?]“; offenes Antwortformat), was für Gründe die Versuchsleiter*innen hinter dem Ausfall von Versuchspersonen vermuten („Was glauben Sie, könnten die Gründe sein, warum die ausgefallenen Versuchspersonen nicht zu Ihrer Untersuchung erschienen sind?“; offenes Antwortformat) und welche relevanten Ideen die Versuchsleiter*innen noch zum Ausfall von Versuchspersonen haben („Haben Sie sonst noch irgendwelche Gedanken zum Thema, die Sie für relevant erachten?“; offenes Antwortformat).

Kontrollitems:

Es gibt bestimmte Bedingungen, welche die Vergleichbarkeit der erhobenen Ausfallquoten beeinträchtigen. Deshalb wurde zusätzlich erhoben, welches Thema die jeweilige Untersuchung hatte („Bitte geben Sie zunächst das Thema der Untersuchung an, auf die Sie sich bei der Beantwortung der Fragen beziehen werden. [So soll vermieden werden, dass mehrfach Daten zur

selben Untersuchung erhoben werden]“; offenes Antwortformat), im Zuge welcher Forschungstätigkeit die Untersuchung durchgeführt wurde („Im Zuge welcher Forschungstätigkeit haben Sie diese Untersuchung durchgeführt?“; Antwortkategorien: Seminar: Praktikum für Forschungsmethodik, Bachelorarbeit, Masterarbeit, Universitäres Forschungsprojekt, Privates Forschungsprojekt, Sonstiges [mit Textfeld]), ob es Teilnahmebedingungen gab („Gab es Kriterien, die die Versuchspersonen erfüllen mussten, um an Ihrer Studie teilnehmen zu können? [Wenn ja, welche]“; offenes Antwortformat) und ob es bei der jeweiligen Untersuchung Besonderheiten gab, aufgrund der sie sich von anderen Untersuchungen unterscheidet („Gab es bei Ihrer Untersuchung Besonderheiten [bezüglich Rekrutierung, Ablauf, etc.] durch die sich Ihre Untersuchung von anderen Untersuchungen unterscheidet?“; offenes Antwortformat).

Ergebnisse der Voruntersuchung

Die Ergebnisse der Voruntersuchung, die die Ausfallquote betreffen, bilden ausschließlich die Situation an der Karl-Franzens-Universität Graz ab, da psychologische Untersuchungen an der Karl-Franzens-Universität Graz eine Besonderheit aufweisen. Psychologie-Studierende an der Karl-Franzens-Universität Graz müssen nachweisen, sich während des Bachelor-Studiums bei Untersuchungen im Ausmaß von mindestens 25 Stunden als Versuchsperson zur Verfügung gestellt zu haben. Für jede Untersuchungsteilnahme erhalten sie sogenannte „Untersuchungsscheine“, die sie benötigen, um ihr Bachelor-Studium abschließen zu können (Universität Graz Institut für Psychologie, n.d.). Diese Untersuchungsscheine werden für die Teilnahme an psychologischen Untersuchungen entsprechend der Dauer der Untersuchung ausgestellt und sollen gewährleisten, dass Personen, die am Institut für Psychologie forschen – etwa im Zuge ihrer Bachelorarbeit, Masterarbeit oder Dissertation – genügend Versuchsteilnehmer*innen finden. Durch diesen extrinsischen Motivationsfaktor könnten sich die Ausfallquoten bei psychologischen Untersuchungen an der Karl-Franzens-Universität Graz jedoch von anderen Stichproben unterscheiden.

Bei der Abgabe wurde dieser Masterarbeit ein USB-Stick mit den erhobenen Daten beigelegt. Auf Anfrage werden die Daten dieser Untersuchung von Em.Univ.-Prof. Dr.rer.nat. Dietrich Albert oder vom Autor dieser Arbeit gerne zur Verfügung gestellt.

Statistische Auswertung

1. Fragestellung der Voruntersuchung: Wie hoch ist die durchschnittliche Ausfallquote bei Untersuchungen an der Karl-Franzens-Universität Graz?

Die Versuchsleiter*innen der 22 in der Stichprobe verbliebenen Untersuchungen berichteten zwischen 25 und 210 Anmeldungen ($\bar{x} = 91.95$; $sd = 55.88$, normalverteilt), 20 bis 160 Teilnehmende ($\bar{x} = 72,60$; $sd = 40.10$, normalverteilt) und zwischen einer und 40 ausgefallenen Versuchspersonen ($\bar{x} = 9.32$; $sd = 10.00$, nicht normalverteilt). Die Widersprüche bei diesen Angaben ergeben sich durch Schätzungen bei den Angaben und durch das Vorselektieren der Stichprobe. Wie bereits unter „Analyse der Ausfallquote“ erläutert, wurden die Anzahl der Teilnehmer*innen und die Anzahl der ausgefallenen Versuchspersonen zur Analyse der Ausfallquote herangezogen. Dabei ergab sich eine durchschnittliche Ausfallquote von 10.66% bei einer Standardabweichung von 6.78%. Die Werte lagen im Bereich zwischen 2.44% und 22.22% (siehe Tabelle 1) und waren in dieser Stichprobe nicht normalverteilt (Schiefe = 0.444, Standardfehler = 0.491, Kurtosis = -1.357, Standardfehler = 0.953).

Deskriptive Statistiken zur Ausfallquote

Mittelwert	10,66%
Median	8,17%
Standard-Abweichung	6,78%
Niedrigste Ausfallquote	2,44%
Höchste Ausfallquote	22,22%

Tabelle 1: Deskriptive Statistiken zu den von 22 Versuchsleiter*innen berichteten Ausfallquoten.

2. Fragestellung der Voruntersuchung: Wird der Ausfall von Versuchspersonen an der Karl-Franzens-Universität Graz als Problem wahrgenommen?

Bei der Frage, ob sie den Ausfall von Versuchspersonen bei ihrer Untersuchung als Problem wahrgenommen haben, gaben fünf Personen an, dass dieser „kein Problem“ dargestellt hat. Die Hälfte der Versuchsleiter*innen – nämlich elf – gaben an, den Ausfall von Versuchspersonen als „eher kein Problem“ wahrgenommen zu haben. Fünf gaben an, dass dieser „eher ein Problem“ dargestellt hat und für eine bzw. einen Versuchsleiter*in war der Ausfall von Versuchspersonen „ein großes Problem“ (siehe Tabelle 2).

<i>Problemwahrnehmung</i>	<i>n</i>
<i>Ein großes Problem</i>	1
<i>Eher ein Problem</i>	5
<i>Eher kein Problem</i>	11
<i>Kein Problem</i>	5

Tabelle 2: Häufigkeit der Antworten auf die Frage „Haben Sie den Ausfall von Versuchspersonen bei Ihrer Untersuchung als Problem wahrgenommen?“ der 22 Versuchsleiter*innen.

3. Fragestellung der Voruntersuchung: Wie werden Untersuchungen an der Karl-Franzens-Universität Graz geplant und durchgeführt?

Bei der Erhebung der gesetzten Rekrutierungsmaßnahmen (siehe Tabelle 3) fiel auf, dass fast alle Befragten, nämlich 20 von 22, Postings in sozialen Medien getätigt hatten. Ausnahmen waren die Versuchsleiter*innen der Untersuchungen zur Doktorarbeit bzw. für das Seminar. Ebenso viele, also 20 der 22 Versuchsleiter*innen, hatten eine Aussendung über einen E-Mail-Verteiler der Universität getätigt. Allerdings sollte beachtet werden, dass ein Teil dieser Versuchsleiter*innen rekrutiert wurde, weil sie eine Aussendung über einen solchen E-Mailverteiler getätigt hatten. Die zwei Personen, die nicht darauf zurückgegriffen hatten, hatten Teilnahmebeschränkungen angegeben, aufgrund welcher eine solche Aussendung nicht sinnvoll gewesen wäre. 15 Versuchsleiter*innen griffen auf die Rekrutierung von Verwandten, Freunden oder Bekannten zurück, zwölf Personen machten einen Aushang an der Universität und drei davon auch außerhalb der Universität. Insgesamt elf Personen nahmen mit zuvor Unbekannten Kontakt per Telefon, Mail oder durch direktes Ansprechen auf. Fünf Personen besuchten Lehrveranstaltungen, um Teilnehmer*innen zu rekrutieren und eine Person schaltete ein Inserat im Internet und in einer Zeitung.

Es ließ sich beobachten, dass Rekrutierungsmaßnahmen, die innerhalb des universitären Rahmens durchgeführt werden konnten („Aushang an der Universität“, „Persönliche Kontaktaufnahme mit zuvor unbekanntem Personen durch direktes Ansprechen“ und „Vorsprechen in Lehrveranstaltungen zum Anwerben von Versuchspersonen“), häufiger genutzt wurden als Maßnahmen, die außerhalb dieses Rahmens gesetzt werden mussten („Aushang außerhalb der Universität“, „Aussendung über einen nichtuniversitären E-Mail-Verteiler“, „Kontaktaufnahme mit zuvor unbekanntem Personen von einer privaten E-Mail-Adresse“, „Telefonische Kontaktaufnahme mit zuvor unbekanntem Personen“, „Inserate im Internet bzw. einer Zeitung“). Auch die Gesamtanzahl an gesetzten Rekrutierungsmaßnahmen wurde berechnet. Eine

Person hatte nur zwei Maßnahmen, vier Personen hatten drei, zehn Personen hatten vier, fünf Personen hatten fünf und zwei Personen hatten sechs Maßnahmen zur Rekrutierung genutzt.

<i>Gesetzte Rekrutierungsmaßnahmen</i>	<i>n</i>
<i>Postings in sozialen Medien (Facebook, Twitter, etc)</i>	<i>20</i>
<i>Aussendung über einen E-Mail-Verteiler der Universität (z.B. student.umfrage)</i>	<i>20</i>
<i>Rekrutierung von Verwandten, Freunden oder Bekannten</i>	<i>15</i>
<i>Aushang an der Universität</i>	<i>12</i>
<i>Persönliche Kontaktaufnahme mit zuvor unbekanntem Personen durch direktes Ansprechen</i>	<i>9</i>
<i>Vorsprechen in Lehrveranstaltungen zum Anwerben von Versuchspersonen</i>	<i>5</i>
<i>Kontaktaufnahme mit zuvor unbekanntem Personen von einer privaten E-Mail-Adresse</i>	<i>4</i>
<i>Aushang außerhalb der Universität</i>	<i>3</i>
<i>Telefonische Kontaktaufnahme mit zuvor unbekanntem Personen</i>	<i>1</i>
<i>Inserate im Internet bzw. einer Zeitung</i>	<i>1</i>
<i>Aussendung über einen nichtuniversitären E-Mail-Verteiler</i>	<i>0</i>

Tabelle 3: Die Häufigkeit der gesetzten Rekrutierungsmaßnahmen durch die 22 Umfrageteilnehmer*innen.

Bezüglich der bei der Rekrutierung angegebenen Untersuchungsmethoden (siehe Tabelle 4) gaben 13 Versuchsleiter*innen an, Fragebögen erwähnt zu haben. Eine Person davon hatte Persönlichkeitsfragebögen, zehn hatten andere Fragebögen und zwei weitere Personen hatten beides angegeben. Drei Personen gaben bei der Rekrutierung an, Leistungstests durchzuführen, jedoch niemand Intelligenztests. Bei neun Untersuchungen wurde eine aktive Teilnahme entweder kognitiv oder physisch gefordert, bei acht wurden physiologische Messungen durchgeführt und bei vier waren Messungen mit technischen Hilfsmitteln wie fMRT, EEG, tDCS oder ein apparativer Versuchsaufbau zu erwarten. Vier der Versuchsleiter*innen gaben bei der Rekrutierung nur an, ob es sich um eine Gruppentestung, eine Einzeltestung oder eine Testung am Computer handelte, ohne jedoch zu erwähnen, welche Methoden zum Einsatz kamen. Die Anzahl an bei der Rekrutierung angegebenen Untersuchungsmethoden betrug in einem Fall keine, in insgesamt sieben Fällen nur eine einzige, in je sechs Fällen zwei bzw. drei Methoden, während je eine Person angab, vier bzw. fünf Methoden bei der Untersuchung einzusetzen.

<i>Angegebene Untersuchungsmethoden</i>	<i>n</i>
<i>Andere Fragebögen</i>	12
<i>Einzeltestung</i>	11
<i>Computertestung</i>	8
<i>Physiologische Messungen (Blutdruck, Puls, etc.)</i>	8
<i>Aktive Teilnahme (Musik, Sport, Videos, etc.)</i>	5
<i>Aktive Übungen (zu kognitiven Prozessen)</i>	4
<i>Persönlichkeitsfragebögen</i>	3
<i>Andere Leistungstests</i>	3
<i>Apparativer Versuchsaufbau (z.B. Messung der Blickbewegung)</i>	2
<i>Gruppentestung</i>	2
<i>tDCS</i>	1
<i>fMRT</i>	1
<i>Interaktion mit anderen Personen</i>	1
<i>Intelligenztests</i>	0
<i>EEG</i>	0

Tabelle 4: Die Häufigkeit der in den Ausschreibungen gemachten Angaben zu den, von den 22 Versuchsleiter*innen berichteten, Untersuchungen.

Von den 22 Versuchsleiter*innen hatten alle außer einer Person, die ein universitäres Forschungsprojekt betreute, den Proband*innen vorab Untersuchungsscheine versprochen. 14 Personen hatten vorab Süßigkeiten sowie fünf die Teilnahme an einer Verlosung angeboten. Eine monetäre Entschädigung für die Teilnahme an der Untersuchung boten drei Personen an, während jeweils eine Person ihre Untersuchung mit der Teilnahme an einer bezahlten Folgeuntersuchung, Gutscheinen oder einem gratis Kaffee bewarb (siehe Tabelle 5).

Drei Versuchsleiter*innen boten nur Untersuchungsscheine als Anreiz an sowie eine Person nur Geld. Elf Personen boten zusätzlich zu den Untersuchungsscheinen einen weiteren Anreiz an, während fünf Personen mit Untersuchungsscheinen, Schokolade und der Teilnahme an einer Verlosung für ihre Studie warben. Wenn Geld als Anreiz geboten wurde, wurde außer Untersuchungsscheinen nie ein weiterer Anreiz geboten.

<i>Gebotene Anreize</i>	<i>n</i>
<i>Untersuchungsscheine</i>	21
<i>Schokolade bzw. andere Süßigkeiten</i>	14
<i>Teilnahme an einer Verlosung unter den Versuchsteilnehmer*innen</i>	5
<i>Geld</i>	3
<i>Teilnahme an einer bezahlten Folgestudie (bei Eignung)</i>	1
<i>Kaffee</i>	1
<i>Keine Belohnung</i>	0

Tabelle 5: Die Häufigkeit der in den Ausschreibungen von den 22 Versuchsleiter*innen gebotenen Anreize für die Teilnahme an den jeweiligen Untersuchungen.

Nur zwei Versuchsleiter*innen gaben an, die zur Untersuchung angemeldeten Versuchspersonen nicht an den Termin der Untersuchung erinnert zu haben. Zwölf Personen erinnerten ihre Teilnehmer*innen vor dem Termin per E-Mail, während elf dies per SMS taten. Drei dieser Personen nutzten beides (siehe Tabelle 6). Dabei wurde nur erhoben wie Versuchspersonen vor der Untersuchung an den Termin erinnert wurden, nicht jedoch zu welchem Zeitpunkt die Erinnerung erfolgte.

<i>Zur Erinnerung genutzte Telekommunikationsmedien</i>	<i>n</i>
<i>Per E-Mail</i>	12
<i>Per SMS</i>	11
<i>Per Anruf</i>	0
<i>Keine Erinnerung</i>	2

Tabelle 6: Die Häufigkeit mit der die 22 Versuchsleiter*innen Medien zur Erinnerung der zur Untersuchung angemeldeten Versuchspersonen genutzt haben.

Für die 22 in dieser Studie verbliebenen Untersuchungen wurde während der Rekrutierung eine Versuchsdauer von zwischen 10 und 240 Minuten angegeben. Die durchschnittliche Dauer betrug 85,68 Minuten mit einer Standardabweichung von 60,16 Minuten (siehe Tabelle 7). Die Versuchsdauer war linkssteil verteilt (Schiefe = 1,208, Standardfehler = 0,491, Kurtosis = 1,631, Standardfehler = 0,953).

<i>Deskriptive Statistiken zur Versuchsdauer</i>	<i>Minuten</i>
<i>Mittelwert</i>	<i>85,68</i>
<i>Median</i>	<i>90</i>
<i>Standard-Abweichung</i>	<i>60,16</i>
<i>Kürzeste Versuchsdauer</i>	<i>10</i>
<i>Längste Versuchsdauer</i>	<i>240</i>

Tabelle 7: Deskriptive Statistiken zu der von den 22 Versuchsleiter*innen bei den Ausschreibungen zu den Untersuchungen angegebenen Versuchsdauer. Angaben in Minuten.

Des Weiteren wurden die Korrelationskoeffizienten der Ausfallquote, der Problemwahrnehmung, der Versuchsdauer, der Anzahl der gesetzten Rekrutierungsmaßnahmen, der Anzahl der angegebenen Untersuchungsmethoden und der Anzahl der verschiedenen, gebotenen Anreize miteinander berechnet (siehe Tabelle 8). Nachdem die Problemwahrnehmung nur ordinal skaliert war, wurden für alle Korrelationen mit der Problemwahrnehmung Rangkorrelationskoeffizienten nach Spearman berechnet. Zur Berechnung aller weiteren Koeffizienten wurde die Pearson Korrelation angewandt.

Signifikante Korrelationen zeigten sich zwischen der Ausfallquote und damit wie sehr der Ausfall von Versuchspersonen als Problem wahrgenommen wurde ($r = -.799$, $p < .05$) sowie zwischen der Ausfallquote und der Anzahl der gebotenen Anreize ($r = -.404$, $p < .10$).⁴ Wenn der Ausfall von Versuchspersonen eher als Problem wahrgenommen wurde, war auch die Ausfallquote höher und umso mehr unterschiedliche Anreize den Versuchspersonen angeboten wurden, umso niedriger war die Ausfallquote. Die Anzahl an unterschiedlichen Anreizen korrelierte negativ mit der angegebenen Versuchsdauer ($r = -.389$, $p < .10$). Je niedriger die Versuchsdauer, umso höher war die Anzahl der gebotenen Anreize. Außerdem zeigte sich, dass bei mehr gesetzten Rekrutierungsmaßnahmen auch mehr verschiedene Anreize geboten wurden ($r = .432$, $p < .05$). Es gilt jedoch zu beachten, dass diese Korrelationen keinerlei Aufschluss über die Kausalität der Zusammenhänge zulassen.

⁴ Wie im Kapitel „Methoden der Voruntersuchung“ beschrieben, erreichen Korrelationskoeffizienten bei kleinen Stichprobengrößen aufgrund des inferenzstatistischen Verfahrens zur Berechnung von Korrelationen erst bei relativ gesehen größeren Koeffizienten das Signifikanzniveau. Da die Voruntersuchung explorativen Charakter hatte, wurden deshalb auch Korrelationskoeffizienten, die ein Signifikanzniveau von $p < .10$ aufwiesen, zur Analyse herangezogen. So sollte vermieden werden, mögliche für die Hauptuntersuchung interessante Erkenntnisse vorzeitig auszuschließen, da in diesem Fall ein β -Fehler (das fälschliche Beibehalten der Null-Hypothese) schwerer wiegt als ein α -Fehler (das fälschliche Verwerfen der Null-Hypothese). Diese Ergebnisse werden erst mit der statistischen Auswertung der Hauptuntersuchung verifiziert bzw. falsifiziert und können nur als Tendenzen angesehen werden.

Korrelationskoeffizienten der Untersuchungsvariablen

	Ausfallquote in Prozent	Problemwahrnehmung (in Rängen 1 bis 4)	Versuchsdauer in Minuten	Anzahl der Rekrutierungsmaßnahmen	Anzahl der Untersuchungsmethoden	Anzahl der Anreize
Ausfallquote in Prozent	-	-.799** p<.001	.286 p=.197	-.049 p=.829	-.107 p=.637	-.404* p=.062
Problemwahrnehmung (in Rängen 1 bis 4)	-.799** p<.001	-	-.263 p=.237	.009 p=.968	-.046 p=.838	.259 p=.245
Versuchsdauer in Minuten	.286 p=.197	-.263 p=.237	-	.206 p=.357	.258 p=.246	-.389* p=.074
Anzahl der Rekrutierungsmaßnahmen	-.049 p=.829	.009 p=.968	.206 p=.357	-	-.132 p=.558	.432** p=.045
Anzahl der Untersuchungsmethoden	-.107 p=.637	-.046 p=.838	.258 p=.246	-.132 p=.558	-	-.312 p=.158
Anzahl der Anreize	-.404* p=.062	.259 p=.245	-.389* p=.074	.432** p=.045	-.312 p=.158	-

Tabelle 8: Die Korrelationskoeffizienten zwischen der Ausfallquote (zwischen 2.44% und 22.22%), Problemwahrnehmung (zwischen 1 und 4; Ein großes Problem = 1, eher ein Problem = 2, eher kein Problem = 3, kein Problem = 4), Versuchsdauer in Minuten (zwischen 10 und 240 Minuten), der Anzahl der gesetzten Rekrutierungsmaßnahmen (zwischen 2 und 6 Maßnahmen), der Anzahl der angegebenen Untersuchungsmethoden (zwischen 0 und 5 Untersuchungsmethoden), der Anzahl der verschiedenen gebotenen Anreize (zwischen 1 und 3 Anreizen); n = 22, *p < .10, **p < .05

Durch die Überprüfung der zugehörigen Scatterplots und Kreuztabellen konnte ausgeschlossen werden, dass es nicht-lineare Zusammenhänge gibt.

Qualitative Auswertung

Zur qualitativen Auswertung der Antworten auf die offenen Fragen wurde die Grounded Theory (Strauss & Corbin, 1996; Strübing, 2014) herangezogen. Dabei handelt es sich um einen Forschungsansatz, der darauf aufbaut, qualitativ erhobene Daten zu kategorisieren und daraus erlangte Erkenntnisse direkt in den weiteren Forschungsverlauf einzuarbeiten. Alle weiteren im Forschungsverlauf erlangten Erkenntnisse werden in der Folge zum Reflektieren und Evaluieren

ren der Kategorien herangezogen, was sich dann wiederum auf den Forschungsverlauf auswirkt. Dadurch ergibt während der Auswertung der qualitativen Daten eine flexible Kategorisierung der Daten, wodurch das Verharren in starren Denkmustern verhindert werden soll. Mithilfe dieses Ansatzes wurden die Antworten der Versuchsleiter*innen zu den offenen Fragen kategorisiert, um daraus Ideen für die weitere Untersuchung abzuleiten.

Die erste offene Frage der Voruntersuchung befasste sich damit, welche Gründe die Versuchspersonen für den Ausfall bei der Untersuchung angaben. Laut den Versuchsleiter*innen gaben die ausgefallenen Versuchspersonen drei verschiedene Arten von Gründen an, warum sie nicht zum vereinbarten Termin erschienen seien. Diese Gründe bezogen sich auf die persönliche Terminkoordinierung (Terminüberschneidungen), Zwischenfälle (Irrtum im Datum, Krankheit, Panne, etc.) oder unzuverlässiges Verhalten (Vergessen des Termins, Verschlafen, Interessensverlust, etc.). Auffallend war, dass der einzige Grund, der sich direkt auf die Untersuchung bezog, der Verlust des Interesses war.

Die Versuchsleiter*innen äußerten außerdem ihre eigenen Ideen, warum Versuchspersonen ausgefallen waren. Als mögliche Gründe wurden die Untersuchungsdauer, die Untersuchungsmethoden, Umweltfaktoren, Unzuverlässigkeit und Interessensverlust aufgrund des Themas bzw. eines zu geringen Anreizes genannt. Die Versuchsleiter*innen erachteten also Umweltfaktoren, die Umstände der Untersuchung bzw. die Persönlichkeit der Versuchspersonen als verantwortlich für den Ausfall von Versuchspersonen.

Auf die Frage, ob die Versuchsleiter*innen noch relevante Gedanken zum Thema hätten, antworteten einige Personen, dass sie Erinnerungen an den Termin, die Form der Kontaktaufnahme (zeitnahe Reaktion auf Anmeldungen, klare Angaben bezüglich Ort und Zeit) und die Vermittlung eines Verbindlichkeitsgefühls gegenüber den Versuchspersonen als wichtig erachteten, um die Ausfallquote zu reduzieren. Als Probleme gaben sie den zur Testung gewählten Zeitpunkt (im Winter, zu Ferienzeiten) bzw. eine Teilung der Testung in Vor- und Hauptuntersuchung an, da viele geeignete Versuchspersonen auf eine Einladung zur Hauptuntersuchung nicht mehr reagiert hätten. Den Ausfall von Versuchspersonen erachteten sie vor allem als Problem für die Planbarkeit der Untersuchung.

Diskussion der Ergebnisse der Voruntersuchung

In der Voruntersuchung zeigte sich, dass Versuchsleiter*innen vermehrt auf Rekrutierungsmethoden setzten, die innerhalb des universitären Kontexts angewandt werden können. Außerdem waren Maßnahmen beliebter, die für den bzw. die Versuchsleiter*in weniger Aufwand bedeuten. Dass in 15 der 22 Fälle auf die Rekrutierung von Verwandten, Freunden und Bekannten gesetzt wurde, könnte auf eine positive Verzerrung der Ausfallquote hindeuten, da diese Personengruppe durch das Näheverhältnis zum bzw. zur Versuchsleiter*in zuverlässiger sein sollte, als unbekannte Personen. In der Stichprobe lag die durchschnittliche Ausfallquote jedoch sowohl für jene Untersuchungen, die Verwandte, Freunde und Bekannte untersuchten (10.32%), als auch für jene die das nicht taten (11.39%), nah am Durchschnitt der gesamten Stichprobe. Der Einfluss eines Näheverhältnisses zum bzw. zur Versuchsleiter*in auf den Ausfall von Versuchspersonen wird in der Hauptuntersuchung explorativ untersucht.

Bei den Untersuchungsmethoden wurde am häufigsten auf Fragebögen und Einzeltestungen gesetzt. Als Anreiz wurden in fast allen Fällen Untersuchungsscheine angeboten, was wohl den Besonderheiten des Psychologiestudiums der Karl-Franzens-Universität geschuldet ist. Zusätzlich wurde jedoch meist nur mit Süßigkeiten geworben. Da Süßigkeiten als Anreiz nie gemeinsam mit Geld angeboten wurden, bedeutet das, entsprechend der festgestellten Korrelation zwischen der Ausfallquote und der Anzahl der gebotenen Anreize, dass vor allem die Kombination aus Untersuchungsscheinen, Süßigkeiten und der Teilnahme an einer Verlosung mit einer niedrigeren Ausfallquote einherging.

Fast alle Versuchsleiter*innen sendeten vor dem Termin Erinnerungen an die Versuchspersonen. Dabei verwendeten sie E-Mails und SMS etwa gleich häufig. Nur zwei Personen versendeten keine Erinnerungen, diese erzielten dennoch unterdurchschnittliche Ausfallquoten von 5% bzw. 6.33%. Da eine statistische Auswertung durch die kleine Gruppengröße nicht möglich war, bleibt eine Beurteilung der Auswirkungen von Erinnerungen vorerst aus.

Dass die Versuchsdauer in der Stichprobe linkssteil verteilt ist, weist darauf hin, dass Untersuchungen, wenn möglich, eher kurz gehalten wurden. Ein Zusammenhang mit der Ausfallquote ($r = .286$, $p = .197$, $n = 22$) konnte nicht festgestellt werden, was aber der kleinen Stichprobe geschuldet sein könnte. Die dazu bereits beschriebene Hypothese, dass die Ausfallquote bei kurzen Untersuchungen geringer ist als bei langen Untersuchungen, wird in der Hauptuntersuchung erneut überprüft.

Mehr als zwei Drittel der Versuchsleiter*innen, insgesamt 16 von 22, sahen den Ausfall von Versuchspersonen als „kein Problem“ oder „eher kein Problem“ an. Es zeigte sich jedoch eine starke Korrelation mit der Ausfallquote ($r = -.799$, $p < .001$). Das heißt eine höhere Ausfallquote ging mit der Wahrnehmung des Ausfalls als größeres Problem bzw. eine niedrigere Ausfallquote mit der Wahrnehmung des Ausfalls als geringeres Problem einher. In vier Fällen, in denen der Ausfall von Versuchspersonen als „eher kein Problem“ wahrgenommen wurde, konnte dennoch eine überdurchschnittliche Ausfallquote zwischen 12,5% und 20% nachgewiesen werden. Eine hohe Ausfallquote muss also nicht unbedingt ein Problem darstellen, sofern die endgültige Stichprobe die Bedingungen der Forscher*innen erfüllt und kein großer zusätzlicher Aufwand entsteht.

Bei der qualitativen Auswertung der drei offenen Fragen wurden die Persönlichkeit der Versuchspersonen, insbesondere unzuverlässiges Verhalten und Probleme mit der persönlichen Terminkoordinierung, unbeeinflussbare Zwischenfälle durch Umwelt und persönliches Umfeld sowie die Umstände der Untersuchung als Gründe für den Ausfall genannt. Ein Großteil dieser Faktoren wurde bei der Bildung der Hypothesen bereits berücksichtigt. Neu hinzu kommen Zwischenfälle im persönlichen Umfeld sowie die persönliche Terminkoordinierung. Nachdem Zwischenfälle dadurch gekennzeichnet sind, dass sie unvorhersehbar sind, bilden sie somit auch einen unkontrollierbaren Faktor, der aber bei jeder Untersuchung relativ gesehen in einem ähnlichen Ausmaß auftreten sollte.

Die empfundene Geschäftigkeit (im Englischen: busyness) ist den Problemen bei der persönlichen Terminkoordinierung ähnlich. Die persönliche Terminkoordinierung ist in diesem Fall als objektive Geschäftigkeit zu sehen und damit anders zu betrachten. Wenn es zu einer Terminüberschneidung kommt, obliegt es der Versuchsperson zu entscheiden, welcher Termin Priorität bekommt. Wird ein bereits vereinbarter Untersuchungstermin zugunsten eines neuen Termins vernachlässigt, so ist dies wohl als unzuverlässiges Verhalten zu werten und damit der Zuverlässigkeit zuzuordnen.

Ergänzung: Auffällige Sonderfälle

In der Folge gilt es, auch Besonderheiten einzelner Untersuchungen Aufmerksamkeit zu widmen, da diese weitere Rückschlüsse auf mögliche Einflussfaktoren auf die Ausfallquote zulassen könnten.

Bei der einzigen Untersuchung, bei der der Ausfall von Versuchspersonen als großes Problem wahrgenommen wurde, lag die Ausfallquote nah am Durchschnitt (10,71%, 56 Anmeldungen, sechs davon ausgefallen). Dies ist wohl darauf zurückzuführen, dass trotz hoher Korrelation mit der Ausfallquote, die Problemwahrnehmung ein subjektives Maß ist. Dass es in diesem Fall zwischen diesen beiden Maßen so eine große Diskrepanz gab, lag möglicherweise an der Erwartungshaltung des bzw. der Versuchsleiter*in. Bei dieser Untersuchung wurde angegeben, dass der Anreiz zur Teilnahme hoch war (Verlosung eines dreitägigen Skiurlaubs für zwei Personen). Außerdem wurden Erinnerungen per SMS und per Mail versandt. Durch die gesetzten Maßnahmen könnte sich die Person mehr Zuspruch erwartet und Probleme gehabt haben, die angestrebte Stichprobengröße zu erreichen, wodurch der Ausfall von Proband*innen stärker ins Gewicht fiel. Auf eine anfänglich hohe Ausfallquote könnte auch durch zusätzlich ergriffene Maßnahmen (z.B. Erinnerungen, zusätzliche Rekrutierungsmaßnahmen) im Laufe der Untersuchung reagiert worden sein, nachdem sie als großes Problem wahrgenommen wurde.

Eine ähnliche Einflussnahme lässt sich auch bei der Untersuchung mit der niedrigsten Ausfallquote (1,96%) beobachten. Da für die Untersuchung nur eine Versuchsdauer von 15 Minuten angesetzt war, hatte der bzw. die Versuchsleiter*in sich dazu entschlossen, zusätzlich Versuchspersonen spontan auf der Straße zu rekrutieren und direkt zu untersuchen. Die Ausfallquote konnte somit gesenkt werden, da kein Termin mehr vereinbart werden musste. Da ein Ausfall von Versuchspersonen durch diese Vorgehensweise verhindert wurde, wurde die Untersuchung aus den Analysen der Ausfallquote ausgeschlossen. Dennoch ist festzuhalten, dass die Ausfallquote nicht als ein Maß zu betrachten ist, das durch die Planung der Untersuchung vorherbestimmt ist, sondern als variable Größe, auf die im Verlauf des Forschungsprozesses noch Einfluss genommen werden kann.

Interessant ist diesbezüglich auch die Beobachtung, dass die Versuchsleiter*innen, wenn es um mögliche Gründe für den Ausfall von Versuchspersonen ging, situationsbezogen attribuierten und die Verantwortung bei den Umständen der Untersuchung (Umweltfaktoren, Untersuchungsmethoden, Versuchsdauer, etc.) suchten. Bei der Frage zu relevanten Gedanken zum Thema nutzten sie jedoch personenbezogene Attribuierung und machten auf Faktoren auf-

merksam, die sie persönlich beeinflussen konnten (Kontaktaufnahme, Erinnerungen an Termine, Erzeugen eines Verbindlichkeitsgefühls). Dies folgt dem Muster der selbstwertdienlichen Attribution (Aronson, Wilson & Akert, 2008, S.116 ff), laut dem Misserfolge – in diesem Fall erhöhter Ausfall von Versuchspersonen – eher situationsbezogen attribuiert werden und der Erfolg von Personen – hier die Reduktion des Ausfalls von Versuchspersonen – sich selbst zugeschrieben wird. Durch eine zeitnahe kritische Auseinandersetzung mit dem Ausfall von Versuchspersonen könnten Versuchsleiter*innen jedoch schon während der Datenerhebung reagieren und Maßnahmen setzen, um den Ausfall von Versuchspersonen zu reduzieren, wie Erinnerungen an den Termin zu versenden oder aber die Art der Kontaktaufnahme zu verbessern.

Zwei Personen berichteten von der Vermutung, dass ihre Untersuchungen für Versuchspersonen besonders ansprechend gewesen seien, da sie für die Teilnahme an ihrer Untersuchung einen besonders hohen Anreiz geschaffen bzw. ein besonders interessantes Thema gewählt hatten. Die Ausfallquoten der beiden Untersuchungen betrugen 10,71% und 8,57% und lagen damit im Bereich der mittleren 20% der angegebenen Ausfallquoten. Das ist möglicherweise ein Hinweis darauf, dass die ansprechende Aufbereitung einer Untersuchung nur für die Anzahl an Anmeldungen zur Untersuchung relevant ist, jedoch nicht dafür, wie viele der angemeldeten Personen tatsächlich zur Untersuchung erscheinen. In der Stichprobe korrelierte die Anzahl der gebotenen Anreize nicht mit der Anzahl der Anmeldungen zu den Untersuchungen ($r = .149$, n.s.). Die Anzahl der gebotenen Anreize gibt jedoch auch keinen Aufschluss darüber, wie ansprechend die einzelnen gebotenen Anreize waren.

Einen Aufschluss darüber könnten die durchschnittlichen Ausfallquoten bei verschiedenen gebotenen Anreizen geben. In den Fällen, in denen den Versuchspersonen Geld geboten wurde, betrug die Ausfallquote im Schnitt 16,93%, wenn Süßigkeiten geboten wurden, betrug sie nur 7,78%. Wurde die Teilnahme an einem Gewinnspiel beworben, betrug die Ausfallquote 6,51%. Durch die geringe Stichprobengröße war ein statistischer Vergleich einzelner Gruppen hier nicht sinnvoll. Wie diese Unterschiede zu beurteilen sind, bleibt offen. An dieser Stelle sei auch nochmals daran erinnert, dass, wenn Geld als Anreiz geboten wurde, dies (mit Ausnahme von Untersuchungsscheinen) immer der einzige Anreiz war. Hingegen wurde oft mit Süßigkeiten und Gewinnspielteilnahmen gemeinsam für Untersuchungen geworben.

Diese Unterschiede bei den Ausfallquoten erinnern an ein klassisches Experiment zur einstellungskonträren Argumentation (im Englischen: counterattitudinal advocacy) von Festinger und Carlsmith (1959). Bei diesem Experiment mussten die Proband*innen zunächst eine Stunde

lang einer langweiligen Tätigkeit nachgehen und dann der nächsten vermeintlichen Versuchsperson (die jedoch in das Experiment eingeweiht war) diese Tätigkeit als spannend verkaufen. Dafür erhielten sie entweder einen Dollar oder 20 Dollar. Festinger und Carlsmith stellten fest, dass die Versuchspersonen, die nur einen Dollar für ihre Falschaussage erhielten, ihre persönliche Einstellung gegenüber der Tätigkeit stärker an die Lüge anpassten als jene, die 20 Dollar dafür erhielten. Dies wird damit erklärt, dass die 20 Dollar entsprechend der Dissonanztheorie (siehe auch Aronson, Wilson & Akert, 2008, S.161 ff.) eine gute Rechtfertigung für die Lüge waren. Personen, die nur einen Dollar bekamen, redeten sich jedoch aufgrund der entstandenen Dissonanz ein, dass die Tätigkeit spannender gewesen sei, als sie es tatsächlich war. Ähnlich zu diesem Experiment könnten Versuchspersonen, die nur einen kleinen Anreiz für eine Untersuchungsteilnahme erhalten, eine alternative Erklärung für ihre Teilnahme haben bzw. für sich finden, wie etwa persönliches Interesse, Hilfsbereitschaft oder den Wunsch, einen Beitrag zur Wissenschaft zu leisten. Da Einstellungsänderungen langfristig wirken, bliebe somit auch die Motivation der Versuchspersonen aufrecht, tatsächlich zum Termin zu erscheinen. Diese, durch den geringen Anreiz erzeugte Einstellungsanpassung, könnte also in weiterer Folge bewirken, dass Personen, die nur einen geringen Anreiz für die Teilnahme erhalten, seltener ausfallen als Personen, die einen wertvolleren Anreiz geboten bekommen.

Hypothese 4.6.: Die Ausfallquote ist bei Untersuchungen kleiner, wenn die Anreize einen kleineren Wert haben, als wenn wertvollere Anreize geboten werden.⁵

Zu einer weiteren Untersuchung, die von der Analyse der Ausfallquote ausgeschlossen wurde, wurde eine Ausfallquote von 37,04% angegeben. Bei dieser Untersuchung wurde die Stichprobe per Voruntersuchung vorselektiert. Durch die angegebenen Zahlen war nicht klar ersichtlich, ob die aussortierten und die ausgefallenen Versuchspersonen gemeinsam als ausgefallene Versuchspersonen berichtet wurden oder nicht. An diesem Fall ist jedoch interessant, dass der bzw. die Versuchsleiter*in angab, dass Schneefall und Kälte ein möglicher Grund für den Ausfall von Versuchspersonen im Testzeitraum waren. Schlechte Wetterbedingungen könnten somit einen Grund für einen wahrscheinlicheren Ausfall von Versuchspersonen bilden.

Hypothese 5.3.: Bei schlechtem Wetter fallen Versuchspersonen eher aus.⁶

⁵ Diese Hypothese wurde dem Themenblock „Motivationale Faktoren“ zugeordnet.

⁶ Diese Hypothese wurde dem Themenblock „Umweltfaktoren“ zugeordnet.

Bezüglich der verwendeten Rekrutierungsmethoden fielen zwei Gruppen besonders auf. Personen, die durch direktes Ansprechen rekrutierten, erzielten eine durchschnittliche Ausfallquote von 8,21%, während Personen, die auch Aushänge außerhalb der Universität tätigten, bei ihren Untersuchungen eine Ausfallquote von 14,13% hatten. Eine Person war in beiden Gruppen vertreten und schrieb zur Rekrutierung: „[Bei der] Form der Kontaktaufnahme [ist] meiner Meinung nach wichtig, ob ProbandInnen tatsächlich erscheinen. Je persönlicher, desto höher das ‚Commitment‘“ Ein persönlicher Kontakt könnte also entscheidend sein, um Versuchspersonen ein Verbindlichkeitsgefühl zu vermitteln und die Ausfallquote somit zu senken.

Hypothese 2.1.: Persönlicher Kontakt bei der Rekrutierung senkt die Ausfallquote.⁷

Probleme bei der Datenerhebung der Voruntersuchung

Einige der vorangegangenen statistischen Analysen sind aufgrund der kleinen Stichprobe nur vorsichtig zu interpretieren. Die kleine Stichprobengröße ist auch auf Fehler in der Datenerhebung zurückzuführen. In der Folge wird kurz auf diese hingewiesen, damit diesen Fehlern in nachfolgenden Studien vorgebeugt werden kann.

Einerseits ist die kleine Stichprobe darauf zurückzuführen, dass die Voraussetzungen zur Teilnahme an der Untersuchung relativ streng und geeignete ehemalige Versuchsleiter*innen deshalb nur schwer zu finden waren. Trotz über 300 per E-Mail versendeter Einladungen sowie der Verbreitung des Fragebogens über soziale Medien öffneten nur 47 Personen den Online-Fragebogen, wovon 17 diesen abbrachen. Zukünftige Befragungen von Versuchsleiter*innen sollten breiter angesetzt werden, indem sie Daten beispielsweise interuniversitär erheben oder über einen längeren Zeitraum direkt andere Studien mit entsprechendem Forschungsdesign begleiten, damit eine größere Zahl an ehemaligen Versuchsleiter*innen erreicht werden kann. In manchen Fällen kann es auch sinnvoll sein, Versuchspersonen an Stelle von Versuchsleiter*innen zu befragen, so wie es auch bei der Hauptuntersuchung dieser Arbeit gehandhabt wurde.

Dass die Stichprobengröße durch den Ausschluss einzelner Untersuchungen weiter verringert wurde, lag wohl an den Fragen zur Ausfallquote, die zu Missverständnissen bei den Angaben geführt haben dürften. Eine konkrete, getrennte Abfrage von Abmeldungen und ausgefallenen Personen sollte zur klaren Interpretierbarkeit der Daten beitragen. Die verwendeten Formulierungen sollten zwar klar auf die gewünschten Daten hinweisen, könnten aber durch ihre Komplexität nicht genau gelesen worden sein und ließen in Sonderfällen dennoch Spielraum zur

⁷ Diese Hypothese wurde dem Themenblock „Gestaltung des Rekrutierungsprozesses“ zugeordnet.

Interpretation. Dementsprechend sollten Fragen zu solchen Kennzahlen in zukünftigen Befragungen kurz, aber prägnant formuliert werden.

Einen Schwachpunkt bei den erhobenen Daten bildet außerdem, dass nicht erhoben wurde, zu welchem Zeitpunkt die Versuchspersonen an den Untersuchungstermin erinnert wurden. Ein unterschiedlicher Zeitraum vor der Untersuchung könnte zu einer unterschiedlichen Wirksamkeit der Erinnerungen führen. Zusätzlich bleibt bei Versuchsleiter*innen, die ihre Versuchspersonen sowohl per Mail als auch per SMS an den Termin erinnert haben, unklar, ob sie die Versuchspersonen per Mail und SMS kontaktiert haben oder nur einen dieser Kommunikationskanäle wählten. Bei der Hauptuntersuchung wurden die Erinnerungen immer einheitlich im gleichen Zeitabstand vor dem Untersuchungstermin versandt, da vorrangig erforscht wurde, ob Erinnerungen überhaupt wirksam sind. Die Versuchspersonen durften selbst auswählen, ob sie ihre Mail-Adresse oder Telefonnummer teilen wollten. Die jeweiligen Angaben wurden zum Versenden von Erinnerungen verwendet. Zukünftige Forschung könnte sich der Frage widmen, ob sich die Effektivität von Erinnerungen an Untersuchungstermine ändert, je nachdem wie lange vor dem Termin sie versandt wurden und ob es einen Unterschied macht, welcher Kommunikationskanal gewählt wurde.

Implikationen für die Hauptuntersuchung

Die Ziele der Voruntersuchung konnten trotz kleiner Stichprobe erreicht werden. Die durchschnittlich erhobene Ausfallquote für Untersuchungen an der Karl-Franzens-Universität Graz beträgt 10,66% mit einer Standardabweichung von 6,78%. Jede/r Versuchsleiter*in gab an, dass Versuchspersonen bei ihrer/seiner Untersuchung ausgefallen sind. Dass 6 der 22 Versuchsleiter*innen den Ausfall von Versuchspersonen eher als Problem oder als großes Problem wahrgenommen haben, unterstreicht die Relevanz der Forschung zu diesem Thema. Entsprechend dem zentralen Grenzwerttheorem sollte die Stichprobe der Hauptuntersuchung mindestens 30 Personen enthalten, die ausgefallen sind, damit sinnvolle statistische Analysen möglich werden. Bei Erreichen der durchschnittlichen Ausfallquote müsste die Gesamtstichprobe 282 Personen enthalten, damit 30 Personen ausfallen. Um diese Zahl entsprechend der erhobenen Standardabweichung der Ausfallquote mit 68,27-prozentiger Wahrscheinlichkeit zu erreichen, muss die Stichprobe zwischen 172 und 774 Personen enthalten.

Es konnte auch ein Überblick darüber gewonnen werden, welche Rekrutierungsmaßnahmen Versuchsleiter*innen an der Karl-Franzens-Universität Graz setzen und welche Untersu-

chungsmethoden sie einsetzen. Beides wurde bei der Planung der Hauptuntersuchung berücksichtigt. Auf den genauen Aufbau der Hauptuntersuchung wird im folgenden Kapitel „Hauptuntersuchung“ genauer eingegangen. Außerdem konnten durch den umfassenden Einblick, den die Versuchsleiter*innen in den Forschungsprozess gegeben haben sowie dank ihrer Erfahrungswerte, die 14 zu Beginn formulierten Hypothesen, um die drei am Ende des Unterkapitels „Ergänzung: Auffällige Sonderfälle“ hergeleiteten Hypothesen erweitert werden. Welche Daten zur Prüfung dieser Hypothesen benötigt werden, wurde ebenfalls bei der Planung der Hauptuntersuchung berücksichtigt.

Hauptuntersuchung

Die Hauptuntersuchung wurde aufgrund der Ergebnisse der Literaturrecherche und der Voruntersuchung so geplant, dass vier verschiedene Ansätze zur Untersuchung des Ausfalls von Versuchspersonen angewandt werden konnten. (1) Die direkte Erhebung von Veränderungen der Ausfallquote während der Datenerhebung (Erhebung objektiver Daten), (2) die Identifikation von Unterschieden zwischen Personen, die zur Hauptuntersuchung erschienen und jenen, die bei der Hauptuntersuchung ausgefallen sind (Erhebung objektiver Daten), (3) die retrospektive Befragung, ob und warum Personen bereits bei Untersuchungen ausgefallen sind (Erhebung subjektiv retrospektiver Daten) sowie (4) die Erhebung der Einschätzung dazu, unter welchen Bedingungen Versuchspersonen wahrscheinlicher bzw. unwahrscheinlicher zu einer Untersuchung erscheinen (Erhebung subjektiv prospektiver Daten).

Um die Ausfallquote direkt erheben zu können, wurden die Versuchspersonen zu einem vorab vereinbarten Untersuchungstermin an die Universität eingeladen. Dort sollten sie eine Fragebogenbatterie zur Erhebung von Persönlichkeitsmerkmalen, motivationalen Faktoren und Umweltfaktoren ausfüllen. Die Fragebogenbatterie wurde in digitaler Form vorgegeben, damit sie als Online-Fragebogen auch an die ausgefallenen Versuchspersonen gesendet und von diesen über das Internet ausgefüllt werden konnte. Bei der Rekrutierung der Versuchspersonen wurde Acht darauf gegeben, dass die Stichprobe ähnlich zustande kommt und sich dadurch ähnlich zusammensetzt wie bei vergleichbaren Untersuchungen an der Karl-Franzens-Universität Graz. Dazu erfolgte die Rekrutierung der Versuchspersonen in mehreren Phasen, in denen jeweils eine andere Rekrutierungsmaßnahme gesetzt wurde, sodass die Ausfallquote bei der Anwendung der Rekrutierungsmaßnahmen auch einzeln erhoben werden konnte. Jeweils die Hälfte der Versuchspersonen, die mit jeder Maßnahme rekrutiert wurden, erhielt eine Erinnerung am Tag vor der Untersuchung. Zusätzlich wurden am Tag der Testung Aufzeichnungen über Umweltfaktoren gemacht (Uhrzeit, Wetter, Temperatur).

Methoden Hauptuntersuchung

Mit der erstellten Fragebogenbatterie (Screenshots der Fragebogenbatterie befinden sich im Anhang unter „Gestaltung der Hauptuntersuchung“) wurden demografische Daten, Daten zur Erfahrung mit wissenschaftlichen Untersuchungen und Gründe zur Teilnahme an der Untersuchung erhoben sowie ob und warum Personen bei Untersuchungen bereits ausgefallen waren und welche Faktoren die Teilnahme der jeweiligen Person an einer wissenschaftlichen Untersuchung begünstigen oder beeinträchtigen. Außerdem wurden Fragebögen vorgegeben, welche

die Persönlichkeit anhand der sogenannten Big-Five, – fünf Persönlichkeitsdimensionen zur ganzheitlichen Beschreibung der Persönlichkeit – die soziale Verantwortung und die Zuverlässigkeit bei wissenschaftlichen Untersuchungen erheben.

Ein Teil der Fragen für diese Erhebung wurde selbst ausgearbeitet. Zusätzlich wurden der Psychomeda Big-Five-Persönlichkeitstest⁸ (B5T; Satow, 2011; Satow, 2012) und die Skala der sozialen Verantwortung⁹ (SV-Skala; Bierhoff, 2012) zur Erhebung der Persönlichkeitsdimensionen und der sozialen Verantwortung herangezogen. Allen Versuchspersonen wurde die Fragebogenbatterie einzeln vorgegeben. In der Folge werden die verwendeten Instrumente, die Stichprobe und die Operationalisierung der Variablen zur Testung der Hypothesen genauer beschrieben.

Aufbau und Ablauf der Untersuchung

Die Reihenfolge der vorgegebenen Messinstrumente wurde so gewählt, dass die Angabe der demografischen Daten, der Kontrollvariablen und der Persönlichkeitsmerkmale zu Beginn und dadurch ohne Einfluss durch die anderen Messinstrumente geschieht. Im Anschluss daran wurde die retrospektive Erhebung durchgeführt, damit Personen, die bereits bei Untersuchungen ausgefallen waren, sich an diese Situation(en) erinnerten und die Fragen zur Teilnahmewahrscheinlichkeit anhand dieser Erinnerung(en) wahrheitsgetreu ausfüllten. Die Frage zur Zuverlässigkeit bei Untersuchungsterminen wurde zum Abschluss gestellt, damit diese subjektive Selbsteinschätzung die Beurteilung der Teilnahmewahrscheinlichkeiten nicht verzerrte.

Wie in der Einleitung des Kapitels „Hauptuntersuchung“ bereits erwähnt, wurde die Fragebogenbatterie so umgesetzt, dass sie über das Internet abrufbar war und bearbeitet werden konnte, um den bei dieser Untersuchung ausgefallenen Versuchspersonen die Teilnahme mit geringen Hindernissen zu ermöglichen. Aus diesem Grund wurde die Fragebogenbatterie zur Hauptuntersuchung auch möglichst kurz gehalten und konnte in etwa 15 bis 20 Minuten vollständig bearbeitet werden. Zur Anmeldung zur Untersuchung sollte bei der Auswahl eines Untersuchungstermins wahlweise eine Telefonnummer oder eine E-Mailadresse angegeben werden,

⁸ Satow, L. (2011). Psychomeda Big-Five-Persönlichkeitstest (B5T): Skalendokumentation und Normen. Forschungsbericht erhältlich unter www.psychomeda.de (Abgerufen am 19. September 2020) und Satow, L. (2012). Big-Five-Persönlichkeitstest (B5T): Test- und Skalendokumentation. Online im Internet: URL: <http://www.dr.satow.de>. (Abgerufen am 19. September 2020)

⁹ Bierhoff, H.-W. (2012). SV. Skala der sozialen Verantwortung [Verfahrensdokumentation aus PSYNDEX Tests-Nr. 9004061 und Fragebogen]. In Leibniz-Zentrum für Psychologische Information und Dokumentation (ZPID) (Hrsg.), Elektronisches Testarchiv. Trier: ZPID. <https://doi.org/10.23668/psycharchives.342> (Abgerufen am 19. September 2020)

um Erinnerungen an die Versuchspersonen senden und Personen, die nicht zum Termin erschienen waren, kontaktieren zu können.

Die Rekrutierung von Teilnehmer*innen für die Untersuchung erfolgte schrittweise hintereinander und umfasste folgende Maßnahmen: Aushang an der Universität, Teilen der Einladung in sozialen Medien, Aussendung über den universitären E-Mailverteiler „student.umfrage“ und die Rekrutierung von Personen am Campus durch direktes Ansprechen. Dadurch war nachvollziehbar, mit welcher Rekrutierungsmethode jede Versuchsperson – auch die ausgefallenen Personen – rekrutiert wurde. Bei der Rekrutierung wurde angegeben, dass untersucht wird, wie möglichst effiziente und ressourcenschonende Forschung zu gewährleisten ist. So sollte verschleiert werden, dass sich die Untersuchung mit dem Ausfall von Versuchspersonen befasst und damit vorbeugen, dass die Zuverlässigkeit, mit der Personen zu dieser Untersuchung erschienen sind, vorab beeinflusst wurde. Familie, Freunde und Bekannte des Versuchsleiters wurden nicht rekrutiert, da diese Gruppe die Ausfallquote eventuell durch das Zeigen der gewünschten Verhaltensweise, also durch besonders zuverlässiges Verhalten, beeinflussen hätte können (Koch, Peter & Müller, 2019, S. 135).

Um die erhobenen objektiven Daten statistisch analysieren zu können, hätte bei der Untersuchung an der Universität eine repräsentative Anzahl an Versuchspersonen ausfallen müssen. Bei der Hauptuntersuchung hätten mindestens 30 Personen ausfallen müssen, damit entsprechend dem zentralen Grenzwerttheorem sinnvolle statistische Analysen möglich gewesen wären. Bei einer, wie bei der Voruntersuchung ermittelten, durchschnittlichen Ausfallquote von 10,66% müsste eine Stichprobe daher zwischen 172 und 774 Personen enthalten, damit entsprechend der erhobenen Standardabweichung mit 68,27-prozentiger Wahrscheinlichkeit 30 Personen ausfallen.

Da nach der Rekrutierung per Aushang, sozialen Medien und student.umfrage absehbar war, dass diese Stichprobengröße bei Einhaltung der geplanten Vorgehensweise nicht erreicht werden konnte und die Ausfälle in einzelnen Gruppen bis dahin zu klein waren, um statistische Schlüsse zu ziehen, wurde die Untersuchung ab diesem Zeitpunkt als Online-Untersuchung im gesamten deutschsprachigen Raum verbreitet. Durch die Umstellung der Datenerhebung via Online-Untersuchung war es nicht mehr möglich, den Ausfall von Versuchspersonen direkt zu erheben. Das hatte zur Folge, dass Gruppenvergleiche zwischen Personen, die zur Untersuchung an der Universität erschienen und jenen, die ausfielen, nicht möglich waren. Dafür konnte jedoch eine Stichprobengröße erreicht werden, mit der die anderen erhobenen Daten mit größerer Aussagekraft ausgewertet werden konnten.

Die Rekrutierung zur Online-Untersuchung geschah wiederum über das Teilen der Umfrage in sozialen Medien, per Mail und auf dem Online-Umfragen-Panel „SurveyCircle“ (<https://www.surveycircle.com/de/>). Bei SurveyCircle handelt es sich um eine Internetplattform, die Forschende dabei unterstützt, Versuchsteilnehmer*innen für Online-Umfragen im deutschsprachigen Raum zu finden. Dabei können mehr bzw. schneller Versuchsteilnehmer*innen gewonnen werden, wenn Forscher*innen selbst an anderen Untersuchungen auf der Plattform teilnehmen.

Für die Teilnahme an der Untersuchung erhielten die Versuchspersonen die Möglichkeit, an einer Verlosung von drei Amazon-Gutscheinen im Wert von je 20€ teilzunehmen. Bei Bedarf erhielten sie auch einen Untersuchungsschein. Personen, die an der Testung an der Universität teilnahmen und persönlich erschienen, bekamen außerdem Süßigkeiten für die Teilnahme.¹⁰

Stichprobe

Insgesamt füllten 222 Personen die Umfrage ganz oder teilweise aus. Davon nahmen 21 an der Untersuchung vor Ort an der Universität teil, während vier Personen bei dieser Untersuchung ausfielen und den Fragebogen nachträglich ausfüllten. 197 Personen nahmen an der als Online-Umfrage gestalteten Version der Testung teil, wovon 105 Personen über das Online-Umfragen-Panel SurveyCircle und 92 Personen über andere Kanäle rekrutiert wurden. Nach Durchsicht der Daten auf Auffälligkeiten bei den Angaben bzw. Klickmuster, die auf offensichtliche Fehler bei der Bearbeitung hindeuteten, wurden 38 Datensätze ausgeschlossen, um das Datenrauschen zu reduzieren.¹¹ Außerdem wurden alle Datensätze ausgeschlossen, in denen auf der Skala der sozialen Erwünschtheit nur sechs oder weniger Punkte erreicht wurden.¹² Damit verblieben 184 Datensätze in der Gesamtstichprobe, davon 20 aus der Untersuchung an der Universität, drei von den ausgefallenen Versuchspersonen sowie 84 der Datensätze vom Online-Panel SurveyCircle und 77, die über andere Kanäle für die Online-Umfrage rekrutiert werden konnten.

Die endgültige Stichprobe setzte sich aus 117 Frauen (63.59%), 66 Männern (35.87%) und einer Person, die keine Angabe zum Geschlecht machte, zusammen. Die Teilnehmer*innen waren im Durchschnitt 28.32 Jahre alt (Minimum = 18 Jahre, Maximum = 64 Jahre, sd = 8.75,

¹⁰ Anmerkung: Es war ursprünglich geplant, die Aussendungen zur Untersuchung über den Testzeitraum zu verändern, um so beispielsweise die gebotenen Anreize zu variieren, dies war jedoch nicht bei allen Rekrutierungsmaßnahmen gleichermaßen möglich, weshalb gänzlich darauf verzichtet wurde.

¹¹ Im Zuge dessen wurden auch zwei der 14 Personen ausgeschlossen, die angegeben hatten, bei mindestens einer Untersuchung ausgefallen zu sein, weil die Fragen zur Teilnahmewahrscheinlichkeit offensichtlich falsch bearbeitet wurden.

¹² Entsprechend der Testdokumentation des B5T kann so eine Täuschungsabsicht „mit einer Power von 43% und einer Fehlerwahrscheinlichkeit von unter 10% erkannt [werden] (Alpha = 0,057)“ (Satow, 2012).

linkssteil verteilt). 118 der Teilnehmer*innen (64.13%) absolvierten zum Zeitpunkt der Testung ein Studium, 102 Teilnehmer*Innen (55.43%) gaben als höchsten abgeschlossenen Bildungsgrad ein Bachelorstudium oder höheren Abschluss an. Außerdem studierten zum Zeitpunkt der Testung 67 Personen (36.41%) Psychologie oder hatten ein Studium der Psychologie absolviert.

Operationalisierung der Konstrukte

In der Folge wird berichtet, welche Instrumente zur Erhebung der Daten eingesetzt wurden. Alle Daten wurden mit LimeSurvey erhoben und dann zur weiteren Verarbeitung mit SPSS konvertiert. Zur Analyse der erhobenen Daten wurde SPSS 26 (Stand: 22. März 2021) verwendet. Mit SPSS wurden zusätzlich zu deskriptiven Statistiken auch Pearson- und Spearman-Korrelationen, Varianzanalysen, Chi²-Tests, Wilcoxon's Vorzeichen-Rang-Tests, Friedman's Rang-Varianzanalysen sowie zugehörige Post-Hoc-Tests berechnet. Wenn die Unterschiede mit mehreren Post-Hoc-Tests überprüft wurden, wurde aufgrund der Vielzahl an durchgeführten statistischen Tests, die Signifikanz erst nach Durchführung einer entsprechenden Bonferroni-Korrektur berichtet. Dadurch sollte eine Alpha-Fehler-Kumulierung vermieden werden.

Demografische Daten

Erhoben wurden das Alter in Jahren und das Geschlecht (Antwortkategorien: Männlich, Weiblich oder keine Antwort) der jeweiligen Versuchsperson. Zusätzlich sollte sie angeben, wie lange sie in Minuten von ihrem Wohnort aus für die Anreise zur Universität benötigt, welche Gründe für die Teilnahme entscheidend waren (vorgegebene Antwortkategorien: Interesse, Versuchsschein benötigt bzw. Punkte für SurveyCircle, Terminfindung einfach bzw. Onlineuntersuchung, Langeweile, Versprochene Belohnung [Süßigkeiten], Untersuchungsdauer passt, Gewinnspielteilnahme, Kenne den Versuchsleiter, Sonstiges [mit Textfeld]) und wie sie auf die Untersuchung aufmerksam wurde (vorgegebene Antwortkategorien: Mail von student.umfrage, Aushang an der Universität, über SurveyCircle, Flyer bekommen, Link zugesendet bekommen, In sozialen Medien gesehen [Facebook, Instagram,...], In einer Vorlesung davon gehört, Von FreundInnen/StudienkollegInnen davon gehört, Kenne den Versuchsleiter, Sonstiges [mit Textfeld]). Sowohl bei den Teilnahmegründen, als auch bei der Frage danach, wie sie auf die Untersuchung aufmerksam wurden, waren die am häufigsten zu erwartenden Antwortmöglichkeiten vorgegeben, konnten aber in einem zusätzlichen Textfeld frei ergänzt werden. Mehrfachantworten waren möglich.

Kontrollvariablen

Zur Kontrolle verschiedener Störvariablen wurde erhoben, ob die jeweilige Versuchsperson Psychologie studiert oder studiert hatte, an wie vielen wissenschaftlichen Untersuchungen sie bereits teilgenommen hatte, ob und wie viele Untersuchungen sie schon als Versuchsleiter*in begleitet hatte sowie der höchste abgeschlossene Bildungsgrad (Antwortkategorien: Pflichtschulabschluss, Lehre, Matura, Bachelor, Master, Magister/Magistra, Doktorat, Sonstiges [mit Textfeld]) und ihre aktuelle Tätigkeit (Antwortkategorien: Schule, Studium, Geringfügig angestellt, Teilzeit angestellt, Vollzeit angestellt, Selbstständig, Sonstiges [mit Textfeld]). Diese Variablen sollten es ermöglichen, Einflüsse durch persönliche Erfahrungen mit der Arbeit im wissenschaftlichen Kontext kontrollieren zu können. Außerdem können Personen, die noch nie an einer Untersuchung teilgenommen haben, auch nicht retrospektiv von einem Ausfall berichten. Je nach Analyseverfahren hätte das eine Verzerrung der Daten zur Folge haben können.

B5T

Der Psychomeda Big-Five-Persönlichkeitstest¹³ wurde von Dr. Lars Satow entwickelt und erhebt die Persönlichkeit von Personen auf den fünf Dimensionen Neurotizismus, Extraversion, Verträglichkeit, Gewissenhaftigkeit und Offenheit (Satow, 2011). Er wurde so konstruiert, dass er bei guter Validität und hoher Reliabilität kürzer als vorhergehende Tests zu den Big-Five ist, wodurch er sich gut für den Einsatz als Online-Test eignet (Satow, 2011). Der Haupttest besteht aus insgesamt 50 Items (Beispielitem: „Ich achte sehr darauf, dass Regeln eingehalten werden.“), je zehn davon wurden jeder Persönlichkeitsdimension zugeordnet. Dabei dient eine vierstufige Likert-Skala („trifft gar nicht zu“, „trifft eher nicht zu“, „trifft eher zu“, „trifft genau zu“) der Beantwortung der Items. Außerdem kann der Test um einige ergänzende Skalen erweitert werden. Für diese Arbeit wurden vier weitere Items herangezogen, um die Ehrlichkeit von Versuchspersonen bei der Beantwortung der Fragen zu überprüfen (Beispielitem: „Ich habe schon mal über andere gelästert oder schlecht über sie gedacht.“; Satow, 2012).

¹³ Satow, L. (2011). Psychomeda Big-Five-Persönlichkeitstest (B5T): Skalendokumentation und Normen. Forschungsbericht erhältlich unter www.psychomeda.de (Abgerufen am 19. September 2020) und Satow, L. (2012). Big-Five-Persönlichkeitstest (B5T): Test- und Skalendokumentation. Online im Internet: URL: <http://www.dr.satow.de>. (Abgerufen am 19. September 2020)

SV-Skala und Zuverlässigkeit

Die Skala der sozialen Verantwortung¹⁴ von Prof. Dr. Hans-Werner Bierhoff besteht aus 22 Items (Beispielitem: „Wenn ein Fremder auf meine Unterstützung angewiesen ist, versuche ich, sie ihm zu geben.“). Jeweils acht davon können herangezogen werden, um die Subskalen „Erfüllung von berechtigten Erwartungen anderer“ und „Befolgung von sozialen Spielregeln“ zu berechnen (Bierhoff, 2012). Die persönliche Zustimmung zu den Items ist auf einer sechsstufigen Likert-Skala zu bewerten, die von „sehr falsch“ bis zu „sehr richtig“ reicht. Die Skala basiert auf der englischen Skala zur Messung der sozialen Verantwortung (Berkowitz & Daniels 1964) und weist in der deutschen Fassung eine zufriedenstellende Reliabilität sowie Konstrukt- und Kriteriumsvalidität auf (Bierhoff, 2012).

Retrospektive Daten

Die Versuchspersonen wurden jeweils dazu befragt, ob sie bereits einmal trotz Anmeldung nicht zu einer wissenschaftlichen psychologischen Untersuchung erschienen waren („Sind Sie schon einmal trotz Anmeldung nicht zu einer psychologischen Untersuchung erschienen?“). Wenn diese Frage mit „ja“ beantwortet wurde, bekam die Versuchsperson zusätzlich Fragen dazu, wie oft sie bei Untersuchungen bereits ausgefallen ist („Wie viele Male sind Sie bereits zu psychologischen Untersuchungen nach Anmeldung nicht erschienen?“) und warum („Welchen Grund / welche Gründe gab es, dass Sie nicht zur Untersuchung / zu den Untersuchungen erschienen sind?“; offenes Antwortformat). Außerdem wurde auf dieselbe Weise erhoben, ob, wie oft und warum Personen sich von wissenschaftlichen psychologischen Untersuchungen wieder abgemeldet hatten („Haben Sie sich schon einmal von einer psychologischen Untersuchung wieder abgemeldet?“, „Wie viele Male haben Sie sich bereits von Untersuchungen nach Anmeldung wieder abgemeldet?“, „Welchen Grund / welche Gründe gab es, dass Sie sich von der Untersuchung / den Untersuchungen wieder abgemeldet haben?“; offenes Antwortformat).

Teilnahmewahrscheinlichkeit

Eine niedrige Ausfallquote wird durch eine durchschnittlich hohe Teilnahmewahrscheinlichkeit der Versuchspersonen bedingt. Um zu erheben, welche Einflussfaktoren einen potentiellen Ausfall aus Sicht der Versuchsteilnehmer*innen begünstigen oder vorbeugen, wurde deshalb eine Skala zur Beurteilung der Teilnahmewahrscheinlichkeit vorgegeben. Anhand der 17 aufgestellten Hypothesen und fünf ergänzender Fragestellungen (siehe folgendes Unterkapitel

¹⁴ Bierhoff, H.-W. (2012). SV. Skala der sozialen Verantwortung [Verfahrensdokumentation aus PSYNDEX Tests-Nr. 9004061 und Fragebogen]. In Leibniz-Zentrum für Psychologische Information und Dokumentation (ZPID) (Hrsg.), Elektronisches Testarchiv. Trier: ZPID. <https://doi.org/10.23668/psycharchives.342> (Abgerufen am 19. September 2020)

„Hypothesen der Hauptuntersuchung“), wurden 38 Items erarbeitet, welche die Merkmale von Untersuchungen, die Rekrutierung, motivationale Faktoren und Umweltfaktoren umfassten (siehe „Fragebogen zu den Teilnahmewahrscheinlichkeiten“ im Anhang). Dabei sollte die jeweilige Versuchsperson angeben, wie die vorgegebenen Bedingungen die Wahrscheinlichkeit beeinflussen, dass sie zu einem Versuchstermin, zu dem sie sich angemeldet hat, auch tatsächlich erscheint. Zu jeder Bedingung war die Änderung der Wahrscheinlichkeit auf einer fünfstufigen Likert-Skala anzugeben (Beispielitem: „Wenn draußen besonders schlechtes Wetter ist...“ „sinkt die Wahrscheinlichkeit stark“, „sinkt die Wahrscheinlichkeit ein wenig“, „bleibt die Wahrscheinlichkeit gleich“, „erhöht sich die Wahrscheinlichkeit ein wenig“, „erhöht sich die Wahrscheinlichkeit stark“). Zusätzlich gab es die Möglichkeit „würde mich von vornherein nicht anmelden“ auszuwählen, um der Verfälschung der angegebenen Wahrscheinlichkeiten durch Personen vorzubeugen, für die einzelne Bedingungen einen Ausschlussgrund für eine Untersuchungsteilnahme darstellen. Die Vorgabe einer Mittelkategorie wurde hier bewusst gewählt, da es nicht auszuschließen ist, dass es Faktoren gibt, die keinen Einfluss auf die Teilnahmewahrscheinlichkeit haben.

Zuverlässigkeit bei Untersuchungsteilnahmen

Zum Abschluss wurde jede/r Teilnehmer*in zusätzlich nach seiner/ihrer Zuverlässigkeit bei Untersuchungsteilnahmen befragt („Für wie wahrscheinlich halten Sie es, dass Sie einen Versuchstermin verschlafen, vergessen oder aus anderen Gründen übersehen?“; 1 = nicht wahrscheinlich, 5 = sehr wahrscheinlich).

Hypothesen der Hauptuntersuchung

In der Folge werden überblicksartig alle aufgestellten Hypothesen aufgelistet und die zur Überprüfung relevanten Variablen angeführt. Die Herleitung der Hypothesen erfolgte bereits im Unterkapitel „Mögliche Einflussfaktoren auf die Ausfallquote“. Ein Verweis auf die zugehörigen Textstellen findet sich bei der jeweiligen Hypothese.

Einflussfaktoren bei der Planung, Gestaltung und Durchführung der Untersuchung

Hypothese 1.1.: Erinnerungen an den Untersuchungstermin reduzieren den Ausfall von Versuchspersonen.

Die theoriebasierte Herleitung der Hypothese erfolgt auf Seite 8 ff. Die Hälfte der Versuchspersonen, die an der Untersuchung an der Universität teilnahmen, erhielten eine Erinnerung an

den Termin. Es wurde mit einem Chi²-Test überprüft, ob bei der Untersuchung an der Universität mehr Personen ausfielen, die keine Erinnerung an den Termin erhielten, als Personen die eine Erinnerung erhielten. Zur Überprüfung der Hypothese wurden auch die Teilnahmewahrscheinlichkeiten unter den Bedingungen erhoben, dass die Versuchspersonen keine Erinnerung („Wenn ich von dem/der Versuchsleiter/in keine Erinnerung an den Termin erhalte...“) bzw. am Vortag der Untersuchung eine Erinnerung an den Untersuchungstermin erhalten („Wenn ich am Vortag eine Erinnerung an den Termin erhalte...“). Mit einem Wilcoxon’s Vorzeichen-Rang-Test wurde überprüft, ob es einen Unterschied in der zentralen Tendenz der beiden Bedingungen gibt.

*Hypothese 1.2.: Die freie Auswahl eines Termins durch die Versuchsteilnehmer*innen reduziert die Ausfallquote.*

Die theoriebasierte Herleitung der Hypothese erfolgt auf Seite 9. Zur Überprüfung der Hypothese wurde erhoben, welche Teilnahmewahrscheinlichkeit die Versuchspersonen unter der Bedingung angeben, dass sie den Untersuchungstermin frei wählen dürfen („Wenn ich den Untersuchungstermin frei wählen kann...“).

Hypothese 1.3.: Die verwendeten Untersuchungsmethoden haben keinen Einfluss auf den Ausfall von Versuchspersonen.

Die theoriebasierte Herleitung der Hypothese erfolgt auf Seite 9 ff. Es wurde zur Prüfung dieser Hypothese die Teilnahmewahrscheinlichkeit erhoben, wenn die Versuchspersonen einen Leistungstest („Wenn ich bei der Untersuchung einen Leistungstest bearbeiten soll...“) oder Fragebögen („Wenn ich bei der Untersuchung Fragebögen bearbeiten soll...“) bearbeiten sollen, wenn aufwendige technische Hilfsmittel zum Einsatz kommen („Wenn bei der Untersuchung aufwendige technische Hilfsmittel zum Einsatz kommen [EEG, fMRT, etc.]...“) sowie, wenn ihre aktive Teilnahme besonders gefordert wird („Wenn bei der Untersuchung meine aktive Teilnahme besonders gefordert ist...“). Es wurde dann mit einer Friedman’s Rang-Varianzanalyse überprüft, ob ein Unterschied zwischen den zentralen Tendenzen dieser Bedingungen besteht.

Hypothese 1.4.: Die Ausfallquote ist bei kurzen Untersuchungen geringer als bei langen Untersuchungen.

Die theoriebasierte Herleitung der Hypothese erfolgt auf Seite 10 ff. Zur Prüfung der Hypothese wurde erhoben, wie sich die Teilnahmewahrscheinlichkeit verändert, wenn die Untersuchung 30 Minuten („Wenn die Untersuchung weniger als 30 Minuten dauert...“), wenn sie zwischen

30 und 90 Minuten („Wenn die Untersuchung zwischen 30 und 90 Minuten dauert...“) und wenn sie mehr als 90 Minuten dauert („Wenn die Untersuchung mehr als 1 1/2 Stunden dauert...“). Mit einer Friedman’s Rang-Varianzanalyse wurde überprüft, ob ein Unterschied zwischen den zentralen Tendenzen dieser Bedingungen besteht.

Gestaltung des Rekrutierungsprozesses

Hypothese 2.1.: Persönlicher Kontakt bei der Rekrutierung senkt die Ausfallquote.

Die theoriebasierte Herleitung der Hypothese erfolgt auf Seite 44. Zur Überprüfung der Hypothese wurde die Teilnahmewahrscheinlichkeit erhoben, wenn der/die Versuchsleiter*in vorab persönlich getroffen wurde („Wenn ich den/die Versuchsleiter/in vor der Untersuchung bereits persönlich getroffen und mit ihr/ihm gesprochen habe...“), wenn vor der Untersuchung Kontakt per Mail oder SMS stattgefunden hat („Wenn ich vor der Untersuchung bereits per Mail oder SMS mit dem/r Versuchsleiter/in kommuniziert habe...“) und wenn es keinen persönlichen Kontakt mit dem/der Versuchsleiter*in gegeben hat („Wenn ich den Untersuchungstermin ohne persönlichen Kontakt mit dem/der Versuchsleiter/in vereinbart haben... [Z.B. durch einen Eintrag in eine Terminliste]“). Mit einer Friedman’s Rang-Varianzanalyse wurde überprüft, ob ein Unterschied zwischen den zentralen Tendenzen dieser Bedingungen besteht.

Personenspezifische Merkmale der Versuchsteilnehmer*innen

Hypothese 3.1.: Zuverlässigere Personen fallen als Versuchspersonen bei wissenschaftlichen Untersuchungen seltener aus.

Die theoriebasierte Herleitung der Hypothese erfolgt auf Seite 14. Zur Überprüfung dieser Hypothese wurde abgefragt, wie zuverlässig Versuchspersonen sich selbst bei der Teilnahme an wissenschaftlichen Untersuchungen einschätzen („Für wie wahrscheinlich halten Sie es, dass Sie einen Versuchstermin verschlafen, vergessen oder aus anderen Gründen übersehen?“). Außerdem wurde ein Item der SV-Skala, das erhebt, wie sehr sich andere Personen auf die oder den Befragte*n verlassen können („Ich gehöre zu der Art von Menschen, auf die andere sich verlassen können.“), herangezogen. Es wurde geprüft, ob ein Zusammenhang der beiden oben erwähnten Items damit besteht, dass die Versuchspersonen bei der aktuellen Untersuchung ausfielen – Versuchspersonen, die ausfielen, bekamen den Fragebogen in digitaler Form zugesendet – bzw. damit, dass sie in der Vergangenheit bereits mindestens einmal trotz Anmeldung nicht zu einem vereinbarten Untersuchungstermin erschienen sind.

Hypothese 3.2.: Gewissenhaftere Personen fallen als Versuchspersonen bei wissenschaftlichen Untersuchungen seltener aus.

Die theoriebasierte Herleitung der Hypothese erfolgt auf Seite 14. Es wurde geprüft, ob ein Zusammenhang der Skala Gewissenhaftigkeit des B5T (Satow, 2011) damit besteht, ob die Versuchspersonen bei der aktuellen Untersuchung ausgefallen sind bzw. damit, ob sie in der Vergangenheit bereits mindestens einmal trotz Anmeldung nicht zu einem vereinbarten Untersuchungstermin erschienen sind.

Hypothese 3.3.: Versuchspersonen, die eine höhere soziale Verantwortung aufweisen, fallen bei wissenschaftlichen Untersuchungen seltener aus.

Die theoriebasierte Herleitung der Hypothese erfolgt auf Seite 14 ff. Es wurde geprüft, ob ein Zusammenhang der sozialen Verantwortung, die mit der SV-Skala (Bierhoff, 2012) erhoben wurde, damit besteht, ob die Versuchspersonen bei der aktuellen Untersuchung ausgefallen sind bzw. damit, ob sie in der Vergangenheit bereits mindestens einmal trotz Anmeldung nicht zu einem vereinbarten Untersuchungstermin erschienen sind.

Motivationale Faktoren

Hypothese 4.1.: Aufwendige Untersuchungen weisen eine höhere Ausfallquote auf.

Die theoriebasierte Herleitung der Hypothese erfolgt auf Seite 16 ff. Zur Überprüfung der Hypothese wurde die Teilnahmewahrscheinlichkeit erhoben, wenn bei der Untersuchung aufwendige technische Hilfsmittel zum Einsatz kommen („Wenn bei der Untersuchung aufwendige technische Hilfsmittel zum Einsatz kommen [EEG, fMRT, etc.]...“), wenn die Anwesenheit der Versuchsperson an mehreren Terminen notwendig ist („Wenn die Untersuchung aus mehreren Terminen besteht...“), wenn es eine Vor- und eine Hauptuntersuchung gibt („Wenn ich nach einer Voruntersuchung zur Hauptuntersuchung eingeladen werde...“) und wenn die aktive Teilnahme der Versuchsperson besonders gefordert ist („Wenn bei der Untersuchung meine aktive Teilnahme besonders gefordert ist...“).

Hypothese 4.2.: Monetäre Anreize für die Teilnahme reduzieren die Ausfallquote.

Die theoriebasierte Herleitung der Hypothese erfolgt auf Seite 17 ff. Damit die Hypothese überprüft werden konnte, wurde die Teilnahmewahrscheinlichkeit erhoben, wenn die Versuchspersonen fünf Euro („Wenn ich 5€ für die Teilnahme erhalte...“), zehn Euro pro Stunde („Wenn ich 10€/Stunde für die Teilnahme erhalte...“) bzw. 30 Euro („Wenn ich 30€ für die Teilnahme an der Untersuchung erhalte...“) für die Teilnahme an der Untersuchung erhalten.

Hypothese 4.3.: Die Möglichkeit der Teilnahme an einer Preisverlosung reduziert die Ausfallquote.

Die theoriebasierte Herleitung der Hypothese erfolgt auf Seite 18. Zur Überprüfung der Hypothese wurde erhoben, welche Teilnahmewahrscheinlichkeit die Versuchspersonen unter der Bedingung angeben, dass sie nach der Untersuchungsteilnahme an einem Gewinnspiel teilnehmen dürfen („Wenn ich nach der Untersuchungsteilnahme bei einem Gewinnspiel mitmachen darf...“).

Hypothese 4.4.: Personen, die dazu verpflichtet sind, sich als Versuchsperson zur Verfügung zu stellen, fallen seltener aus.

Die theoriebasierte Herleitung der Hypothese erfolgt auf Seite 18 ff. Es wurde mit einem Chi²-Test überprüft, ob bei der Untersuchung an der Universität mehr Personen ausgefallen sind, die keinen Versuchsschein benötigten, als Personen die einen Versuchsschein für die Untersuchungsteilnahme erhielten. Personen die bei der ursprünglichen Untersuchung ausgefallen waren, hatten die Möglichkeit, sich nach der nachträglichen Teilnahme an der Untersuchung über das Internet einen Versuchsschein ausstellen zu lassen.

Hypothese 4.5.: Versuchspersonen mit Interesse am Thema der Untersuchung fallen seltener aus als Versuchspersonen ohne Interesse am Thema.

Die theoriebasierte Herleitung der Hypothese erfolgt auf Seite 19. Zur Prüfung der Hypothese wurden die Teilnahmegründe an der Untersuchung erhoben. Es wurde mit einem Chi²-Test überprüft, ob es einen Unterschied bei der Anzahl der ausgefallenen Personen und der Anzahl jener Personen, die tatsächlich an der Untersuchung an der Universität teilgenommen haben, bezüglich dessen gibt, dass sie sich aus Interesse an der Untersuchung angemeldet haben. Außerdem wurde die Teilnahmewahrscheinlichkeit unter den Bedingungen erhoben, dass Versuchspersonen das Thema der Untersuchung nicht interessant („Wenn ich das Thema der Untersuchung nicht interessant finde...“) bzw. besonders interessant („Wenn ich das Thema der Untersuchung besonders interessant finde...“) finden. Mit einem Wilcoxon's Vorzeichen-Rang-Test wurde überprüft, ob es einen Unterschied in der zentralen Tendenz der beiden Bedingungen gibt.

Hypothese 4.6.: Die Ausfallquote ist bei Untersuchungen kleiner, wenn die Anreize einen kleineren Wert haben, als wenn wertvollere Anreize geboten werden.

Die theoriebasierte Herleitung der Hypothese erfolgt beginnend auf Seite 42 ff. Zur Überprüfung der Hypothese wurde die Teilnahmewahrscheinlichkeit erhoben, wenn die Versuchspersonen, nach der Teilnahme an der Untersuchung an einem Gewinnspiel teilnehmen dürfen („Wenn ich nach der Untersuchungsteilnahme bei einem Gewinnspiel mitmachen darf...“), bzw. wenn sie Süßigkeiten („Wenn ich Süßigkeiten für die Teilnahme bekomme...“), fünf Euro („Wenn ich 5€ für die Teilnahme erhalte...“), zehn Euro pro Stunde („Wenn ich 10€/Stunde für die Teilnahme erhalte...“) bzw. 30 Euro („Wenn ich 30€ für die Teilnahme an der Untersuchung erhalte...“) für die Teilnahme an der Untersuchung erhalten. Mit einer Friedman’s Rang-Varianzanalyse wurde überprüft, ob und wie sich die zentralen Tendenzen dieser Bedingungen unterscheiden.

Umweltfaktoren

Hypothese 5.1.: Personen, die eine aufwendigere Anreise haben, fallen bei wissenschaftlichen Untersuchungen eher aus.

Die theoriebasierte Herleitung der Hypothese erfolgt auf Seite 21. Zur Überprüfung der Hypothese wurde unter den Personen, die sich zur Untersuchung an der Universität angemeldet haben, erhoben, wie lange sie für die Anreise zur Universität Graz in Minuten benötigen („Wie lange benötigen Sie für gewöhnlich von Ihrem Wohnort aus bis zur Karl-Franzens-Universität Graz [in Minuten]:“)

Hypothese 5.2.: Versuchspersonen, die eine hohe Geschäftigkeit empfinden, fallen bei wissenschaftlichen Untersuchungen eher aus.

Die theoriebasierte Herleitung der Hypothese erfolgt auf Seite 21. Zur Prüfung der Hypothese wurde die Teilnahmewahrscheinlichkeit unter verschiedenen Bedingungen erhoben, die eine erhöhte Geschäftigkeit bedingen. Die Teilnahmewahrscheinlichkeit wurde erhoben unter der Bedingung, dass (kurz) vor dem Untersuchungstermin der/die Chef*in der Versuchsperson sie bittet bei der Arbeit einzuspringen („Wenn mein/e Chef/in mich [kurz] vor dem Untersuchungstermin bittet bei der Arbeit einzuspringen...“), sie von einem Familienmitglied um Hilfe gebeten wird („Wenn ein Familienmitglied mich vor dem Untersuchungstermin um Hilfe bittet...“), bzw. ein/e Freund*in sie spontan um Hilfe bittet („Wenn ein/e Freund/in mich vor dem Untersuchungstermin spontan um Hilfe bittet...“).

Hypothese 5.3.: Bei schlechtem Wetter fallen Versuchspersonen eher aus.

Die theoriebasierte Herleitung der Hypothese erfolgt auf Seite 43. Zur Überprüfung der Hypothese wurden Aufzeichnungen über Temperatur in Grad Celsius und des Wetters (Aufzeichnungen mit den Kategorien: sonnig, teilweise sonnig, bewölkt und regnerisch) zu den vereinbarten Untersuchungsterminen geführt sowie die Teilnahmewahrscheinlichkeit unter den Bedingungen erhoben, dass draußen besonders schönes („Wenn draußen besonders schönes Wetter ist...“), leicht bewölkt („Wenn draußen leicht bewölkt Wetter ist...“) bzw. schlechtes Wetter („Wenn draußen besonders schlechtes Wetter ist...“) ist. Mit einer Friedman’s Rang-Varianzanalyse wurde dann überprüft, ob ein Unterschied zwischen den zentralen Tendenzen der so erhobenen Teilnahmewahrscheinlichkeiten besteht.

Ergänzende Fragestellungen

Folgende Fragestellungen wurden ergänzend explorativ untersucht. Sie umfassen Einflussfaktoren die möglicherweise einen Einfluss auf die Ausfallquote haben, aber zu denen es nicht genug Evidenz gibt um eine gerichtete Hypothese zu formulieren. Die Relevanz der jeweiligen Themen wurde teils im Unterkapitel „Mögliche Einflussfaktoren auf die Ausfallquote“ und teils im Kapitel „Diskussion der Voruntersuchung“ besprochen. Ein Verweis auf die zugehörigen Textstellen findet sich bei der jeweiligen Hypothese.

Ergänzende Fragestellung 6.1.: Haben die verwendeten Rekrutierungsmaßnahmen Einfluss auf die Ausfallquote?

Inwiefern die verwendeten Rekrutierungsmaßnahmen für den Ausfall von Versuchspersonen relevant sind, wird auf Seite 11 ff ausgeführt. Um zu überprüfen ob, die verwendeten Rekrutierungsmaßnahmen einen Einfluss auf die Ausfallquote haben, wurde erhoben, wie viele Personen durch welche Maßnahme rekrutiert wurden (Aushang an der Universität, Teilen der Einladung in sozialen Medien und die Rekrutierung von Personen am Campus durch direktes Ansprechen) und ob sich die Ausfallquoten dieser Gruppen unterscheiden.

Ergänzende Fragestellung 6.2.: Hat die Verwendung der Höflichkeitsform Einfluss auf die Ausfallquote?

Inwiefern dieses Thema für den Ausfall von Versuchspersonen relevant ist, wird auf Seite 12 ausgeführt. Zur Überprüfung, ob die Verwendung der Höflichkeitsform bei der Rekrutierung einen Einfluss auf die Ausfallquote hat, wurde die Teilnahmewahrscheinlichkeit erhoben wenn Versuchspersonen in der Ausschreibung zur Untersuchung mit „Du“ („Wenn ich in der Ausschreibung zur Untersuchung mit „Du“ angesprochen werde...“) bzw. mit „Sie“ („Wenn ich in

der Ausschreibung zur Untersuchung mit „Sie“ angesprochen wurde...“) angesprochen werden. Mit einem Wilcoxon's Vorzeichen-Rang-Test wurde überprüft, ob es einen Unterschied in der zentralen Tendenz dieser beiden Bedingungen gibt.

*Ergänzende Fragestellung 6.3.: Hat ein Näheverhältnis zum/zur Versuchsleiter*in Einfluss auf die Ausfallquote?*

Inwiefern das Näheverhältnis zum/zur Versuchsleiter*in für den Ausfall von Versuchspersonen relevant ist, wird auf Seite 39 ausgeführt. Der Einfluss des Näheverhältnisses wurde durch Erhebung der Teilnahmewahrscheinlichkeit unter den Bedingungen, dass die Untersuchung von einem/einer Freund*in („Wenn die Untersuchung von einem/r Freund/in geleitet wird...“) bzw. einer/einem Fremden („Wenn die Untersuchung von einem/r Fremden geleitet wird...“) geleitet wird. Mit einem Wilcoxon's Vorzeichen-Rang-Test wurde überprüft, ob es einen Unterschied in der zentralen Tendenz dieser beiden Bedingungen gibt.

Ergänzende Fragestellung 6.4.: Hat das Geschlecht der Versuchsperson Einfluss auf die Ausfallquote?

Inwiefern dieses Thema für den Ausfall von Versuchspersonen relevant ist, wird auf Seite 15 ausgeführt. Es wurde überprüft, ob ein Zusammenhang zwischen dem angegebenen Geschlecht (männlich, weiblich, keine Angabe) damit besteht, ob die Versuchspersonen bei der aktuellen Untersuchung ausgefallen sind bzw. damit, ob sie in der Vergangenheit bereits mindestens einmal trotz Anmeldung nicht zu einem vereinbarten Untersuchungstermin erschienen sind.

Ergänzende Fragestellung 6.5.: Hat die Uhrzeit des Untersuchungstermins Einfluss auf die Ausfallquote?

Inwiefern dieses Thema für den Ausfall von Versuchspersonen relevant ist, wird auf Seite 20 ausgeführt. Der Einfluss der Uhrzeit des Untersuchungstermins auf die Ausfallquote wurde überprüft, indem festgehalten wurde, zu welcher Tageszeit sich Versuchspersonen zur Untersuchung anmeldeten und ob es Unterschiede bei der Ausfallquote bezüglich der Tageszeit gibt. Außerdem wurde die Teilnahmewahrscheinlichkeit dafür erhoben, dass die Untersuchung vor 8:30 Uhr („Wenn die Untersuchung vor 8:30 Uhr startet...“), zwischen 9:00 und 17:00 Uhr („Wenn die Untersuchung zwischen 9:00 Uhr und 17:00 Uhr startet...“) oder nach 18:30 („Wenn die Untersuchung nach 18:30 Uhr startet...“) beginnt. Mit einer Friedman's Rang-Varianzanalyse wurde überprüft, ob ein Unterschied zwischen den zentralen Tendenzen dieser Bedingungen besteht.

Kontrollitems

Zur Kontrolle, ob die Versuchspersonen die Untersuchung ehrlich und richtig bearbeiten, wurden die soziale Erwünschtheit mit dem B5T (Satow, 2012) sowie ein Item erhoben, mit dem geprüft werden sollte, ob die Items zur Teilnahmewahrscheinlichkeit aufmerksam gelesen wurden („Wenn ich mich nicht an den Termin erinnere...“). Außerdem wurde erhoben ob die Versuchspersonen Psychologie studierten oder studiert hatten. So sollte der Einfluss kontrolliert werden können, den ein Psychologiestudium möglicherweise auf die Zuverlässigkeit von Versuchspersonen bei Untersuchungsteilnahmen hat, da davon ausgegangen wurde, dass ein Teil der Teilnehmer*innen aus dieser Subpopulation rekrutiert werden wird.

Ergebnisse Hauptuntersuchung

Die statistische Auswertung erfolgte entsprechend den im Kapitel „Stand der Forschung“ ausgearbeiteten Themenblöcken: Einflussfaktoren bei der Planung, Gestaltung und Durchführung der Untersuchung, Gestaltung des Rekrutierungsprozesses, personenspezifische Merkmale der Versuchsteilnehmer*innen, motivationale Faktoren und Umweltfaktoren. Die nach der Voruntersuchung zusätzlich aufgestellten Hypothesen wurden ebenfalls diesen Themenblöcken zugeordnet. Die ergänzenden Fragestellungen wurden anschließend separat geprüft.

Bei der Abgabe wurde dieser Masterarbeit ein USB-Stick mit den erhobenen Daten beigelegt. Auf Anfrage werden die Daten dieser Untersuchung von Em.Univ.-Prof. Dr.rer.nat. Dietrich Albert oder vom Autor dieser Arbeit gerne zur Verfügung gestellt.

Überprüfung der Homogenität der Stichprobe

Die drei unterschiedlichen Stichproben (Untersuchung an der Universität, das Online-Panel SurveyCircle und weitere Online-Umfrage – in der Folge bezieht sich „Online-Umfrage“ auf jenen Teil der online rekrutierten Personen, die nicht über SurveyCircle rekrutiert wurden) wurden anhand der demografischen Daten auf Unterschiede geprüft. Dabei wurden Gruppenunterschiede bezüglich des Alters ($F_{2,1} = 4.27, p < .05$) sowie der Verteilung der teilnehmenden Männer und Frauen ($\chi^2_{2, n=183} = 7.78, p < .05$), der Zahl der Studierenden ($\chi^2_{2, n=184} = 26.71, p < .001$), der Anzahl der Personen die zum Zeitpunkt der Untersuchung Psychologie studierten bzw. studiert hatten ($\chi^2_{2, n=184} = 30.30, p < .001$), der Verteilung des höchsten abgeschlossenen Bildungsgrades ($\chi^2_{12, n=177} = 36.75, p < .001$), der Zahl der Vollzeitbeschäftigten ($\chi^2_{2, n=184} = 22.71, p < .001$) und der Anzahl der Selbstständigen ($\chi^2_{2, n=184} = 7.22, p < .05$) festgestellt. Aufgrund der festgestellten demografischen Unterschiede wurden die Daten sowohl für die Gesamtstichprobe als

auch für die einzelnen Stichproben gruppenweise ausgewertet. Wenn ein statistischer Wert in einer der Gruppen von dem entsprechenden Wert in der Gesamtstichprobe abwich, wird darauf explizit hingewiesen und dieser zusätzlich berichtet. Eine Ausnahme bilden alle Werte bei denen eine der Gruppen dadurch eine Größe von weniger als zehn Personen erreicht hätte. Diese Daten wurden aufgrund mangelnder Repräsentativität nicht berichtet.

Ausfall von Versuchspersonen

Zur Untersuchung an der Universität meldeten sich 27 Personen an. 21 Personen erschienen zum Termin, fünf Personen fielen aus und eine Person meldete sich am Tag der Untersuchung wegen Kopfschwehs ab. Somit ergibt sich, entsprechend der bei der Voruntersuchung verwendeten Formel, eine Ausfallquote von 18.52% (5 von 27).

In der Gesamtstichprobe berichteten zwölf von 184 Personen (6.52%), dass sie bereits mindestens einmal bei einer wissenschaftlichen psychologischen Untersuchung ausgefallen sind (Maximum = 6 Ausfälle; $\bar{x} = 0.14$, $sd = 0.64$, linkssteil verteilt), während 26 von 184 Personen (14.13%) sich schon mindestens einmal von einer Untersuchung wieder abgemeldet hatten (Maximum = 5 Abmeldungen; $\bar{x} = 0.26$, $sd = 0.74$, linkssteil verteilt). Fünf Personen davon hatten sich mindestens einmal von einer Untersuchung abgemeldet und waren mindestens einmal ausgefallen. Im Durchschnitt hatten die Versuchspersonen vor dieser Untersuchung an 9.58 Untersuchungen an Universitäten teilgenommen (Maximum = 100 Versuchsteilnahmen, $sd = 14.24$, Median = 3.5, linkssteil verteilt), wobei 60 von 184 Personen zuvor noch nie an solch einer Untersuchung teilgenommen hatten. Alle Personen, die angegeben hatten, bei mindestens einer Untersuchung ausgefallen zu sein, hatten zuvor an mindestens vier psychologischen Untersuchungen an Universitäten teilgenommen.

Auswertung der Teilnahmewahrscheinlichkeiten

Die Antworten zu den Items der Teilnahmewahrscheinlichkeiten sind ordinal skaliert, weshalb sie anhand von Häufigkeiten analysiert wurden. Die vollständige Häufigkeitstabelle befindet sich im Anhang (siehe Tabelle zu ergänzenden Daten 1 unter „Häufigkeitstabelle zur Teilnahmewahrscheinlichkeit“ im Anhang). Den Antwortkategorien wurden folgende Zahlen zur Angabe von Median und Modus zugewiesen: Die Wahrscheinlichkeit... „...sinkt stark“ = 1, „...sinkt ein wenig“ = 2, „...bleibt gleich“ = 3, „...steigt ein wenig“ = 4, „...steigt stark“ = 5.

Testung der Hypothesen

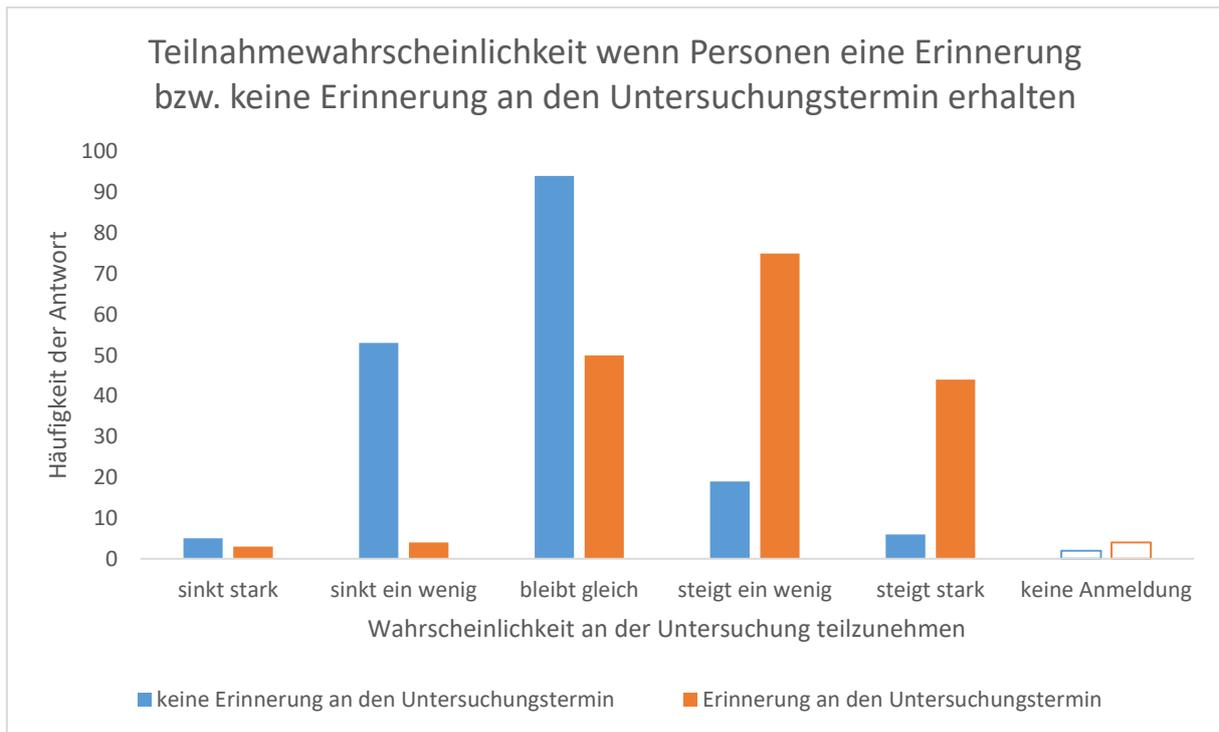
In der Folge werden die Ergebnisse der Hauptuntersuchung zu den aufgestellten Hypothesen berichtet.

Einflussfaktoren bei der Planung, Gestaltung und Durchführung der Untersuchung

Hypothese 1.1.: Erinnerungen an den Untersuchungstermin reduzieren den Ausfall von Versuchspersonen.

Bei dem Teil der Untersuchung, der an der Universität stattfand, erhielten elf Personen eine Erinnerung und 14 Personen keine Erinnerung. Zwei Personen wurden durch direktes Ansprechen rekrutiert. Der Untersuchungstermin wurde jeweils für eineinhalb Stunden nach der Rekrutierung vereinbart. Diese Personen wurden bei der Anzahl der Erinnerungen nicht berücksichtigt. Alle fünf ausgefallenen Versuchspersonen, die an der Untersuchung an der Universität teilnehmen sollten, haben keine Erinnerung an den Termin erhalten. Es wurde ein Chi²-Test durchgeführt, um zu testen, ob Personen, die keine Erinnerung an den Untersuchungstermin erhalten, häufiger ausfallen. Da zwei Zellen eine erwartete Häufigkeit besitzen, die weniger als fünf beträgt, hatte der Test keine Aussagekraft. Somit war eine sinnvolle statistische Auswertung dieser Daten aufgrund der kleinen Stichprobengröße nicht möglich.

Mehr als jeweils die Hälfte (119 von 180 bzw. 94 von 179) der Teilnehmer*innen in der Gesamtstichprobe gab an, dass die Wahrscheinlichkeit zur Untersuchung zu erscheinen „ein wenig steigt“ oder „stark steigt“, wenn die Versuchspersonen am Vortag eine Erinnerung an den Termin erhalten (siehe Diagramm 1; $n_{\text{gesamt}} = 180$, $n_{\text{steigt ein wenig}} = 75$, $n_{\text{steigt stark}} = 44$; Median = 4, Modus = 4) und dass die Wahrscheinlichkeit gleich bleibt, wenn sie keine Erinnerung erhalten (siehe Diagramm 1; $n_{\text{gesamt}} = 179$, $n_{\text{bleibt gleich}} = 94$; Median = 3, Modus = 3). Von den Personen, die an der Untersuchung an der Universität teilnahmen, gab mehr als die Hälfte an, dass die Wahrscheinlichkeit zum Termin zu erscheinen „ein wenig sinkt“, wenn sie keine Erinnerung erhalten ($n_{\text{gesamt}} = 23$, $n_{\text{sinkt ein wenig}} = 13$; Median = 2, Modus = 2).



*Diagramm 1: Häufigkeitsverteilungen der Teilnahmewahrscheinlichkeit, wenn Versuchspersonen von dem/der Versuchsleiter*in keine Erinnerung an den Untersuchungstermin erhalten (n = 179) bzw. wenn sie am Vortag der Untersuchung eine Erinnerung an den Termin erhalten (n = 180).*

Ein Wilcoxon's Vorzeichen-Rang-Test ergab, dass die Teilnahmewahrscheinlichkeit höher eingeschätzt wird, wenn Versuchspersonen eine Erinnerung erhalten, als wenn sie nicht an den Termin erinnert werden ($z = -8.248, p < .001, n = 174$). Die Effektstärke nach Cohen entspricht einem starken Effekt ($r = .63$). Erinnerungen an den Untersuchungstermin reduzieren den Ausfall von Versuchspersonen.

*Hypothese 1.2.: Die freie Auswahl eines Termins durch die Versuchsteilnehmer*innen reduziert die Ausfallquote.*

Mehr als zwei Drittel (124 von 180) der Versuchsteilnehmer*innen gab an, dass die Wahrscheinlichkeit tatsächlich an der Untersuchung teilzunehmen „ein wenig steigt“ oder „stark steigt“, wenn sie den Untersuchungstermin frei wählen können ($n_{\text{gesamt}} = 180, n_{\text{steigt ein wenig}} = 73, n_{\text{steigt stark}} = 51$; Median = 4, Modus = 4). Die freie Auswahl des Untersuchungstermins erhöht die Wahrscheinlichkeit, dass Versuchspersonen tatsächlich an der Untersuchung teilnehmen und reduziert somit die Ausfallquote.

Hypothese 1.3.: Die verwendeten Untersuchungsmethoden haben keinen Einfluss auf den Ausfall von Versuchspersonen.

Mehr als die Hälfte (96 von 180) der Versuchspersonen gab an, dass die Teilnahmewahrscheinlichkeit „gleich bleibt“, wenn sie mit technisch aufwendigen Hilfsmitteln (EEG, fMRT, etc.) untersucht werden ($n_{\text{gesamt}} = 180$, $n_{\text{bleibt gleich}} = 96$; Median = 3, Modus = 3). Außerdem gaben jeweils mehr als zwei Drittel der Versuchspersonen an, dass die Teilnahmewahrscheinlichkeit „gleich bleibt“, wenn sie einen Leistungstest bearbeiten sollen ($n_{\text{gesamt}} = 177$, $n_{\text{bleibt gleich}} = 120$; Median = 3, Modus = 3), wenn sie Fragebögen bearbeiten sollen ($n_{\text{gesamt}} = 179$, $n_{\text{bleibt gleich}} = 147$; Median = 3, Modus = 3) oder ihre aktive Teilnahme bei der Untersuchung besonders gefordert ist ($n_{\text{gesamt}} = 179$, $n_{\text{bleibt gleich}} = 127$; Median = 3, Modus = 3). Eine Friedman's Rang-Varianzanalyse ergab, dass es einen Unterschied in der zentralen Tendenz dieser Bedingungen gibt ($\chi^2_{3, n=166} = 13.918$, $p < .01$). Nach als Post-Hoc-Tests durchgeführten paarweisen Vergleichen, bei denen die Bonferroni-Korrektur angewandt wurde, war jedoch kein signifikanter Unterschied in der zentralen Tendenz der Teilnahmewahrscheinlichkeiten bei Einsatz der unterschiedlichen Untersuchungsmethoden nachweisbar. Welche Untersuchungsmethoden verwendet werden, hat keinen Einfluss auf die Ausfallquote.

Hypothese 1.4.: Die Ausfallquote ist bei kurzen Untersuchungen geringer als bei langen Untersuchungen.

In der Gesamtstichprobe gaben mehr als drei Viertel (165 von 180) der Personen an, dass die Teilnahmewahrscheinlichkeit gleich bleibt ein wenig steigt oder stark steigt, wenn die Untersuchung weniger als 30 Minuten dauert (siehe Diagramm 2; $n_{\text{gesamt}} = 180$, $n_{\text{bleibt gleich}} = 79$, $n_{\text{steigt ein wenig}} = 57$, $n_{\text{steigt stark}} = 29$; Median = 3, Modus = 3) und ebenso mehr als drei Viertel (155 von 180), dass sie gleich bleibt, ein wenig sinkt oder stark sinkt, wenn die Untersuchung mehr als 90 Minuten dauert (siehe Diagramm 2; $n_{\text{gesamt}} = 180$, $n_{\text{bleibt gleich}} = 82$, $n_{\text{sinkt ein wenig}} = 53$, $n_{\text{sinkt stark}} = 20$; Median = 3, Modus = 3). Außerdem gaben mehr als zwei Drittel (134 von 179) an, dass die Wahrscheinlichkeit gleich bleibt, wenn die Untersuchung zwischen 30 und 90 Minuten dauert (siehe Diagramm 2; $n_{\text{gesamt}} = 179$, $n_{\text{bleibt gleich}} = 134$; Median = 3, Modus = 3).

Besonderheiten zeigten sich bei der Stichprobe des Online-Panels SurveyCircle, bei dem jeweils mindestens die Hälfte (50 bzw. 71 von 84) der Personen angab, dass die Teilnahmewahrscheinlichkeit ein wenig oder stark steigt, wenn die Untersuchung weniger als 30 Minuten dauert ($n_{\text{gesamt}} = 84$, $n_{\text{steigt ein wenig}} = 35$, $n_{\text{steigt stark}} = 15$; Median = 4, Modus = 4) bzw. gleich bleibt, ein wenig oder stark sinkt, wenn sie länger als eineinhalb Stunden dauert (41 von 84; $n_{\text{gesamt}} = 84$, $n_{\text{bleibt gleich}} = 32$, $n_{\text{sinkt ein wenig}} = 31$, $n_{\text{sinkt stark}} = 10$; Median = 2, Modus = 3).

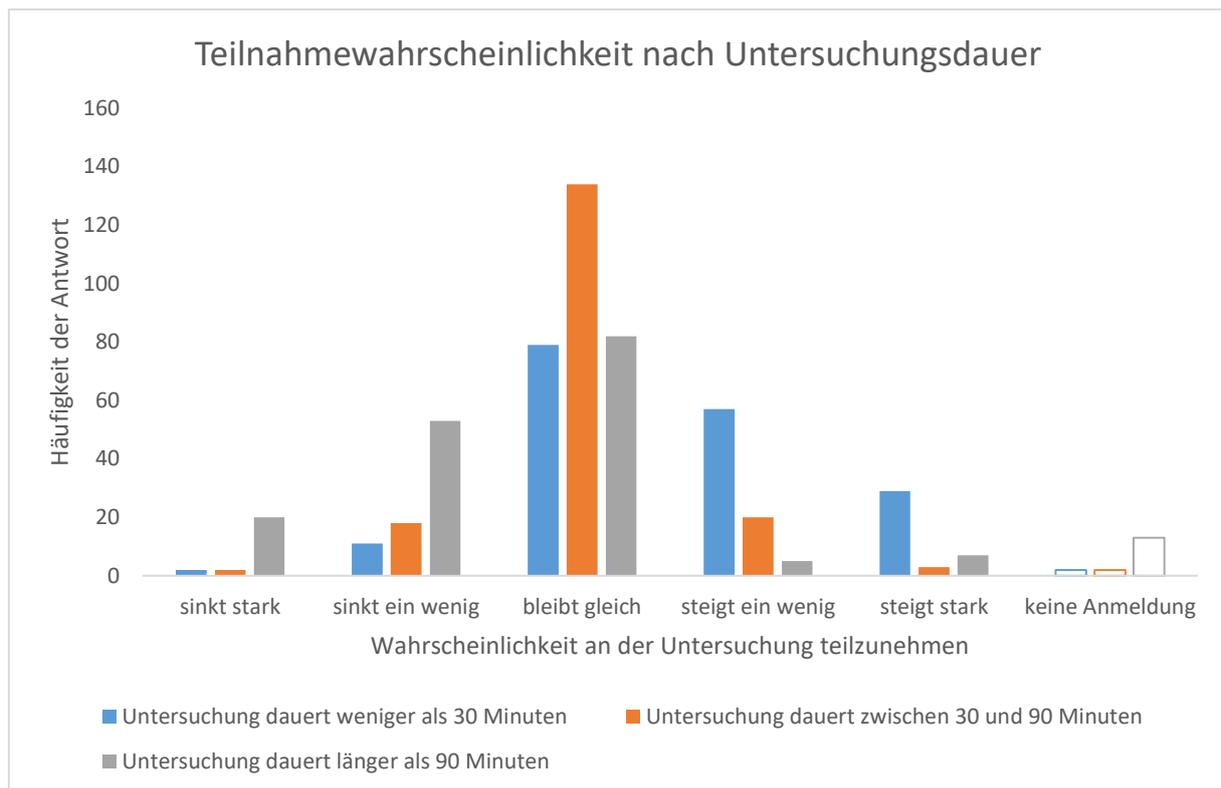


Diagramm 2: Häufigkeitsverteilungen der Teilnahmewahrscheinlichkeit, wenn die Untersuchung weniger als 30 Minuten dauert ($n = 180$), wenn sie zwischen 30 und 90 Minuten dauert ($n = 179$) bzw. wenn sie mehr als 90 Minuten dauert ($n = 180$).

Eine Friedman's Rang-Varianzanalyse ergab, dass sich die drei Bedingungen bezüglich ihrer zentralen Tendenz voneinander unterscheiden ($\chi^2_{2, n=162} = 85.115, p < .001$). Bei Untersuchungen die länger als 90 Minuten dauern ist die Einschätzung der Teilnahmewahrscheinlichkeit kleiner als bei Untersuchungen, die zwischen 30 und 90 Minuten dauern ($z = 3.500, p_{\text{angepasst}} < .01$, Effektstärke nach Cohen: $r = 0.27$) bzw. als bei Untersuchungen die weniger als 30 Minuten dauern ($z = 7.083, p_{\text{angepasst}} < .001$, Effektstärke nach Cohen: $r = 0.56$). Die Teilnahmewahrscheinlichkeit wird außerdem bei Untersuchungen höher eingeschätzt, die weniger als 30 Minuten dauern, als bei Untersuchungen, die zwischen 30 und 90 Minuten dauern ($z = 3.583, p_{\text{angepasst}} < .01$, Effektstärke nach Cohen: $r = 0.28$). Bei kürzeren Untersuchungen ist die Teilnahmewahrscheinlichkeit höher und damit die Ausfallquote geringer, als bei längeren Untersuchungen.

Zusammenfassung der Ergebnisse betreffend der Einflussfaktoren bei der Planung, Gestaltung und Durchführung der Untersuchung

Es wurde festgestellt, dass Erinnerungen an den Untersuchungstermin und eine freie Wahl des Untersuchungstermins die Teilnahmewahrscheinlichkeit von Versuchspersonen erhöhen und somit die Ausfallquote senken. Außerdem ist die Ausfallquote bei kürzeren Untersuchungen

niedriger als bei längeren Untersuchungen. Welche Untersuchungsmethoden bei einer Untersuchung eingesetzt werden, hat hingegen keinen Einfluss auf den Ausfall von Versuchspersonen.

Gestaltung des Rekrutierungsprozesses

Hypothese 2.1.: Persönlicher Kontakt bei der Rekrutierung senkt die Ausfallquote.

Jeweils mehr zwei Drittel oder mehr (123 von 179 bzw. 120 von 180) der Versuchspersonen gab an, dass sich die Teilnahmewahrscheinlichkeit ein wenig oder stark erhöht, wenn sie vor der Untersuchung den oder die Versuchsleiter*in persönlich getroffen und mit ihr oder ihm gesprochen haben (siehe Diagramm 3; 123 von 179; $n_{\text{gesamt}} = 179$, $n_{\text{steigt ein wenig}} = 90$, $n_{\text{steigt stark}} = 33$; Median = 4, Modus = 4) bzw. wenn sie per Mail oder SMS mit dem oder der Versuchsleiter*in kommuniziert haben (siehe Diagramm 3; 120 von 180; $n_{\text{gesamt}} = 180$, $n_{\text{steigt ein wenig}} = 83$, $n_{\text{steigt stark}} = 37$; Median = 4, Modus = 4). Außerdem gaben zwei Drittel (122 von 179) an, dass die Wahrscheinlichkeit gleich bleibt, wenn sie den Untersuchungstermin ohne persönlichen Kontakt vereinbart haben (siehe Diagramm 3; $n_{\text{gesamt}} = 179$, $n_{\text{bleibt gleich}} = 122$; Median = 3, Modus = 3).

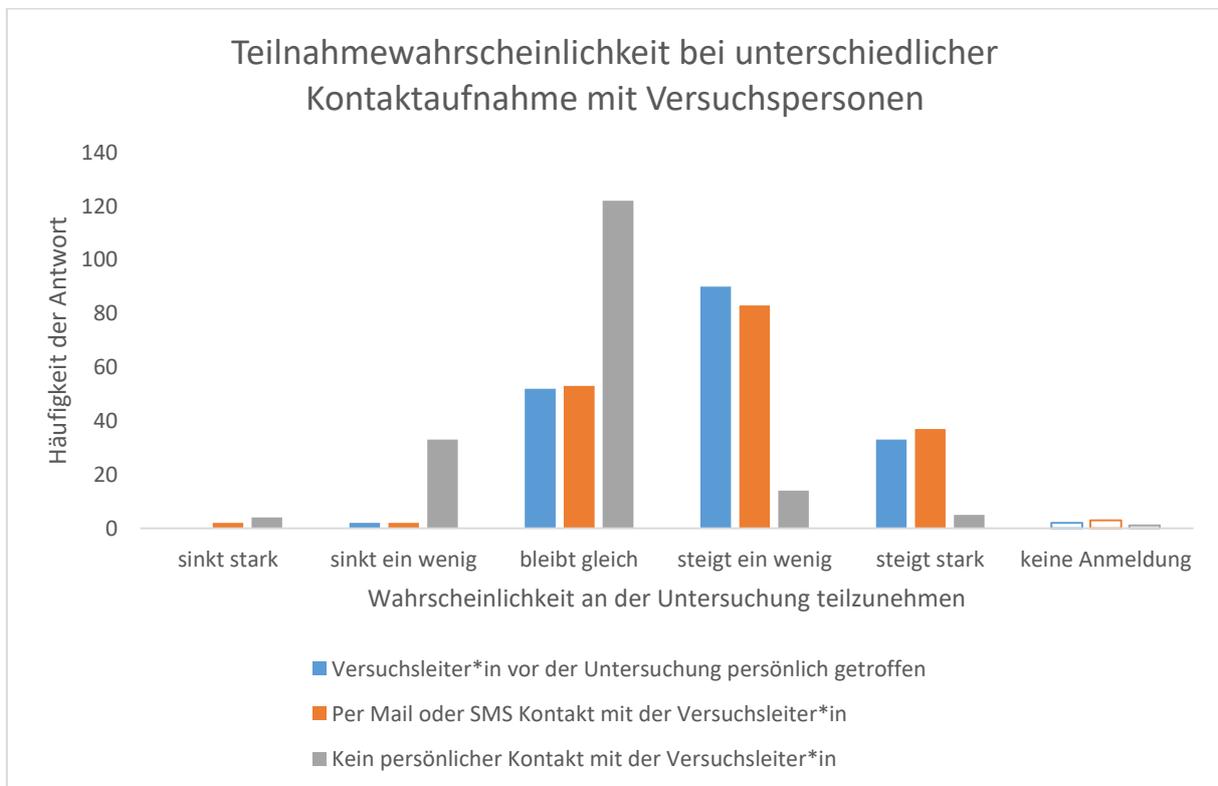


Diagramm 3: Häufigkeitsverteilungen der Teilnahmewahrscheinlichkeit, wenn die Versuchspersonen den/die Versuchsleiter*in vor der Untersuchung persönlich treffen und mit ihm/ihr sprechen ($n = 179$), mit dem/der Versuchsleiter*in per Mail oder SMS kommunizieren ($n = 180$) bzw. wenn kein persönlichen Kontakt zum/zur Versuchsleiter*in besteht ($n = 179$).

Durch eine Friedman's Rang-Varianzanalyse konnte ein Unterschied zwischen diesen Bedingungen festgestellt werden ($\chi^2_{2, n=175} = 158.036, p < .001$). Die Teilnahmewahrscheinlichkeit wird niedriger eingeschätzt, wenn der Untersuchungstermin ohne persönlichen Kontakt vereinbart wird, als wenn mit dem bzw. der Versuchsleiter*in persönlich gesprochen wird ($z = -8.499, p < .001$, Effektstärke nach Cohen: $r = .64$) oder wenn mit dieser bzw. diesem per Mail oder SMS kommuniziert wurde ($z = -8.739, p < .001$, Effektstärke nach Cohen: $r = .66$). Versuchspersonen erscheinen mit höherer Wahrscheinlichkeit zu einem Untersuchungstermin und reduzieren damit die Ausfallquote, wenn sie davor mit dem bzw. der Versuchsleiter*in persönlich gesprochen oder per Mail kommuniziert haben, als wenn der Termin ohne persönlichen Kontakt vereinbart wurde.

Personenspezifische Merkmale der Versuchsteilnehmer*innen

Hypothese 3.1.: Zuverlässigere Personen fallen als Versuchspersonen bei wissenschaftlichen Untersuchungen seltener aus.

Ein sinnvoller statistischer Vergleich der Zuverlässigkeit zwischen den Personen, die bei der Untersuchung an der Universität ausfielen und jenen, die erschienen, war aufgrund der kleinen Gruppengrößen nicht möglich.

Um den Zusammenhang zwischen der Zuverlässigkeit und ob Versuchspersonen bei Untersuchungen ausgefallen sind, überprüfen zu können, wurde ein Mediansplit bei dem Item „Für wie wahrscheinlich halten Sie es, dass Sie einen Versuchstermin verschlafen, vergessen oder aus anderen Gründen übersehen?“ (Median = 1; Mediansplit zwischen 1 und 2; 1 = „nicht wahrscheinlich“, 5 = „sehr wahrscheinlich“) und bei dem Item „Ich gehöre zu der Art von Menschen, auf die andere sich verlassen können.“ der SV-Skala (Median = „richtig“; Mediansplit zwischen „richtig“ und „sehr richtig“) durchgeführt. Aufgrund der geringen Zahl an Personen in der Gesamtstichprobe die ausgefallen sind, wäre bei Verwendung der ursprünglichen Kategorien in mehreren Zellen des Chi²-Tests, erwartete Häufigkeiten aufgetreten, die weniger als fünf betragen, wodurch die Tests nicht aussagekräftig gewesen wären.

In der Gesamtstichprobe besteht ein Zusammenhang zwischen dem Ausfall von Versuchspersonen bei mindestens einer Untersuchung und der Wahrscheinlichkeit, mit der Personen angegeben haben, einen Untersuchungstermin zu verschlafen, zu vergessen oder aus anderen Gründen zu übersehen ($\chi^2_{1, n=180} = 6.673, p_{\text{exakt}} < .05$, Effektstärke: $\phi = -.193$). Außerdem zeigte sich ein Zusammenhang zwischen der Empfindung von Personen, dass andere sich auf sie verlassen können und damit, ob sie bei mindestens einer Untersuchung ausgefallen sind ($\chi^2_{1, n=184} = 6.673,$

$p_{\text{exakt}} < .05$, Effektstärke: $\phi = .246$). Schließt man jene Personen aus, die noch keine oder nur wenige Gelegenheiten hatten, um auszufallen und begrenzt die Stichprobe entsprechend dem Median auf jene Hälfte, die zuvor an zumindest vier Untersuchungen teilgenommen hatte, bleiben der Zusammenhang mit dem „Verschlafen, vergessen oder übersehen eines Untersuchungstermins“ erhalten und die Effektstärke steigt ($\chi^2_{1, n=90} = 13.089$, $p_{\text{exakt}} < .01$, Effektstärke: $\phi = .381$). Der Zusammenhang damit, dass Andere sich auf die Versuchsperson verlassen können, konnte in diesem Fall nicht überprüft werden, da die erwartete Häufigkeit des Chi²-Tests in zwei Zellen weniger als fünf beträgt. Personen die sich selbst als zuverlässiger einschätzen, fallen bei Untersuchungen mit niedrigerer Wahrscheinlichkeit aus.

Zusätzlich konnte ein kleiner Zusammenhang zwischen der Einschätzung der Versuchspersonen, wie wahrscheinlich sie einen Untersuchungstermin verschlafen, vergessen oder aus anderen Gründen übersehen und der Häufigkeit, mit der diese Personen bei Untersuchungen ausgefallen sind, festgestellt werden ($r = .215$, $p < .01$). Je häufiger Personen bei Untersuchungen ausgefallen sind, umso größer schätzen sie die Wahrscheinlichkeit ein, einen Untersuchungstermin zu verschlafen, zu vergessen oder aus anderen Gründen zu übersehen. Begrenzt man die Stichprobe wiederum auf jene Personen, die bereits an mindestens vier Untersuchungen teilgenommen hatten, so zeigt sich ein mittelstarker Zusammenhang zwischen den beiden Variablen ($r = .385$, $p < .001$). Zwischen der Häufigkeit des Ausfalls und der Angabe, wie sehr sich andere Personen auf die Versuchspersonen verlassen können, besteht kein Zusammenhang ($r = -.092$, n.s.). Auch eine Begrenzung der Stichprobe auf jene Personen, die bereits an mindestens vier Untersuchungen teilgenommen hatten, ändert daran nichts ($r = -.145$, n.s.). Personen die sich als weniger zuverlässig einschätzen, fallen bei Untersuchungen häufiger aus.

Hypothese 3.2.: Gewissenhaftere Personen fallen als Versuchspersonen bei wissenschaftlichen Untersuchungen seltener aus.

Ein statistischer Vergleich der Gewissenhaftigkeit zwischen den Personen, die bei der Untersuchung an der Universität ausgefallen sind und jenen, die zum Termin erschienen sind, war aufgrund der kleinen Gruppengrößen nicht möglich.

Durch einen Chi²-Test zeigte sich kein Zusammenhang zwischen der Gewissenhaftigkeit und damit, ob Versuchspersonen bereits bei wissenschaftlichen Untersuchungen ausgefallen sind ($\chi^2_{1, n=184} = .148$, n.s.). Die Ausprägung der Gewissenhaftigkeit korreliert jedoch schwach damit, wie oft Versuchspersonen schon bei wissenschaftlichen Untersuchungen ausgefallen sind ($r = -.148$, $p < .05$). Weniger gewissenhafte Menschen fallen häufiger bei Untersuchungen aus. Begrenzt man die Stichprobe wiederum auf Personen, die an mindestens vier Untersuchungen

teilgenommen hatten, zeigt sich kein Zusammenhang zwischen der Gewissenhaftigkeit und dem Ausfall von Personen ($\chi^2_{1, n=92} = .009$, n.s.) bzw. mit der Häufigkeit des Ausfalls ($r = -.188$, n.s.). Die Gewissenhaftigkeit von Versuchspersonen hat keinen Einfluss darauf, ob und wie häufig Versuchspersonen bei wissenschaftlichen Untersuchungen ausfallen.

Hypothese 3.3.: Versuchspersonen, die eine höhere soziale Verantwortung aufweisen, fallen bei wissenschaftlichen Untersuchungen seltener aus.

Es zeigt sich für den Ausfall bei mindestens einer Untersuchung, weder eine Assoziation mit der Gesamtskala der sozialen Verantwortung ($\chi^2_{1, n=184} = 2.646$, n.s.), noch mit einer der Subskalen „Befolgung von sozialen Spielregeln“ ($\chi^2_{1, n=184} = .385$, n.s.) und „Erfüllung von berechtigten Erwartungen anderer“ ($\chi^2_{1, n=184} = 4.659$, n.s.). Auch eine Begrenzung der Stichprobe auf Personen, die zuvor an mindestens vier Untersuchungen teilgenommen hatten, ändert daran nichts (SV-Skala: $\chi^2_{1, n=92} = 3.621$, n.s.; Befolgung von sozialen Spielregeln: $\chi^2_{1, n=92} = .847$, n.s.; Erfüllung von berechtigten Erwartungen anderer: $\chi^2_{1, n=92} = 3.510$, n.s.). Die soziale Verantwortung von Personen hat keinen Einfluss darauf, ob diese Personen bei wissenschaftlichen Untersuchungen ausfallen.

Es besteht jedoch ein schwacher Zusammenhang zwischen der Häufigkeit, mit der Personen bei Untersuchungen ausgefallen sind und damit, wie sehr sie die berechtigten Erwartungen anderer erfüllen ($r = -.157$, $p < .05$), jedoch keiner mit der gesamten SV-Skala ($r = -.140$, n.s.) und der Befolgung von sozialen Spielregeln ($r = -.021$, n.s.). Betrachtet man nur Personen, die zuvor an mindestens vier Untersuchungen teilgenommen hatten, ist dieser Zusammenhang ein wenig stärker ($r = -.209$, $p < .05$). Für die Befolgung sozialer Spielregeln ($r = -.004$, n.s.) und die Gesamtskala ($r = -.177$, n.s.) lässt sich auch hier keine signifikante Korrelation nachweisen. Personen, die bei mehreren Untersuchungen ausfallen, erfüllen die berechtigten Erwartungen anderer weniger.

Zusammenfassung der Ergebnisse betreffend personenspezifischer Merkmale

Die Ergebnisse zeigen, dass zuverlässigere Personen seltener bei wissenschaftlichen Untersuchungen ausfallen, während die Gewissenhaftigkeit und die Soziale Verantwortung von Versuchspersonen keinen Einfluss auf die Ausfallquote haben.

Hypothese 4.1.: Aufwendige Untersuchungen weisen eine höhere Ausfallquote auf.

Jeweils mehr als die Hälfte (96 von 180 bzw. 127 von 179) der Versuchspersonen gab an, dass die Teilnahmewahrscheinlichkeit gleich bleibt, wenn bei der Untersuchung aufwendige technische Hilfsmittel wie EEG oder fMRT zum Einsatz kommen ($n_{\text{gesamt}} = 180$, $n_{\text{bleibt gleich}} = 96$; Median = 3, Modus = 3) oder ihre aktive Teilnahme besonders gefordert ist ($n_{\text{gesamt}} = 179$, $n_{\text{bleibt gleich}} = 127$; Median = 3, Modus = 3). Außerdem gab mehr als zwei Drittel (146 von 179) der Versuchspersonen an, dass die Wahrscheinlichkeit teilzunehmen gleich bleibt, ein wenig steigt oder stark steigt, wenn sie nach einer Voruntersuchung zur Hauptuntersuchung eingeladen werden ($n_{\text{gesamt}} = 179$, $n_{\text{bleibt gleich}} = 82$, $n_{\text{steigt ein wenig}} = 46$, $n_{\text{steigt stark}} = 18$; Median = 3, Modus = 3). Aufwendigere Untersuchungsdesigns senken nur dann die Teilnahmewahrscheinlichkeit, wenn eine Untersuchung aus mehreren Terminen besteht. In diesem Fall gab mehr als die Hälfte (113 von 179) der Versuchspersonen an, dass die Teilnahmewahrscheinlichkeit sinken würde ($n_{\text{sinkt ein wenig}} = 88$, $n_{\text{sinkt stark}} = 25$; Median = 2, Modus = 2). Bei den Personen, die bei der Untersuchung an der Universität teilnahmen, wurde eine andere Verteilung beobachtet ($n_{\text{gesamt}} = 23$, $n_{\text{sinkt stark}} = 2$, $n_{\text{sinkt ein wenig}} = 6$, $n_{\text{bleibt gleich}} = 8$, $n_{\text{steigt ein wenig}} = 6$, $n_{\text{keine Anmeldung}} = 1$; Median = 3, Modus = 3). Untersuchungen mit mehreren Terminen weisen also eine niedrigere Teilnahmewahrscheinlichkeit auf, während die Teilnahmewahrscheinlichkeit eher gleich bleibt, wenn eine Untersuchung aus Vor- und Hauptuntersuchung besteht, wenn ihre aktive Teilnahme besonders gefordert ist oder technische Hilfsmittel zum Einsatz kommen. Aufwendigere Untersuchungen können, müssen aber keine höhere Ausfallquote aufweisen.

Hypothese 4.2.: Monetäre Anreize für die Teilnahme reduzieren die Ausfallquote.

Bezüglich monetärer Anreize gab je mehr als die Hälfte der Teilnehmer*innen an, dass die Teilnahmewahrscheinlichkeit ein wenig oder stark steigt, wenn sie für die Teilnahme an einer Untersuchung 5€ (siehe Diagramm 4; 94 von 180; $n_{\text{gesamt}} = 180$, $n_{\text{steigt ein wenig}} = 67$, $n_{\text{steigt stark}} = 27$; Median = 4, Modus = 3) bzw. 10€ pro Stunde erhalten (siehe Diagramm 4; 132 von 180; $n_{\text{gesamt}} = 180$, $n_{\text{steigt ein wenig}} = 67$, $n_{\text{steigt stark}} = 65$; Median = 4, Modus = 4). Wenn 30€ für eine Untersuchungsteilnahme geboten werden, gab mehr als drei Viertel (138 von 179) der Teilnehmer*innen an, dass die Teilnahmewahrscheinlichkeit steigt (siehe Diagramm 4; $n_{\text{gesamt}} = 179$, $n_{\text{steigt ein wenig}} = 52$, $n_{\text{steigt stark}} = 86$; Median = 5, Modus = 5).

Bei der Untersuchung an der Universität und den Personen die über SurveyCircle rekrutiert wurden gab es Abweichungen von der Gesamtstichprobe. Mehr als die Hälfte (13 von 23) der

Personen, die bei der Untersuchung an der Universität teilnahmen, gab an, dass die Wahrscheinlichkeit an einer Untersuchung teilzunehmen, stark steigt, wenn sie 10€ pro Stunde dafür erhalten ($n_{\text{gesamt}} = 23$, $n_{\text{steigt stark}} = 13$; Median = 5, Modus = 5). Von den Teilnehmer*innen der Umfrage über „SurveyCircle“ gaben drei Viertel (63 von 84) an, dass die Teilnahmewahrscheinlichkeit ein wenig oder stark steigt, wenn sie 30€ für die Teilnahme bekommen ($n_{\text{gesamt}} = 84$, $n_{\text{steigt ein wenig}} = 31$, $n_{\text{steigt stark}} = 32$ Median = 4, Modus = 5). Der Einsatz monetärer Anreize für die Teilnahme an Untersuchungen erhöht die Teilnahmewahrscheinlichkeit und reduziert damit die Ausfallquote.

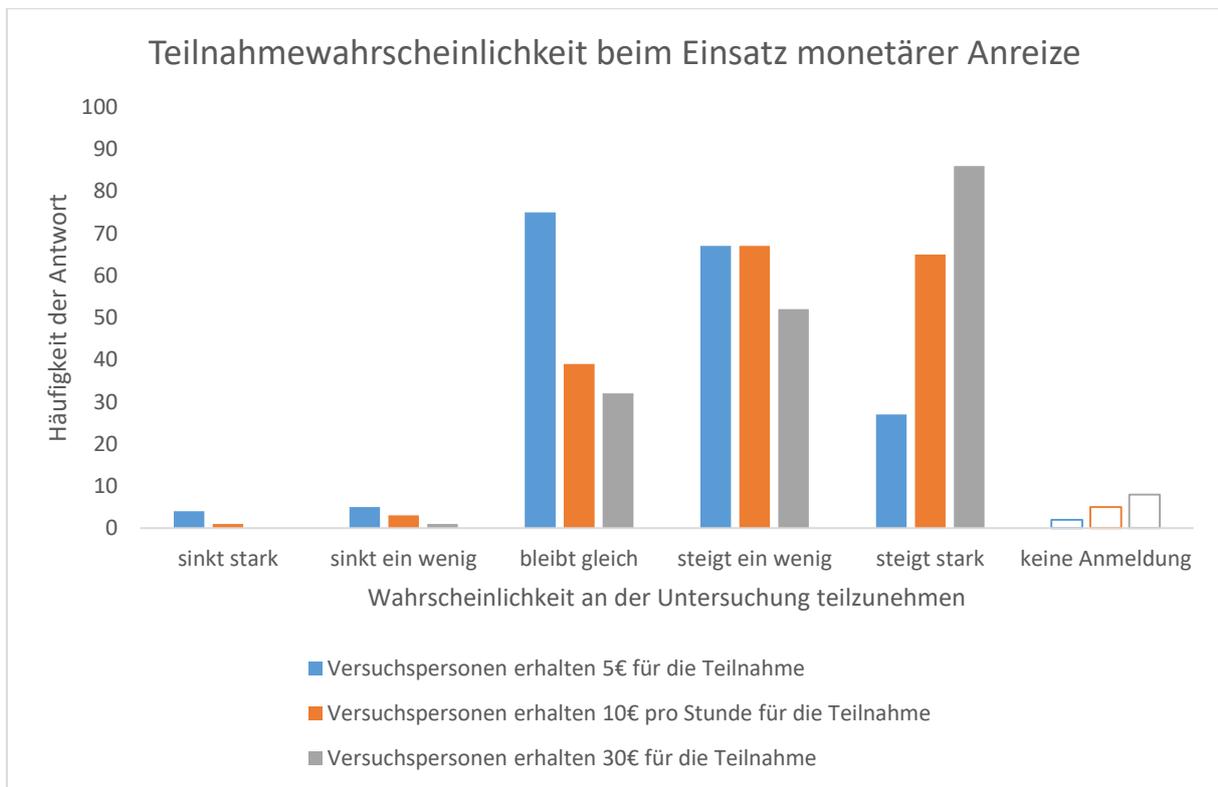


Diagramm 4: Häufigkeitsverteilungen der Teilnahmewahrscheinlichkeit, wenn Versuchspersonen für die Teilnahme an der Untersuchung 5€ ($n = 180$), 10€ pro Stunde ($n = 180$) bzw. 30€ ($n = 179$) erhalten.

Hypothese 4.3.: Die Möglichkeit der Teilnahme an einer Preisverlosung reduziert die Ausfallquote.

Mehr als die Hälfte (114 von 180) der Versuchspersonen gab an, dass die Teilnahmewahrscheinlichkeit gleich bleibt, wenn sie nach der Untersuchung an einem Gewinnspiel teilnehmen dürfen, während mehr als ein Drittel (61 von 180) angab, dass die Wahrscheinlichkeit ein wenig oder stark steigt ($n_{\text{gesamt}} = 180$, $n_{\text{bleibt gleich}} = 114$, $n_{\text{steigt ein wenig}} = 50$, $n_{\text{steigt stark}} = 11$; Median = 3, Modus = 3). Um eine klare Aussage über den Einfluss einer Preisverlosung bzw. eines Gewinnspiels auf die Ausfallquote treffen zu können, müsste die Verteilung dieser Teilnahmewahrscheinlichkeit mit einem Referenzitem verglichen werden. Die Erhebung solch eines Items

wurde bei der Planung dieser Untersuchung allerdings nicht berücksichtigt. Die Möglichkeit der Teilnahme an einer Preisverlosung hat entweder keinen Einfluss auf die Ausfallquote oder senkt diese. Eine genauere Aussage dazu kann anhand der vorliegenden Daten nicht getroffen werden.

Hypothese 4.4.: Personen, die dazu verpflichtet sind, sich als Versuchsperson zur Verfügung zu stellen, fallen seltener aus.

Von den 26 Personen, die an der Untersuchung an der Universität nachträglich online oder direkt teilnahmen, benötigten 18 einen Untersuchungsschein. Von den vier ausgefallenen Personen, die an der Untersuchung teilnahmen, benötigten drei einen Untersuchungsschein. Eine sinnvolle statistische Auswertung dieser Daten war aufgrund der kleinen Stichprobengröße nicht möglich.

Hypothese 4.5.: Versuchspersonen mit Interesse am Thema der Untersuchung fallen seltener aus als Versuchspersonen ohne Interesse am Thema.

Eine der vier bei der Untersuchung an der Universität ausgefallenen Personen – die nachträglich online an der Untersuchung teilgenommen hatten – gab an, dass sie sich aus Interesse an der Untersuchung angemeldet hatte. Insgesamt machten 14 der 26 Versuchspersonen diese Angabe. Eine sinnvolle statistische Auswertung dieser Daten war aufgrund der kleinen Stichprobengröße nicht möglich.

Mehr als drei Viertel (148 von 179) der Personen in der Gesamtstichprobe gaben an, dass die Teilnahmewahrscheinlichkeit ein wenig oder stark steigt, wenn sie das Thema einer Untersuchung besonders interessant finden ($n_{\text{gesamt}} = 179$, $n_{\text{steigt ein wenig}} = 86$, $n_{\text{steigt stark}} = 62$; siehe Diagramm 5; 148 von 179; Median = 4, Modus = 4). Mehr als zwei Drittel (135 von 180) der Versuchspersonen gaben an, dass die Teilnahmewahrscheinlichkeit gleich bleibt ($n = 50$), ein wenig ($n = 58$) oder stark sinken würde ($n = 27$), wenn sie sich zu einer Untersuchung anmelden, deren Thema sie nicht interessant finden, (siehe Diagramm 5; $n_{\text{gesamt}} = 180$, $n_{\text{bleibt gleich}} = 50$, $n_{\text{sinkt ein wenig}} = 58$, $n_{\text{sinkt stark}} = 27$; Median = 2, Modus = 2). Dabei gab es jedoch Abweichungen bei den Personen, die an der Untersuchung an der Universität teilnahmen. Mehr als drei Viertel (19 von 23) dieser Personen gab an, dass die Wahrscheinlichkeit gleich bleibt oder ein wenig sinkt, wenn sie das Thema der Untersuchung nicht interessant finden ($n_{\text{gesamt}} = 23$, $n_{\text{bleibt gleich}} = 11$, $n_{\text{sinkt ein wenig}} = 8$; Median = 3, Modus = 3).

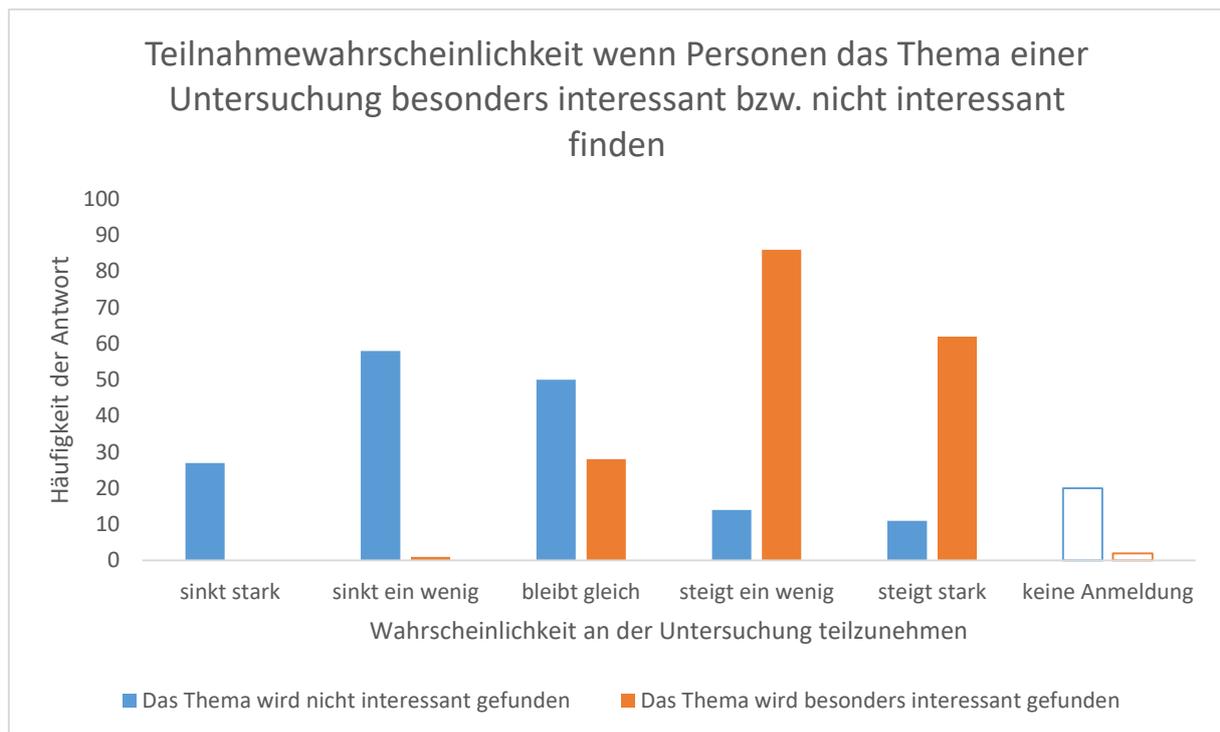


Diagramm 5: Häufigkeitsverteilungen der Teilnahmewahrscheinlichkeit, wenn Versuchspersonen das Thema der Untersuchung nicht interessant ($n = 180$) bzw. besonders interessant ($n = 179$) finden.

Ein Wilcoxon's Vorzeichen-Rang-Test ergab, dass die Einschätzung der Teilnahmewahrscheinlichkeit signifikant höher ist, wenn Versuchspersonen das Thema der Untersuchung besonders interessant finden, als wenn sie das Thema nicht interessant finden ($z = 9.185$, $p < .001$, $n = 157$). Die Effektstärke liegt bei 0.73 und entspricht damit einem starken Effekt. Versuchspersonen mit Interesse am Thema der Untersuchung nehmen mit höherer Wahrscheinlichkeit tatsächlich an dieser Untersuchung teil, als Personen die das Thema der Untersuchung nicht interessiert.

Hypothese 4.6.: Die Ausfallquote ist bei Untersuchungen kleiner, wenn die Anreize einen kleineren Wert haben, als wenn wertvollere Anreize geboten werden.

Mehr als drei Viertel (146 von 179) der Versuchspersonen gaben an, dass die Teilnahmewahrscheinlichkeit gleich bleibt, wenn sie Süßigkeiten für die Teilnahme erhalten ($n_{\text{gesamt}} = 179$, $n_{\text{bleibt gleich}} = 146$; Median = 3, Modus = 3). Wie bereits berichtet, gab mehr als die Hälfte (114 von 180) der Versuchspersonen an, dass die Teilnahme bei einem Gewinnspiel die Teilnahmewahrscheinlichkeit nicht beeinflusst ($n_{\text{gesamt}} = 180$, $n_{\text{bleibt gleich}} = 114$, $n_{\text{steigt ein wenig}} = 50$, $n_{\text{steigt stark}} = 11$; Median = 3, Modus = 3). Wenn Geld geboten wurde, gab mehr als die Hälfte der Versuchspersonen an, dass die Wahrscheinlichkeit zur Untersuchung zu erscheinen ein wenig oder stark steigt, wenn sie 5€ (94 von 180; $n_{\text{gesamt}} = 180$, $n_{\text{steigt ein wenig}} = 67$, $n_{\text{steigt stark}} = 27$; Median =

4, Modus = 3) bzw. 10€ pro Stunde dafür erhalten (132 von 180; $n_{\text{gesamt}} = 180$, $n_{\text{steigt ein wenig}} = 67$, $n_{\text{steigt stark}} = 65$; Median = 4, Modus = 4). Die Teilnahmewahrscheinlichkeit steigt laut mehr als zwei Drittel (138 von 179) der Versuchspersonen ein wenig oder stark, wenn sie 30€ für eine Teilnahme erhalten ($n_{\text{gesamt}} = 179$, $n_{\text{steigt ein wenig}} = 52$, $n_{\text{steigt stark}} = 86$; Median = 5, Modus = 5). Die angegebene Teilnahmewahrscheinlichkeit unterscheidet sich zwischen den unterschiedlichen Anreizen (Friedman's Rang-Varianzanalyse: $\chi^2_{4, n=166} = 333.269$, $p < .001$). Eine durchgeführte Post-Hoc Analyse inklusive Bonferroni-Korrektur zeigte, dass wertvollere Belohnungen eine höhere Teilnahmewahrscheinlichkeit nach sich ziehen (zwischen $z = -5.050$, $p_{\text{angepasst}} < .001$, Effektstärke: $r = 0.39$ und $z = -12.305$, $p_{\text{angepasst}} < .001$, Effektstärke: $r = 0.96$ – Daten im Detail in Tabelle 9 auf S.77). Die Teilnahmewahrscheinlichkeit unterscheidet sich nur nicht, wenn Süßigkeiten oder eine Gewinnspielteilnahme als Anreiz versprochen werden ($z = -2.586$, $p_{\text{angepasst}} = .097$), bei einer Gewinnspielteilnahme oder 5€ als Entschädigung und auch nicht ($z = -2.464$, $p_{\text{angepasst}} = .137$), wenn Personen 10€ pro Stunde oder 30€ für die Teilnahme erhalten würden ($z = -1.874$, $p_{\text{angepasst}} = .609$). Die Teilnahmewahrscheinlichkeit ist bei Untersuchungen, bei denen ein wertvollerer Anreiz zur Teilnahme geboten wird, höher als bei Untersuchungen bei denen die Anreize einen kleineren Wert haben.

Zusammenfassung der Ergebnisse betreffend der motivationalen Faktoren

Der Einsatz monetärer Anreize reduziert den Ausfall von Versuchspersonen an einer Untersuchung. Dabei gilt auch, je wertvoller ein gebotener Anreiz ist, umso stärker motiviert er Versuchspersonen auch tatsächlich zur Untersuchung zu erscheinen. Besonderes Interesse am Untersuchungsthema senkt die Ausfallquote ebenfalls, während es den Ausfall von Versuchspersonen wahrscheinlicher macht, wenn diese kein Interesse am Thema der Untersuchung haben. Auch aufwendige Untersuchungen können sich negativ auf die Ausfallquote auswirken – etwa wenn die Untersuchung aus mehreren Terminen besteht – müssen dies aber nicht tun. Wie sich Preisverlosungen bzw. Gewinnspiele auf die Untersuchung auswirken konnte anhand der Daten nicht eindeutig geklärt werden, es besteht die Möglichkeit, dass sie die Ausfallquote senken oder aber keinen Einfluss auf die Ausfallquote haben. Ebenso unklar bleibt, ob und wie sich Versuchsscheine auf die Ausfallquote auswirken, da dies anhand der erhobenen Daten nicht analysiert werden konnte.

Unterschiede zwischen Bedingungen bezüglich der Teilnahmewahrscheinlichkeit

<i>Bedingung unterscheidet sich</i>	<i>von Bedingung</i>	<i>z</i>	<i>p</i>	<i>r</i>
<i>Wenn ich Süßigkeiten für die Teilnahme bekomme...</i>	<i>Wenn ich nach der Untersuchungsteilnahme bei einem Gewinnspiel mitmachen darf...</i>	-2.59	<i>n.s.</i>	-
<i>Wenn ich Süßigkeiten für die Teilnahme bekomme...</i>	<i>Wenn ich 5€ für die Teilnahme erhalte...</i>	-5.05	<.001	0.39
<i>Wenn ich Süßigkeiten für die Teilnahme bekomme...</i>	<i>Wenn ich 10€/Stunde für die Teilnahme erhalte...</i>	-10.43	<.001	0.81
<i>Wenn ich Süßigkeiten für die Teilnahme bekomme...</i>	<i>Wenn ich 30€ für die Teilnahme an der Untersuchung erhalte...</i>	-12.31	<.001	0.96
<i>Wenn ich nach der Untersuchungsteilnahme bei einem Gewinnspiel mitmachen darf...</i>	<i>Wenn ich 5€ für die Teilnahme erhalte...</i>	-2.46	<i>n.s.</i>	-
<i>Wenn ich nach der Untersuchungsteilnahme bei einem Gewinnspiel mitmachen darf...</i>	<i>Wenn ich 10€/Stunde für die Teilnahme erhalte...</i>	-7.85	<.001	0.61
<i>Wenn ich nach der Untersuchungsteilnahme bei einem Gewinnspiel mitmachen darf...</i>	<i>Wenn ich 30€ für die Teilnahme an der Untersuchung erhalte...</i>	-9.72	<.001	0.75
<i>Wenn ich 5€ für die Teilnahme erhalte...</i>	<i>Wenn ich 10€/Stunde für die Teilnahme erhalte...</i>	-5.38	<.001	0.42
<i>Wenn ich 5€ für die Teilnahme erhalte...</i>	<i>Wenn ich 30€ für die Teilnahme an der Untersuchung erhalte...</i>	-7.26	<.001	0.56
<i>Wenn ich 10€/Stunde für die Teilnahme erhalte...</i>	<i>Wenn ich 30€ für die Teilnahme an der Untersuchung erhalte...</i>	-1.87	<i>n.s.</i>	-

Tabelle 9: Bedingungen die sich bezüglich der Teilnahmewahrscheinlichkeit unterscheiden. Berichtet werden die standardisierte Teststatistik z, mit einer Bonferroni-Korrektur angepasste p-Werte und die Effektstärke r einer Friedman's Rang-Varianzanalyse.

Hypothese 5.1.: Personen, die eine aufwendigere Anreise haben, fallen bei wissenschaftlichen Untersuchungen eher aus.

Die Dauer der Anreise zur Universität wurde nur bei den Personen erhoben, die sich zur Untersuchung an der Universität angemeldet haben (Minimum = 2 Minuten, Maximum = 50 Minuten, \bar{x} = 19.78 Minuten, sd = 9.826 Minuten, linkssteil verteilt). Dadurch war die Stichprobengröße zu klein für eine statistische Auswertung der Daten. Die ausgefallenen Versuchspersonen gaben an, zur Universität eine Anreise von 2, 15, 20 oder 30 Minuten zu haben.

Hypothese 5.2.: Versuchspersonen, die eine hohe Geschäftigkeit empfinden, fallen bei wissenschaftlichen Untersuchungen eher aus.

Jeweils mehr als zwei Drittel der Versuchspersonen gaben an, dass die Teilnahmewahrscheinlichkeit ein wenig oder stark sinkt, wenn ihr/e Chef*in sie (kurz) vor dem Untersuchungstermin bittet, bei der Arbeit einzuspringen (143 von 180; $n_{\text{gesamt}} = 180$, $n_{\text{sinkt ein wenig}} = 69$, $n_{\text{sinkt stark}} = 74$; Median = 2, Modus = 1), sie ein Familienmitglied vor dem Untersuchungstermin um Hilfe bittet (138 von 180; $n_{\text{gesamt}} = 180$, $n_{\text{sinkt ein wenig}} = 70$, $n_{\text{sinkt stark}} = 68$; Median = 2, Modus = 2). Mehr als die Hälfte (116 von 180) der Versuchspersonen gab an, dass die Teilnahmewahrscheinlichkeit sinkt, wenn ein/e Freund*in sie vor dem Untersuchungstermin spontan um Hilfe bittet ($n_{\text{gesamt}} = 180$, $n_{\text{sinkt ein wenig}} = 79$, $n_{\text{sinkt stark}} = 37$; Median = 2, Modus = 2). Versuchspersonen die aufgrund äußerer Umstände eine hohe Geschäftigkeit empfinden, haben eine geringere Teilnahmewahrscheinlichkeit bei Untersuchungen.

Hypothese 5.3.: Bei schlechtem Wetter fallen Versuchspersonen eher aus.

Von den 27 Personen, die sich für die Untersuchung an der Universität anmeldeten, fielen vier von 16 bei bewölktem Wetter und eine von zwei Personen bei regnerischem Wetter aus. Bei sonnigem Wetter (sechs Termine) oder teilweise sonnigem Wetter (zwei Termine) fielen keine Versuchspersonen aus. Zu den gewählten Untersuchungsterminen hatte es zwischen 6°C und 21°C. Zu den Terminen zu denen Versuchspersonen ausfielen, hatte es 7°C, 8°C, 12°C und 15°C.

Je mehr als drei Viertel der Versuchspersonen gaben an, dass die Teilnahmewahrscheinlichkeit gleich bleibt, wenn draußen besonders schönes Wetter (131 von 180; $n_{\text{gesamt}} = 180$, $n_{\text{bleibt gleich}} = 131$; Median = 3, Modus = 3) oder besonders schlechtes Wetter ist (133 von 180; $n_{\text{gesamt}} = 180$, $n_{\text{bleibt gleich}} = 133$; Median = 3, Modus = 3). Ist draußen leicht bewölktes Wetter, so bleibt die Wahrscheinlichkeit einer Teilnahme für 90% der Versuchspersonen gleich (162 von 180; n_{gesamt}

= 180, $n_{\text{bleibt gleich}} = 162$; Median = 3, Modus = 3). Eine Friedman's Rang-Varianzanalyse ergab, dass es einen Unterschied in der zentralen Tendenz bei unterschiedlichem Wetter gibt ($\chi^2_{2, n=170} = 19.359$, $p < .001$). Nach Post-Hoc Tests, bei denen eine Bonferroni-Korrektur angewandt wurde, war jedoch kein signifikanter Unterschied zwischen einzelnen Bedingungen nachweisbar. Die Teilnahmewahrscheinlichkeit ändert sich nicht, wenn draußen besonders schönes, besonders schlechtes oder leicht bewölktetes Wetter ist.

Zusammenfassung der Ergebnisse betreffend der Umweltfaktoren

Bei der Untersuchung der Umweltfaktoren konnte festgestellt werden, dass die gefühlte Geschäftigkeit (im Englischen: busyness) den Ausfall von Versuchspersonen wahrscheinlicher macht. Das Wetter hat hingegen keinen Einfluss darauf, ob Versuchspersonen tatsächlich zum vereinbarten Untersuchungstermin erscheinen. Ob und welchen Einfluss eine aufwendigere Anreise auf die Ausfallquote hat, bleibt unklar, da dies anhand der erhobenen Daten nicht analysiert werden konnte.

Ergänzende Fragestellungen

Ergänzende Fragestellung 6.1.: Haben die verwendeten Rekrutierungsmaßnahmen Einfluss auf die Ausfallquote?

Von den 27 Personen, die sich zur Untersuchung an der Universität angemeldet hatten, sind zwei der acht Personen, die sich über den Aushang an der Universität angemeldet hatten und drei der zwölf Personen, die sich über den Mail-Verteiler angemeldet hatten, ausgefallen. Bei der Rekrutierung über soziale Medien (fünf Personen) oder durch direktes Ansprechen (zwei Personen) gab es keine Ausfälle. Eine Person, die sich über die Ausschreibung in den sozialen Medien angemeldet hatte, hat sich jedoch am Tag der Untersuchung wieder abgemeldet. Unter all diesen Bedingungen haben sich zu wenige Personen angemeldet, um einen sinnvollen statistischen Vergleich zuzulassen.

Ergänzende Fragestellung 6.2.: Hat die Verwendung der Höflichkeitsform Einfluss auf die Ausfallquote?

Wie die Versuchspersonen in Aussendungen angesprochen werden, hatte für einen großen Teil der Versuchspersonen keinen Einfluss auf die Teilnahmewahrscheinlichkeit. 86.67% (156 von 180) gaben an, dass die Wahrscheinlichkeit gleich bleibt, wenn sie mit „Sie“ angesprochen werden ($n_{\text{gesamt}} = 180$, $n_{\text{bleibt gleich}} = 156$; Median = 3, Modus = 3) und 86.59% (155 von 179) gaben an, dass sie gleich bleibt, wenn sie mit „Du“ angesprochen werden ($n_{\text{gesamt}} = 179$, n_{bleibt}

gleich = 155; Median = 3, Modus = 3). Ein Wilcoxon's Vorzeichen-Rang-Test zeigte keinen Unterschied zwischen diesen beiden Bedingungen ($z = -1.768$, n.s., $n = 177$). Ob Versuchspersonen mit „Du“ oder „Sie“ angesprochen werden hat keinen Einfluss auf die Teilnahmewahrscheinlichkeit.

*Ergänzende Fragestellung 6.3.: Hat ein Näheverhältnis zum/zur Versuchsleiter*in Einfluss auf die Ausfallquote?*

Mehr als zwei Drittel (122 von 180) der Versuchspersonen berichteten, dass die Teilnahmewahrscheinlichkeit gleich bleibt, wenn die Untersuchung von einer oder einem Fremden geleitet wird ($n_{\text{gesamt}} = 180$, $n_{\text{bleibt gleich}} = 122$; Median = 3, Modus = 3). Im Gegensatz dazu gaben zwei Drittel (117 von 180) der Personen an, dass sich die Teilnahmewahrscheinlichkeit stark erhöht, wenn sie sich zu einer Untersuchung anmelden, bei der eine mit ihnen befreundete Person als Versuchsleiter*in fungiert ($n_{\text{gesamt}} = 180$, $n_{\text{steigt stark}} = 117$; Median = 5, Modus = 5). Durch einen Wilcoxon's Vorzeichen-Rang-Test zeigte sich, dass die Teilnahmewahrscheinlichkeit signifikant höher ist, wenn der bzw. die Proband*in mit dem/der Versuchsleiter*in der Untersuchung befreundet ist ($z = -9.402$, $p < .001$, $n = 162$, Effektstärke nach Cohen: $r = .74$). Versuchspersonen erscheinen mit höherer Wahrscheinlichkeit zum Untersuchungstermin, wenn sie mit dem bzw. der Versuchsleiter*in befreundet sind, als wenn sie diese bzw. diesen nicht kennen.

Ergänzende Fragestellung 6.4.: Hat das Geschlecht der Versuchsperson Einfluss auf die Ausfallquote?

Vier der fünf Personen, die bei der Untersuchung an der Universität ausgefallen sind, waren weiblich. Da die fünfte ausgefallene Person, den Fragebogen nicht nachträglich ausgefüllt hat, konnte für sie kein Geschlecht erhoben werden. Insgesamt meldeten sich für die Untersuchung an der Universität 18 Frauen und acht Männer an.¹⁵ Durch die kleine Gruppengröße ist ein sinnvoller statistischer Rückschluss nicht möglich.

In der Gesamtstichprobe zeigt sich kein Unterschied in der Verteilung der Geschlechter zwischen den Personen, die bei mindestens einer Untersuchung ausgefallen sind und jenen, die noch nie ausgefallen sind ($\chi^2_{1, n=183} = .042$, n.s.). Auch bei einer Begrenzung der Stichprobe auf

¹⁵ Nachdem eine der ausgefallenen Versuchspersonen nicht erreichbar war, konnte nicht festgestellt werden, welchem Geschlecht sich diese Person zuordnet.

Personen mit mindestens vier Versuchsteilnahmen, gibt es keinen Unterschied zwischen Männern und Frauen ($\chi^2_{1, n=91} = .089$, n.s.). Es gibt keinen Zusammenhang zwischen dem Ausfall von Versuchspersonen damit, ob diese männlich oder weiblich sind.

Ergänzende Fragestellung 6.5.: Hat die Uhrzeit des Untersuchungstermins Einfluss auf die Ausfallquote?

Für die Untersuchung an der Universität wählte jeweils nur eine Person einen Termin um 8:30 Uhr bzw. um 9:00 Uhr. Dies waren die beiden Termine, die am frühesten stattfinden sollten. Zu beiden Terminen fiel die jeweilige Versuchsperson aus. Je eine weitere Person fiel um 11:00 bzw. 17:00 Uhr aus. Weitere Untersuchungstermine wurden zwischen 9:30 und 18:30 Uhr ausgewählt und wahrgenommen. Am meisten Personen, nämlich sieben, meldeten sich für 11:00 Uhr an.

Mehr als drei Viertel (143 von 180) der Versuchspersonen gab an, dass die Teilnahmewahrscheinlichkeit gleich bleibt oder sinkt, wenn die Untersuchung vor 8:30 Uhr startet ($n_{\text{bleibt gleich}} = 75$; $n_{\text{sinkt ein wenig}} = 45$, $n_{\text{sinkt stark}} = 23$; Median = 3, Modus = 3). Außerdem gab mehr als drei Viertel (149 von 180) an, dass die Wahrscheinlichkeit gleich bleibt oder sinkt, wenn die Untersuchung nach 18:30 startet ($n_{\text{bleibt gleich}} = 88$; $n_{\text{sinkt ein wenig}} = 45$, $n_{\text{sinkt stark}} = 16$; Median = 3, Modus = 3) und mehr als zwei Drittel (127 von 179), dass die Wahrscheinlichkeit gleich bleibt, wenn die Untersuchung zwischen 9:00 Uhr und 17:00 Uhr startet ($n_{\text{gesamt}} = 179$, $n_{\text{bleibt gleich}} = 127$; Median = 3, Modus = 3). Eine Friedman's Rang-Varianzanalyse ergab, dass sich die zentrale Tendenz je nach Uhrzeit zu der Termin stattfindet unterscheidet ($\chi^2_{2, n=141} = 40.826$, $p < .001$). Ein paarweiser Vergleich, der als Post-Hoc-Test inklusive Bonferroni-Korrektur durchgeführt wurde, ergab, dass die Teilnahmewahrscheinlichkeit höher eingeschätzt wird, wenn eine Untersuchung zwischen 9:00 und 17:00 Uhr beginnt, als wenn sie vor 8:30 Uhr ($z = -4.556$, $p_{\text{angepasst}} < .001$, Effektstärke nach Cohen: $r = .38$) oder nach 18:30 Uhr startet ($z = 3.662$, $p_{\text{angepasst}} < .01$, Effektstärke nach Cohen: $r = .31$). Es gibt keinen Unterschied in der zentralen Tendenz der Teilnahmewahrscheinlichkeit, wenn die Untersuchung vor 8:30 oder nach 18:30 startet ($z = -0.893$, n.s.). Bei Untersuchungen die vor 8:30 oder nach 18:30 starten ist die Teilnahmewahrscheinlichkeit der Versuchspersonen geringer als bei Untersuchungen die zwischen 9:00 und 17:00 beginnen.

Kontrollitems

Mehr als die Hälfte (93 von 179) der Teilnehmer*innen in der Gesamtstichprobe gab an, dass die Wahrscheinlichkeit einer Untersuchungsteilnahme stark sinkt, wenn sie sich nicht an den Termin erinnern ($n_{\text{gesamt}} = 179$, $n_{\text{sinkt stark}} = 93$; Median = 1, Modus = 1). Von den Personen, die über das Online-Panel SurveyCircle rekrutiert wurden, gaben drei Viertel (63 von 84) an, dass die Wahrscheinlichkeit ein wenig oder stark sinkt, wenn sie sich nicht an den Termin erinnern ($n_{\text{gesamt}} = 84$, $n_{\text{sinkt ein wenig}} = 23$, $n_{\text{sinkt stark}} = 40$; Median = 2, Modus = 1).

In der Gesamtstichprobe sind mehr Psychologiestudierende und Psycholog*innen schon mindestens einmal bei Untersuchungen ausgefallen als Personen, die nicht Psychologie studierten ($\chi^2_{1, n=91} = 5.075$, $p < .05$). Psycholog*innen und Psychologiestudierende haben im Schnitt an deutlich mehr Untersuchungen teilgenommen hatten ($\bar{x} = 4.22$; $sd = 12.44$), als der Rest der Stichprobe ($\bar{x} = 0.50$; $sd = 3.08$; $t_{70.663} = -2.412$, $p < .05$), worauf der Unterschied beim Ausfall bei Untersuchungen zurückzuführen sein könnte. Begrenzt man die Stichprobe auf Personen mit mindestens vier Untersuchungsteilnahmen, so sind Psycholog*innen und Psychologiestudierende gleich häufig zumindest einmal bei einer Untersuchung ausgefallen wie andere Personen ($\chi^2_{1, n=92} = .021$, n.s.). Psycholog*innen und Psychologiestudierende fallen bei wissenschaftlichen Untersuchungen gleich häufig aus wie andere Personen.

Gründe für den Ausfall von Versuchspersonen

Von den 14 Personen, die angaben, bereits mindestens einmal bei einer wissenschaftlichen Untersuchung als Versuchsperson ausgefallen zu sein (inklusive der zwei von der statistischen Analyse ausgeschlossenen), haben acht als Grund dafür genannt, den Termin vergessen zu haben. Als weitere Gründe wurde zweimal „verschlafen“ sowie je einmal „Faulheit“, „kurzfristige Planänderung“, „Desinteresse“ und „Rausch“ genannt. All diese Gründe beschreiben unzuverlässiges Verhalten. Zusätzlich wurden in vier Fällen medizinische Gründe für den Ausfall genannt.

Diskussion

Zuverlässiges Verhalten bei Untersuchungsteilnahmen zeigt sich als komplexes Konstrukt, das von mehreren Faktoren positiv oder negativ beeinflusst werden kann, das aber auch durch viele Faktoren und Entscheidungen unbeeinflusst bleibt. Vor allem Erinnerungen an den Untersuchungstermin, eine freie Wahl des Termins, eine Form von direktem Kontakt mit den potentiellen Teilnehmer*innen, ein Näheverhältnis zu Teilnehmer*innen, ein für viele Personen interessantes Thema, eine Untersuchungsdauer von unter 30 Minuten und monetäre Anreize helfen, den Ausfall von Versuchspersonen zu reduzieren. Auf der anderen Seite führen eine höhere Geschäftigkeit der Versuchspersonen, uninteressante Themen, eine Untersuchungsdauer von mehr als 90 Minuten, Untersuchungstermine vor 8:30 Uhr bzw. nach 18:30 Uhr und Untersuchungsdesigns, die mehrere Terminen vorsehen, dazu, dass mehrere Personen mit höherer Wahrscheinlichkeit ausfallen. Auch wenn auf ein paar dieser Faktoren, wie die Geschäftigkeit der Versuchsteilnehmer*innen oder das notwendige Untersuchungsdesign, mitunter kein oder nur kaum Einfluss genommen werden kann, haben Versuchsleiter*innen dennoch gute Möglichkeiten, die Wahrscheinlichkeit, dass die Versuchspersonen auch tatsächlich an ihren Untersuchungen teilnehmen, zu beeinflussen.

Um festzustellen, welche dieser Faktoren den größten positiven bzw. negativen Einfluss auf die Teilnahmewahrscheinlichkeit haben, wurden zwei Friedman's Rang-Varianzanalysen durchgeführt. Einmal für alle Items bei denen die Versuchspersonen im Median angegeben haben, dass die Wahrscheinlichkeit, dass sie tatsächlich zum Termin erscheinen, ein wenig oder stark sinkt („Wenn mein/e Chef*in mich [kurz] vor dem Untersuchungstermin bittet bei der Arbeit einzuspringen...“, Median = 1; „Wenn ein Familienmitglied mich vor dem Untersuchungstermin um Hilfe bittet...“, Median = 2; „Wenn ein/e Freund*in mich vor dem Untersuchungstermin um Hilfe bittet...“, Median = 2; „Wenn die Untersuchung aus mehreren Terminen besteht...“, Median = 2; „Wenn ich das Thema der Untersuchung nicht interessant finde...“, Median = 2; „sinkt die Wahrscheinlichkeit ein wenig“ = 2, „sinkt die Wahrscheinlichkeit stark“ = 1) und einmal für alle Items bei denen die Versuchspersonen im Median angegeben haben, dass die Wahrscheinlichkeit, dass sie tatsächlich zum Termin erscheinen, ein wenig oder stark steigt („Wenn ich 5€ für die Teilnahme erhalte...“, Median = 4; „Wenn ich 10€/Stunde für die Teilnahme erhalte...“, Median = 4; „Wenn ich 30€ für die Teilnahme an der Untersuchung erhalte...“, Median = 5; „Wenn ich am Vortag eine Erinnerung an den Termin erhalte...“, Median = 4; „Wenn die Untersuchung von einem/r Freund*in geleitet wird...“, Median = 5; „Wenn ich den/die Versuchsleiter*in vor der Untersuchung bereits persönlich getroffen und mit ihm/ihr

gesprächen habe...“, Median = 4; „Wenn ich vor der Untersuchung bereits per Mail oder SMS mit dem/der Versuchsleiter*in kommuniziert habe...“, Median = 4; „Wenn ich den Untersuchungstermin frei wählen kann...“, Median = 4; „Wenn ich das Thema der Untersuchung besonders interessant finde...“, Median = 4; „steigt die Wahrscheinlichkeit ein wenig“ = 4, „steigt die Wahrscheinlichkeit stark“ = 5). Diese ergaben, dass es sowohl bei den negativen Einflussfaktoren ($\chi^2_{8, n=153} = 214.899, p < .001$) als auch bei den positiven Einflussfaktoren einen Unterschied bezüglich der Teilnahmewahrscheinlichkeit gibt ($\chi^2_{4, n=134} = 76.549, p < .001$).

Bei den Faktoren, die einen negativen Einfluss auf die Ausfallquote haben, zeigte sich durch Post-Hoc-Tests, dass von dem/der Chef*in benötigte Hilfe, die Teilnahmewahrscheinlichkeit stärker senkt als geringes Interesse, mehrere Untersuchungstermine oder von Freund*innen benötigte Hilfe (siehe Tabelle 10 auf S.85). Auch wenn ein Familienmitglied der Versuchsperson vor dem Untersuchungstermin spontan Hilfe benötigt sinkt die Teilnahmewahrscheinlichkeit stärker als bei geringem Interesse am Thema der Untersuchung, wenn die Untersuchung aus mehreren Untersuchungsterminen besteht oder als wenn von Freund*innen spontan Hilfe benötigt wird (siehe Tabelle 10 auf S.85).

Die positiven Einflussfaktoren zeigten durch Post-Hoc Tests ein komplizierteres Verhältnis zueinander (siehe Tabelle 11 auf S.86). Als stärkster Einfluss auf eine hohe Teilnahmewahrscheinlichkeit erwies es sich, wenn ein/e Freund*in die Untersuchung leitet. Diese Bedingung unterschied sich von allen anderen, mit Ausnahme davon, wenn 30€ für eine Untersuchung geboten werden. Wenn 30€ geboten werden, steigt die Wahrscheinlichkeit eher als bei einer freien Terminwahl, dem Senden von Erinnerungen, persönlichem Kontakt, Kontakt via Mail oder SMS bzw. als wenn 5€ als Anreiz für die Teilnahme geboten werden. Besonders interessante Themen steigern die Teilnahmewahrscheinlichkeit eher als persönlicher Kontakt oder Kontakt via Mail oder SMS und 5€ als Anreiz für die Teilnahme. Im Gegensatz dazu berichteten die Versuchspersonen, dass bei 5€ als Anreiz die Teilnahmewahrscheinlichkeit seltener steigt als bei einer freien Terminwahl, 10€ pro Stunde als Anreiz, besonderem Interesse, 30€ als Anreiz oder wenn ein oder eine Freund*in die Untersuchung leitet.

Unterschiede zwischen Bedingungen welche die Teilnahmewahrscheinlichkeit negativ beeinflussen

<i>Bedingung unterscheidet sich</i>	<i>von Bedingung</i>	<i>z</i>	<i>p</i>	<i>r</i>
<i>Wenn ein Familienmitglied mich vor dem Untersuchungstermin um Hilfe bittet...</i>	<i>Wenn ein/e Freund*in mich vor dem Untersuchungstermin um Hilfe bittet...</i>	3.59	<.01	0.29
<i>Wenn ein Familienmitglied mich vor dem Untersuchungstermin um Hilfe bittet...</i>	<i>Wenn die Untersuchung aus mehreren Terminen besteht...</i>	4.42	<.001	0.36
<i>Wenn ein Familienmitglied mich vor dem Untersuchungstermin um Hilfe bittet...</i>	<i>Wenn ich das Thema der Untersuchung nicht interessant finde...</i>	5.58	<.001	0.45
<i>Wenn mein/e Chef*in mich (kurz) vor dem Untersuchungstermin bittet bei der Arbeit einzuspringen...</i>	<i>Wenn ein/e Freund*in mich vor dem Untersuchungstermin um Hilfe bittet...</i>	3.48	<.01	0.28
<i>Wenn mein/e Chef*in mich (kurz) vor dem Untersuchungstermin bittet bei der Arbeit einzuspringen...</i>	<i>Wenn die Untersuchung aus mehreren Terminen besteht...</i>	-4.31	<.001	0.35
<i>Wenn mein/e Chef*in mich (kurz) vor dem Untersuchungstermin bittet bei der Arbeit einzuspringen...</i>	<i>Wenn ich das Thema der Untersuchung nicht interessant finde...</i>	5.47	<.001	0.44

Tabelle 10: Bedingungen die sich bezüglich der Teilnahmewahrscheinlichkeit unterscheiden. Berichtet werden die standardisierte Teststatistik z, mit einer Bonferroni-Korrektur angepasste p-Werte und die Effektstärke r einer Friedman's Rang-Varianzanalyse.

Unterschiede zwischen Bedingungen welche die Teilnahmewahrscheinlichkeit positiv beeinflussen

<i>Bedingung unterscheidet sich</i>	<i>von Bedingung</i>	<i>z</i>	<i>p</i>	<i>r</i>
<i>Wenn ich 5€ für die Teilnahme erhalte...</i>	<i>Wenn ich den Untersuchungstermin frei wählen kann...</i>	-3.55	<.05	0.31
<i>Wenn ich 5€ für die Teilnahme erhalte...</i>	<i>Wenn ich 10€/Stunde für die Teilnahme erhalte...</i>	-5.13	<.001	0.44
<i>Wenn ich 5€ für die Teilnahme erhalte...</i>	<i>Wenn ich das Thema der Untersuchung besonders interessant finde...</i>	-5.59	<.001	0.48
<i>Wenn ich 5€ für die Teilnahme erhalte...</i>	<i>Wenn ich 30€ für die Teilnahme an der Untersuchung erhalte...</i>	-7.27	<.001	0.63
<i>Wenn ich 5€ für die Teilnahme erhalte...</i>	<i>Wenn die Untersuchung von einem/r Freund*in geleitet wird...</i>	-10.17	<.001	0.88
<i>Wenn ich den/die Versuchsleiter*in vor der Untersuchung bereits persönlich getroffen und mit ihm/ihr gesprochen habe...</i>	<i>Wenn ich das Thema der Untersuchung besonders interessant finde...</i>	-3.29	<.05	0.28
<i>Wenn ich den/die Versuchsleiter*in vor der Untersuchung bereits persönlich getroffen und mit ihm/ihr gesprochen habe...</i>	<i>Wenn ich 30€ für die Teilnahme an der Untersuchung erhalte...</i>	-4.97	<.001	0.43
<i>Wenn ich den/die Versuchsleiter*in vor der Untersuchung bereits persönlich getroffen und mit ihm/ihr gesprochen habe...</i>	<i>Wenn die Untersuchung von einem/r Freund*in geleitet wird...</i>	-7.86	<.001	0.68
<i>Wenn ich vor der Untersuchung bereits per Mail oder SMS mit dem/der Versuchsleiter*in kommuniziert habe...</i>	<i>Wenn ich das Thema der Untersuchung besonders interessant finde...</i>	-3.20	<.05	0.28
<i>Wenn ich vor der Untersuchung bereits per Mail oder SMS mit dem/der Versuchsleiter*in kommuniziert habe...</i>	<i>Wenn ich 30€ für die Teilnahme an der Untersuchung erhalte...</i>	-4.89	<.001	0.42
<i>Wenn ich vor der Untersuchung bereits per Mail oder SMS mit dem/der Versuchsleiter*in kommuniziert habe...</i>	<i>Wenn die Untersuchung von einem/r Freund*in geleitet wird...</i>	-7.78	<.001	0.67

<i>Wenn ich am Vortag eine Erinnerung an den Termin erhalte...</i>	<i>Wenn ich 30€ für die Teilnahme an der Untersuchung erhalte...</i>	-4.42	<.001	0.38
<i>Wenn ich am Vortag eine Erinnerung an den Termin erhalte...</i>	<i>Wenn die Untersuchung von einem/r Freund*in geleitet wird...</i>	-7.31	<.001	0.63
<i>Wenn ich den Untersuchungstermin frei wählen kann...</i>	<i>Wenn ich 30€ für die Teilnahme an der Untersuchung erhalte...</i>	-3.73	<.01	0.32
<i>Wenn ich den Untersuchungstermin frei wählen kann...</i>	<i>Wenn die Untersuchung von einem/r Freund*in geleitet wird...</i>	-6.62	<.001	0.57
<i>Wenn ich 10€/Stunde für die Teilnahme erhalte...</i>	<i>Wenn die Untersuchung von einem/r Freund*in geleitet wird...</i>	-5.04	<.001	0.44
<i>Wenn ich das Thema der Untersuchung besonders interessant finde...</i>	<i>Wenn die Untersuchung von einem/r Freund*in geleitet wird...</i>	-4.57	<.001	0.39

Tabelle 11: Bedingungen die sich bezüglich der Teilnahmewahrscheinlichkeit unterscheiden. Berichtet werden die standardisierte Teststatistik z, mit einer Bonferroni-Korrektur angepasste p-Werte und die Effektstärke r einer Friedman's Rang-Varianzanalyse.

Diese Ergebnisse nehmen allerdings nur Bezug auf die Teilnahmewahrscheinlichkeit. Die Auswirkungen der untersuchten Einflussfaktoren sind jedoch auch immer in Anbetracht der allgemeinen Effizienz einer Untersuchung differenziert zu hinterfragen. Beispielsweise ist trotz einer Steigerung der Teilnahmewahrscheinlichkeit und der damit verbundenen Senkung der Ausfallquote festzuhalten, dass das Auszahlen monetärer Anreize aufgrund der entstehenden Kosten nicht notwendigerweise zu einer Steigerung der Effizienz der Untersuchung führt. Es gilt abzuwägen, ob die Ausweitung der Rekrutierungsmaßnahmen und die Optimierung der Ausschöpfungsquote nicht ressourcenschonender sind, um die gewünschte Stichprobengröße durch die Reduzierung der Ausfallquote zu erreichen, als die Inanspruchnahme zusätzlicher finanzieller Ressourcen.

Es ist also, wie in der Einleitung erwähnt, zu beachten, dass für effiziente Forschung nicht allein Ausfallquote bzw. Teilnahmewahrscheinlichkeit entscheidend sind. Es gilt, ein gutes Mittelmaß zwischen der Ausschöpfungsquote und der Ausfallquote zu finden. Bei der Umsetzung eines Untersuchungsplans, der zwar die Ausfallquote sehr gering hält, aber dafür auch nur eine geringe Ausschöpfungsquote aufweist, kann das Erreichen der gewünschten Stichprobengröße ebenso zum Problem werden wie ein Plan, der zwar die Ausschöpfungsquote optimiert, dafür aber mit einer hohen Ausfallquote zu kämpfen hat. Insofern ist eine effektive Rekrutierung

ebenso wichtig, wie dabei den Grundstein für zuverlässiges Verhalten zu legen. Welche Maßnahmen zur Optimierung der Ausschöpfungsquote gesetzt werden können, erklärt Keusch (2015) ausführlich am Beispiel von Online-Umfragen.

Die Ausschöpfungsquote kann aber auch durch die Umstände von Untersuchungen beeinflusst werden. Vor allem Personen, die sonst möglicherweise ein erhöhtes Ausfallrisiko zeigen würden, melden sich unter Bedingungen, die sie vermeiden möchten, oft einfach nicht zu Untersuchungen an. Diese Bedingungen wirken sich somit nicht auf die Ausfallquote, sondern nur auf die Ausschöpfungsquote aus. In der Gesamtstichprobe wurde im Zuge der Erhebung der Teilnahmewahrscheinlichkeiten auch erhoben, unter welchen Bedingungen sich Versuchspersonen nicht zu einer Untersuchung anmelden würden. Unter zwei Bedingungen gaben mehr als 10% der Versuchspersonen an, sich nicht zu einer solchen Untersuchung anzumelden. Am meisten Personen, nämlich 27, würden sich nicht zu einer Untersuchung anmelden, die vor 8:30 Uhr beginnt. 20 Personen würden sich nicht zu einer Untersuchung anmelden, deren Thema sie nicht interessiert. Bei sieben weiteren Bedingungen gaben mehr als 5% der Versuchspersonen an, sich nicht zu einer solchen Untersuchung anzumelden. So ist das bei 17 der Fall, wenn die Untersuchung von einer oder einem Freund*in geleitet wird und bei 14 Personen, wenn die Untersuchung nach 18:30 Uhr startet. Je 13 Personen würden sich nicht anmelden, wenn die Untersuchung aus mehreren Terminen besteht, mehr als eineinhalb Stunden dauert oder ein/e Freund*in sie vor der Untersuchung um Hilfe bittet. Elf Personen schließen eine Anmeldung aus, wenn technisch aufwendige Verfahren eingesetzt werden und zehn, wenn ein Familienmitglied am Tag der Untersuchung Hilfe benötigt. Auch wenn nur ein Teil dieser Umstände beeinflusst werden kann, ist unter diesen Bedingungen doch mit einer niedrigeren Ausschöpfungsquote zu rechnen.

Die Ausschöpfungsquote bei Laboruntersuchungen kann nicht nur durch das Vermeiden gewisser Umstände verbessert werden, sondern auch durch die Reduktion der Abmeldungen. Selbst wenn Personen, die sich von Untersuchungen wieder abmelden, dabei meistens kaum bis keine Ressourcen in Anspruch nehmen, kann so die Effizienz gesteigert werden. In der Gesamtstichprobe gaben 26 Personen (14.13%) an, sich von einer Untersuchung schon einmal wieder abgemeldet zu haben (Maximum = 9 Abmeldungen, $\bar{x} = 0.39$; $sd = 1.19$). Als Grund für die Abmeldung wurde in 19 Fällen angegeben, dass eine Teilnahme aufgrund anderer Termine oder Stress zeitlich doch nicht möglich war. In zwölf Fällen waren gesundheitliche Ursachen für eine Abmeldung verantwortlich, in je zwei Fällen die Gestaltung der Untersuchung bzw.

unzuverlässiges Verhalten sowie in einem Fall eine Verspätung durch die öffentlichen Verkehrsmittel.

Um zu überprüfen, ob es sich beim Ausfall von Versuchspersonen und bei der Abmeldung von bereits vereinbarten Untersuchungsterminen um ähnliche Konstrukte handelt, wurden diverse Korrelationen mit der Häufigkeit, mit der Versuchspersonen sich von wissenschaftlichen Untersuchungen wieder abgemeldet hatten, berechnet (siehe Tabelle 12). Zwischen der Häufigkeit der Abmeldungen und der Häufigkeit der Ausfälle zeigte sich in der Gesamtstichprobe ein mittlerer Zusammenhang ($r = .493, p < .001$). Je häufiger Personen sich von Untersuchungen wieder abgemeldet haben, umso häufiger sind sie auch schon bei Untersuchungen ausgefallen. Eine Kontrolle dieses Zusammenhangs durch die Anzahl der Untersuchungsteilnahmen verändert diesen Zusammenhang kaum ($r = .463, p < .001$). Außerdem zeigt sich, dass je häufiger sich Personen von Untersuchungen wieder abgemeldet haben, sie umso seltener die berechtigten Erwartungen anderer erfüllen, ($r = -.164, p < .05$) und dass sie eher zu Neurotizismus neigen ($r = .188, p < .05$). Bei einer Begrenzung der Stichprobe auf Personen, die bereits bei mindestens vier Untersuchungen teilgenommen hatten, veränderten sich die Korrelationen mit der Häufigkeit der Abmeldungen nur geringfügig (Häufigkeit der Ausfälle: $r = .467, p < .001$; berechnete Erwartungen anderer: $r = -.240, p < .05$; Neurotizismus: $r = .216, p < .05$). Bei Personen, die bereits an mindestens vier Untersuchungen teilgenommen hatten, zeigte sich außerdem, dass je höher die Anzahl ihrer Abmeldungen ist, Versuchspersonen es umso niedriger einschätzten, dass andere sich auf sie verlassen können ($r = -.259, p < .05$). Auch wenn es einige Personen gibt, die bei Untersuchungen sowohl ausgefallen sind als sich auch davon wieder abgemeldet haben, zeigen sich doch unterschiedliche Zusammenhangsstrukturen. Somit sollten beide Verhaltensweisen getrennt betrachtet werden. Abmeldungen von der Untersuchung müssten bei der Planung der Untersuchung unabhängig vom Ausfall vorgebeugt werden, soweit das überhaupt möglich ist. Auf diesem Gebiet ist jedoch eigenständige Forschung notwendig.

Korrelationen mit der Häufigkeit, mit der Versuchspersonen sich von wissenschaftlichen Untersuchungen wieder abgemeldet hatten

	Häufigkeit der Abmeldungen	Häufigkeit der Abmeldungen (nur Personen mit mindestens vier Teilnahmen bei wissenschaftlichen Untersuchungen)
Häufigkeit der Ausfälle	$r = .493^{**}$ $p < .001$ $n = 184$	$r = .467^{**}$ $p < .001$ $n = 92$
Soziale Verantwortung	$r = -.014$ $p = .846$ $n = 184$	$r = -.011$ $p = .915$ $n = 92$
Erfüllung berechtigter Erwartungen anderer	$r = -.164^*$ $p < .05$ $n = 184$	$r = -.240^*$ $p < .05$ $n = 92$
Befolgung sozialer Spielregeln	$r = .073$ $p = .327$ $n = 184$	$r = .134$ $p = .203$ $n = 92$
Andere können sich auf mich verlassen***	$r = -.144$ $p = .051$ $n = 184$	$r = -.259^*$ $p < .05$ $n = 92$
Wahrscheinlichkeit einen Untersuchungstermin zu verschlafen, vergessen oder zu übersehen***	$r = .012$ $p = .872$ $n = 180$	$r = -.016$ $p = .881$ $n = 90$
Neurotizismus	$r = .188^*$ $p < .05$ $n = 184$	$r = .216^*$ $p < .05$ $n = 92$
Extraversion	$r = -.011$ $p = .884$ $n = 184$	$r = .016$ $p = .877$ $n = 92$
Offenheit	$r = -.040$ $p = .590$ $n = 184$	$r = -.099$ $p = .346$ $n = 92$
Gewissenhaftigkeit	$r = -.007$ $p = .923$ $n = 184$	$r = .026$ $p = .807$ $n = 92$
Verträglichkeit	$r = -.093$ $p = .210$ $n = 184$	$r = -.130$ $p = .215$ $n = 92$

Tabelle 12: Die Korrelationskoeffizienten zwischen der Häufigkeit, mit der sich Versuchspersonen von wissenschaftlichen Untersuchungen wieder abgemeldet hatten für die Gesamtstichprobe und Personen mit mindestens vier Untersuchungsteilnahmen.

Korrelationen wurden berechnet für die Häufigkeit der Abmeldungen (zwischen 0 und 5 Abmeldungen) mit der Häufigkeit mit der Versuchspersonen bereits bei wissenschaftlichen Untersuchungen

ausgefallen sind (zwischen 0 und 6 Ausfällen), dem Punktwert auf der Skala der Sozialen Verantwortung (zwischen 44 und 77 Punkten), der zugehörigen Subskalen Erfüllung berechtigter Erwartungen anderer (zwischen 14 und 31 Punkten) und Befolgung sozialer Spielregeln (zwischen 9 und 33 Punkten), dem Item „Ich gehöre zu der Art von Menschen, auf die andere sich verlassen können“ (zwischen „eher falsch“ und „sehr richtig“), der Angabe für wie wahrscheinlich die Versuchsperson es hält, dass sie einen Untersuchungstermin verschläft, vergisst oder aus anderen Gründen übersieht (zwischen 1 und 4; 1 = nicht wahrscheinlich, 5 = sehr wahrscheinlich) sowie mit den erzielten Punkten auf den Subskalen Neurotizismus (zwischen 12 und 40 Punkten), Extraversion (zwischen 16 und 38 Punkten), Offenheit (zwischen 19 und 37 Punkten), Gewissenhaftigkeit (zwischen 17 und 36 Punkten) und Verträglichkeit (zwischen 19 und 40 Punkten) des B5T.

* $p < .05$, ** $p < .01$, ***Spearman-Korrelationskoeffizienten

Einige der untersuchten Einflussfaktoren zeigten keinen oder nur wenig Einfluss auf die Ausfallquote bzw. die Teilnahmewahrscheinlichkeit. Die Identifizierung solcher Faktoren ist insofern von Bedeutung, da es erlaubt, während der Planung einer Untersuchung den Fokus auf jene Faktoren zu legen, die wirklich von Bedeutung sind. Sauermann und Roach (2013) schreiben nach dem erfolglosen Versuch die Ausschöpfungsquote bei Online-Umfragen zu verbessern, passend dazu, dass alle Forscher*innen Entscheidungen bezüglich solcher Faktoren treffen müssen und ihre Ergebnisse ein Hinweis darauf seien, dass diese ihre Zeit und Anstrengungen lieber für die Optimierung anderer Design-Parameter aufwenden sollten. Bezüglich der Parameter, die Versuchsleiter*innen bei den Untersuchungen beeinflussen können, bedeutet das, dass sie sich wegen der Ausfallquote keine Gedanken darüber machen müssen, welche Untersuchungsmethoden sie einsetzen, ob sie eine Voruntersuchung zur Selektion der Teilnehmer*innen einsetzen, ob die Rekrutierung einen Monat oder direkt vor der Testung beginnt, ob zum Untersuchungszeitpunkt mit schlechtem Wetter zu rechnen ist oder ob sie Versuchspersonen mit „Du“ oder „Sie“ anschreiben oder ansprechen.

Das Item, das die Zuverlässigkeit bei Untersuchungen direkt erhob („Für wie wahrscheinlich halten Sie es, dass Sie einen Versuchstermin verschlafen, vergessen oder aus anderen Gründen übersehen?“) wurde extra für diese Untersuchung entworfen. Es zeigte sowohl mit dem Ausfall von Versuchspersonen als auch mit der Häufigkeit des Ausfalls einen Zusammenhang. Um den Kontext dieses Items besser zu verstehen, wurde versucht, seine inhaltliche Validität anhand der anderen erhobenen Variablen zur Verlässlichkeit und Zuverlässigkeit zu bestimmen. Am stärksten korrelierte die Zuverlässigkeit bei Untersuchungen mit dem Item der SV-Skala, das abfragte, wie sehr Personen von anderen als verlässlich eingeschätzt werden ($r = -.321$, $p < .001$) sowie der Gewissenhaftigkeit ($r = -.314$, $p < .001$), wobei die Zuverlässigkeit oft als Subfacette dieser betrachtet wird (Dudley, Orvis, Lebiecki & Cortina, 2006; Hough, 1992; Hough & Ones, 2001; Olaru et al, 2019; Roberts et al, 2005; Stewart, 1999). Zusätzlich zeigten

sich kleinere Zusammenhänge mit der sozialen Verantwortung ($r = -.281, p < .001$) und der Erfüllung berechtigter Erwartungen anderer ($r = -.197, p < .001$), nicht aber mit der Befolgung sozialer Spielregeln ($r = -.123, n.s.$). Diese durchwegs mittleren bis kleinen Korrelationen lassen vermuten, dass das Item zur Zuverlässigkeit bei Untersuchungen zwar ein verwandtes, aber doch eigenständiges Konstrukt erfasst. Der wohl bedeutendste Unterschied zu den anderen erhobenen Variablen besteht darin, dass nicht versucht wird, ein Persönlichkeitsmerkmal zu erfassen. Stattdessen wurden durch die Formulierung: „Für wie wahrscheinlich halten Sie es, dass Sie einen Versuchstermin verschlafen, vergessen oder aus anderen Gründen übersehen?“ spezifische unzuverlässige Verhaltensweisen direkt abgefragt, die von mehreren Faktoren als nur einem bestimmten Persönlichkeitsmerkmal abhängen könnten.

Eine Besonderheit des Ausfalls von Versuchspersonen stellen auch die Unterschiede zwischen verschiedenen Subpopulationen dar. Auch wenn bei dieser Studie nicht explizit nach Unterschieden zwischen verschiedenen Personengruppen gesucht wurde, sind diese doch durch Unterschiede zwischen den verschiedenen Teilstichproben ersichtlich. Bei den erhobenen Daten wurde das durch eine Varianzanalyse bei der Häufigkeit der Abmeldungen ($F_{2,179} = 3.874, p < .05$) und bei der SV-Skala ($F_{2,179} = 5.921, p < .01$) sowie ihren beiden Subskalen, der Erfüllung berechtigter Erwartungen anderer ($F_{2,179} = 4.179, p < .05$) und der Befolgung sozialer Spielregeln ($F_{2,179} = 9.455, p < .001$) ersichtlich. Dennoch kamen die statistischen Analysen zu diesen Variablen bei allen Gruppen zu den gleichen Schlüssen. Außerdem zeigte eine Varianzanalyse Unterschiede zwischen den Teilstichproben bezüglich der durchschnittlichen Teilnahmewahrscheinlichkeit der Versuchspersonen unter allen Bedingungen ($F_{2,179} = 7.805, p < .01$). Die Gruppe an Personen, die an der Untersuchung an der Universität teilnahm, zeigte eine höhere durchschnittliche Teilnahmewahrscheinlichkeit ($\bar{x} = 3.23; sd = 0.18$) als Personen, die an der Online-Umfrage teilnahmen ($\bar{x} = 3.04; sd = 0.20$). Diese Unterschiede könnten auf die unter „Überprüfung der Homogenität der Stichprobe“ berichtete, unterschiedliche demografische Zusammensetzung der Gruppen zurückzuführen sein.

Diese unterschiedliche Ausprägung der durchschnittlichen Teilnahmewahrscheinlichkeit ist ein Hinweis darauf, dass die berichteten Einflussfaktoren zwar als Grundstruktur zum Ausfall von Versuchspersonen angesehen werden können, dass diese jedoch nicht unbedingt universell gelten müssen. Bei der Planung einer Untersuchung ist es wichtig, sich über die Bedürfnisse der Personen in der geplanten Stichprobe Gedanken zu machen, sich in diese hineinzusetzen

und den Rekrutierungsprozess, Anreize und Maßnahmen entsprechend vorzubereiten. Die Ergebnisse dieser Arbeit bilden nur den Grundstein der Überlegungen, die notwendig sind, um die Ausfallquote bestmöglich zu optimieren.

Im Vergleich zur durchschnittlich berichteten Ausfallquote von 10.66% in der Voruntersuchung war die Ausfallquote bei dem Teil der Untersuchung, der an der Universität durchgeführt wurde, höher (19.23%). Ein möglicher Grund dafür könnte der Verzicht auf die Rekrutierung von Freunden, Familie und Bekannten bei diesem Teil der Untersuchung sein, um die Ausfallquote nicht zu verfälschen. In Anbetracht dessen, dass Versuchspersonen angegeben haben, dass die Teilnahmewahrscheinlichkeit stark steigt, wenn eine Untersuchung von einem bzw. einer Freund*in geleitet wird, könnte der Verzicht darauf, diese Personengruppe zu rekrutieren, einen Teil dieses Unterschieds bedingt haben. Bei der Voruntersuchung gaben 15 der 22 Versuchsleiter*innen, deren Ausfallquote erfasst wurde, an, Verwandte, Freunde und Bekannte rekrutiert zu haben. Betrachtet man die Ausfallquoten der Voruntersuchung, so zeigt sich kein Unterschied zwischen Untersuchungen bei denen Verwandte, Freunde und Bekannte rekrutiert wurden und jenen, die darauf verzichtet haben ($U = 48.50$, $z = -.282$, n.s.). Eventuell stellt der Verzicht, diese Personengruppe zu rekrutieren, einen begünstigenden, aber nicht notwendigerweise auftretenden Einflussfaktor auf die Ausfallquote dar. Weitere Forschung ist hier noch notwendig.

Zu hinterfragen ist auch, ob die Ergebnisse einen „nonresponse bias“ abbilden, da weniger Personen, die berichteten bereits einmal bei einer Untersuchung ausgefallen zu sein, in der Gesamtstichprobe waren (6.52% der Teilnehmer*innen) als eine durchschnittliche Ausfallquote von 10.66% vermuten lassen würde. Allerdings berichtete ein beachtlicher Teil der Personen in der Gesamtstichprobe, dass sie noch nie an einer wissenschaftlichen Untersuchung an einer Universität teilgenommen haben (32.61%). Bedenkt man, dass Personen, die noch nie an einer solchen Untersuchung teilgenommen haben, auch nicht ausfallen können, ist der Anteil an Personen, die ausgefallen sind, weniger überraschend. Die Wahrscheinlichkeit eines Ausfalls von einzelnen Personen steigt mit einer zunehmenden Anzahl an Untersuchungsteilnahmen. Schließt man alle Personen aus, die noch nie teilgenommen haben, so macht der Anteil jener Personen, die bereits mindestens einmal ausgefallen sind, 9.68% aus. Bei mindestens zwei Teilnahmen sind es 10.91%, bei mindestens drei sind es 12% und bei mindestens vier Teilnahmen sind es 13.04%. Unter Anbetracht dieser Zahlen ist davon auszugehen, dass die erhobene Gesamtstichprobe einen repräsentativen Teil an ausgefallenen Personen abbildet.

Einschränkungen dieser Studie

Es gibt einige Einschränkungen, die die Ergebnisse dieser Studie mit sich bringen. Diese sind durch die Planung der Untersuchung, aber auch durch die Tatsache zustande gekommen, dass sich initial zu wenige Personen durch die gesetzten Rekrutierungsmaßnahmen zur ursprünglich geplanten Hauptuntersuchung angemeldet haben. Das betrifft zum Teil die Überprüfung einzelner Hypothesen aufgrund fehlender Daten, die Qualität der Daten und die Heterogenität der Gesamtstichprobe.

Zwei der Hypothesen dieser Arbeit waren aufgrund der zu kleinen Stichprobe bei der Untersuchung an der Universität nicht überprüfbar. Diese betrafen den Einfluss von aufwendigen Anreisen und von Untersuchungsscheinen auf die Ausfallquote. Durch die Verteilung der Anreizeiten lässt sich schließen, dass auch eine wenig aufwendige Anreise von zwei bis 15 Minuten zumindest kein Ausschlusskriterium für den Ausfall von Versuchspersonen ist. Ähnliches gilt für die Vergabe von Untersuchungsscheinen als Anreiz zur Teilnahme. Drei der vier Personen, die bei der Untersuchung an der Universität ausgefallen sind, haben angegeben, einen Untersuchungsschein zu benötigen. Somit liegt die Vermutung nahe, dass diese Art von Anreiz Versuchspersonen nicht davon abhält, bei Untersuchungen auszufallen.

Auch die Hypothese zu den Auswirkungen der verschiedenen Rekrutierungsmethoden konnte nur zum Teil überprüft werden. Durch die geringe Anzahl an Anmeldungen war ein Vergleich der Gruppen, die sich aufgrund verschiedener Rekrutierungsmaßnahmen angemeldet haben, nicht möglich.

Ein Großteil der Versuchspersonen gab an, dass die Teilnahmewahrscheinlichkeit steigt, wenn sie Geld für eine Untersuchung bekommen. In der Voruntersuchung wurde jedoch festgestellt, dass vor allem die Anzahl an unterschiedlichen Anreizen einen Einfluss auf die Ausfallquote hatte. Dies konnte in der Hauptuntersuchung nicht überprüft werden, da die gebotenen Anreize nicht variiert werden konnten. Auch wenn Süßigkeiten und Gewinnspiele per se keinen oder nur wenig Einfluss auf die Teilnahmewahrscheinlichkeit haben, könnten diese gemeinsam in einer ansprechend gestalteten Einladung doch einen positiven Einfluss haben. Außerdem wurde nur erhoben, ob die Möglichkeit an einem Gewinnspiel teilzunehmen die Teilnahmewahrscheinlichkeit erhöht, ohne auf den Wert des Gewinns einzugehen. Wie in der Einleitung erwähnt, waren die Auswirkungen bei Gewinnspielen in anderen Forschungsfeldern am stärksten, wenn wenige, dafür aber hohe Gewinne eingesetzt wurden (Milward, Lynskey & Strang, 2014; Sauermann & Roach, 2013). Bei einer Untersuchung, von der in der Voruntersuchung berichtet

wurde, schilderte der bzw. die Versuchsleiter*in einen besonders hohen Gewinn als Anreiz zur Teilnahme eingesetzt zu haben. Die Ausfallquote der Untersuchung lag mit 10,71% dennoch nahe am Durchschnitt von 10,66%. Ob sich der Einfluss von Gewinnspielen auf die Ausfallquote ändert, wenn hohe Gewinne eingesetzt werden, bleibt offen und muss weiter erforscht werden.

Zusätzlich bleiben einige mögliche Einflussfaktoren auf die Ausfallquote unerforscht, auf die bei der Durchführung der Untersuchung nicht entsprechend eingegangen werden konnte. Dazu gehören der Zeitpunkt der Aussendung, kulturell bedingte Unterschiede der Ausfallquote und die Teilnahme an Untersuchungen zur Unterhaltung oder um einen Beitrag zur Wissenschaft zu leisten, unabhängig vom Interesse am Thema. Um den Einfluss dieser Faktoren auf die Ausfallquote einschätzen zu können, ist weitere Forschung nötig.

Die Analyse des Kontrollitems zur Erfassung der Teilnahmewahrscheinlichkeit („Wenn ich mich nicht an den Termin erinnere...“) zeigte, dass ein Teil der Versuchspersonen dieses Item entweder nicht genau gelesen hat oder die Aufgabenstellung nicht allen Personen ganz klar war. Durch Rückmeldungen zur Untersuchung wurde deutlich, dass dieses Item aufgrund der Aufgabenstellung zu Verwirrung geführt hat, weshalb Personen die kontraintuitiv geantwortet haben („...steigt die Wahrscheinlichkeit stark“) nicht automatisch ausgeschlossen wurden. Dennoch wurden alle Datensätze, bei denen nicht mit „...sinkt die Wahrscheinlichkeit ein wenig“ oder „...sinkt die Wahrscheinlichkeit stark“ geantwortet wurde, genau auf Klickmuster oder unsachgemäße Bearbeitung der Fragen überprüft.

Die Bearbeitung dieses Items war scheinbar vor allem bei den von SurveyCircle rekrutierten Personen ein Problem (Median = 2, Modus = 1). Dafür gibt es zwei mögliche Erklärungsansätze. Erstens war die Aufgabenstellung zur Erhebung der Teilnahmewahrscheinlichkeit möglicherweise nicht klar genug und für alle verständlich formuliert. Sollte dieser Erhebungsbogen wiederverwendet werden, wird empfohlen, die Instruktion als Fallbeispiel aus Perspektive der Versuchsperson zu formulieren (Zum Beispiel: „Stellen Sie sich vor, Sie haben sich zu einer Untersuchung angemeldet. Bitte geben Sie an, ob die Wahrscheinlichkeit, dass Sie den Untersuchungstermin tatsächlich wahrnehmen, unter folgenden Bedingungen sinkt, steigt oder gleich bleibt.“). Zweitens könnte die Qualität der Daten von SurveyCircle aufgrund des zugrundeliegenden Systems geringer sein als die der anderen beiden Stichproben. SurveyCircle ist eine Plattform auf der Online-Umfragen in einer Rangliste präsentiert werden. Für jede absolvierte Untersuchung bekommen Teilnehmer*innen Punkte auf ihr Konto gutgeschrieben. Die Anzahl

der gutgeschriebenen Punkte richtet sich danach, welche Position eine Umfrage auf der Rangliste hat. Umfragen, die einen besseren Rang belegen, geben mehr Punkte. Welchen Rang eine Umfrage einnimmt, wird dadurch bestimmt, wie viele Punkte dem Account, über den die Umfrage aktiviert wurde, im Vergleich zu allen anderen gerade aktiven Umfragen bereits gutgeschrieben wurden (<https://www.surveycircle.com/de/>). Wurden mit einem Account mehr Punkte gesammelt, können also auch schneller mehr Versuchspersonen rekrutiert werden. Dementsprechend sind diese Punkte für viele Teilnehmer*innen die größte Motivation an anderen Untersuchungen teilzunehmen. Von den 105 Personen, die für diese Untersuchung über SurveyCircle rekrutiert wurden, gaben 73 als Teilnahmegrund die Punkte für das Online-Panel SurveyCircle an, wobei dies für 49 Personen sogar der einzige Teilnahmegrund war. Ohne Kontrollinstanz können auf SurveyCircle durch schnelles „Durchklicken“ durch möglichst viele Umfragen in kurzer Zeit viele Punkte gesammelt werden, wodurch allerdings die Qualität der Daten sinkt und einige Datensätze ganz auszuschließen sind.

Wie bereits in der Einleitung erläutert, könnten auch Untersuchungsscheine, ähnlich den Punkten bei SurveyCircle, dazu führen, dass die Qualität der Daten abnimmt (Koch, Peter & Müller, 2019, S.124), aber auch dazu, dass die Zahl der Anmeldungen sinkt. Je nach Länge der Untersuchung und dem damit einhergehenden Stundenausmaß der Untersuchungsscheine könnte der Nutzen, an einer kurzen Untersuchung teilzunehmen, den Aufwand nicht überwiegen, sodass Versuchspersonen sich tendenziell eher zu weniger, dafür aber längeren Untersuchungen anmelden. Außerdem warnen Porter, Whitcomb und Weitzer (2004) vor einem sogenannten „Survey Fatigue“, also davor, dass die Ausschöpfungsquote bei Untersuchungen sinkt, wenn versucht wurde, die Versuchspersonen aus der Stichprobe zuvor schon zur Teilnahme an einer oder mehreren anderen Untersuchungen zu bewegen. Durch die Masse an Einladungen zu Untersuchungen, die Studierende erhalten, dürfte so die Rekrutierung erschwert werden. Zusätzlich könnten Personen, die genug Untersuchungsscheine gesammelt haben, aufgrund des Gefühls ihren Beitrag geleistet zu haben, das Interesse verlieren an weiteren Untersuchungen teilzunehmen. Eine Kombination dieser Faktoren könnte auch erklären, warum sich so wenige Personen bei dieser Studie zur Teilnahme an der eher kurz gehaltenen Untersuchung an der Universität bereiterklärt haben.

Außerdem besteht durch die Korrelation zwischen der Häufigkeit des Ausfalls und der Zuverlässigkeit, bei Untersuchungen zu erscheinen, Grund zu der Annahme, dass einige Personen eine grundlegende Tendenz zum Ausfall haben. Die Wahrscheinlichkeit, dass diese Personen ausfallen, steigt somit mit der Anzahl an Untersuchungen, zu denen sie sich anmelden. Dies

könnte zu einer Unterschätzung der gefunden Effekte führen, weil Antworten von Personen, die diese Tendenz zum Ausfall aufweisen, aber zum Untersuchungszeitpunkt noch nie ausgefallen sind, die Effektstärken mindern. Nachdem es sich hier um Grundlagenforschung handelt, ist die Überprüfung der Ergebnisse in der Praxis somit der nächste Schritt. Dazu würde sich die Begleitung, Beobachtung und der Vergleich von wissenschaftlichen Untersuchungen, die an Universitäten ohnehin durchgeführt werden, anbieten.

Empfehlung für zukünftige Forschung

Effizienter Forschung liegen drei Bausteine zugrunde: eine hohe Ausschöpfungsquote, um die gewünschte Stichprobengröße erreichen zu können, eine geringe Ausfallquote, um den Forschungsprozess ressourcenschonend zu halten und eine hohe Qualität der erhobenen Daten. Dabei muss eine hohe Ausfallquote nicht unbedingt ein Problem darstellen. Sie kann durch das Erreichen einer guten Ausschöpfungsquote kompensiert werden, wenn die endgültige Stichprobengröße den Ansprüchen der Forscher*innen genügt. Die Optimierung der Ausschöpfungsquote und das Erreichen der notwendigen Stichprobengröße wurde in den letzten Jahren jedoch immer aufwendiger und schwerer (Tourangeau, 2017), was nicht zuletzt Phänomenen wie dem „Survey Fatigue“ geschuldet ist (Porter, Whitcomb und Weitzer, 2004). Durch diese Entwicklung kommt der Ausfallquote immer mehr Bedeutung zu, nicht nur um Ressourcen zu schonen, sondern auch, um effizient und schnell eine ausreichend große Stichprobe zu generieren.

Forschende können dem Ausfall von Versuchspersonen auf verschiedene Arten vorbeugen. Etwa dadurch, den Versuchspersonen Erinnerungen an den genauen Untersuchungstermin und Ort zu senden. Die Wahl des Untersuchungstermins sollte den Versuchsteilnehmer*innen zuvor so frei wie möglich stehen. Dabei sollten optimaler Weise auch Termine zur Verfügung stehen, die nicht vor 8:30 Uhr oder nach 18:30 stattfinden, da zu solchen Terminen die Ausfallquote erhöht ist. Ebenso sind kürzere Untersuchungen hilfreich, um die Ausfallquote gering zu halten und auch, wenn für manche Forschungsdesigns mehrere Termine notwendig sind, sollte dies in der Planung, wenn möglich, vermieden werden.

Ein direkter Austausch mit Versuchspersonen, egal ob persönlich, via E-Mail oder SMS, unterstützt die Teilnahmewahrscheinlichkeit und kann dazu genutzt werden, bei den Versuchspersonen Interesse am Thema der Untersuchung zu wecken. Das Interesse am Thema erhöht die Wahrscheinlichkeit, dass Versuchspersonen tatsächlich an der Untersuchung teilnehmen, während sie weniger wahrscheinlich zur Untersuchung erscheinen, wenn sie das Thema nicht interessant finden. Dementsprechend wichtig ist es, die Rekrutierungsmaßnahmen interessant zu gestalten und den Versuchsteilnehmer*innen Interesse am Thema zu vermitteln.

Monetäre Anreize können helfen, den Ausfall von Versuchspersonen zu reduzieren. Wenn Versuchspersonen mehr Geld für eine Teilnahme geboten wird, erscheinen diese wahrscheinlicher zur Untersuchung. Dennoch gilt im Einzelfall abzuwägen, ob das Ausmaß, in dem die Ausfallquote reduziert wird, die Mehrkosten, die durch den Einsatz finanzieller Ressourcen entstehen,

aufwiegen. Wie die Voruntersuchung zeigt, kann auch ohne den Einsatz monetärer Anreize eine niedrige Ausfallquote erreicht werden.

Unzuverlässigem Verhalten kann durch diese Maßnahmen also gut vorgebeugt werden, aufgrund von Zwischenfällen und persönlichen Neigungen zu unzuverlässigem Verhalten ist es aber nicht möglich, dem Ausfall von Versuchspersonen ganz vorzubeugen. Insgesamt gibt es einige Faktoren, die den Ausfall von Versuchspersonen begünstigen, aber nicht vermeidbar sind. Es können immer wieder Zwischenfälle auftreten, die verhindern, dass Versuchspersonen zur Untersuchung erscheinen. Verkehrsprobleme, Unfälle, gesundheitliche Probleme oder aber das spontane Erbitten von Hilfe aus dem nahen Umfeld der Versuchspersonen stellen einen Faktor dar, auf den Versuchsleiter*innen keinen Einfluss haben. Wenn den Versuchspersonen die Kontaktdaten der Versuchsleiter*innen leicht zugänglich sind, ist aber zumindest die Möglichkeit vorhanden, dass aus einem spontanen Ausfall eine vorhersehbare Abmeldung wird.

Die Ausfallquote ist dementsprechend nicht als ein Maß zu betrachten, dass durch die Planung der Untersuchung vorherbestimmt ist, sondern als variable Größe, auf die auch im Verlauf des Forschungsprozesses noch Einfluss genommen werden kann. Wenn zu Beginn der Testungen schon auffällt, dass die Ausfallquote erhöht ist, kann mehr direkter Kontakt mit potentiellen Versuchspersonen forciert werden sowie weitere Anreize und mehr Flexibilität bei den Terminen geboten werden. Sollten noch keine Erinnerungen an die Termine versendet worden sein, kann auch darauf zurückgegriffen werden. Außerdem kann der Ausfall von Versuchspersonen umgangen werden, indem diese direkt zum Untersuchungszeitpunkt in der Nähe des Untersuchungsortes spontan rekrutiert werden. Dies erfordert jedoch sowohl von dem bzw. der Versuchsleiter*in, als auch von den Teilnehmer*innen ein gewisses Maß an Spontanität und Flexibilität.

Die wohl zuverlässigste Gruppe an Versuchspersonen bilden Freunde, Familie und Bekannte. Wenn die gewünschte Stichprobengröße nur schwer zu erreichen ist, wird oft auf diese Personengruppe zurückgegriffen. Während sich das zwar positiv auf die Ausfallquote auswirkt, können die erhobenen Daten dadurch an Qualität verlieren. Insbesondere ist das der Fall, wenn die Versuchspersonen wissen, was erforscht werden soll oder die Untersuchung eine hohe Augenscheinvalidität aufweist (Koch, Peter & Müller, 2019, S. 135).

Allgemein ist anzuraten, sich vor der Rekrutierung gut zu überlegen, aus welchen Personengruppen sich die endgültige Stichprobe zusammensetzen wird und welche Bedürfnisse diese

Gruppen haben. Somit können Themen für die Teilnehmer*innen besonders ansprechend gestaltet werden, Anreize ausgewählt werden, die die Personen besonders motivieren und Termine angeboten werden, zu denen einem Großteil der Versuchspersonen vermutlich nichts dazwischen kommt. Bei Personengruppen wie Studierenden, die oft mit wenig Geld auskommen müssen, wirken sich monetäre Anreize eventuell stärker aus als bei Personen die vollzeitbeschäftigt sind und nur aus Interesse an einer Untersuchung teilnehmen. Bei arbeitenden Personen oder Gruppen, die ohnehin an einer gewissen Anzahl an Untersuchungen teilnehmen müssen, könnte dafür die Dauer der Untersuchung eine größere Rolle spielen.

Diese Arbeit zeigt, dass die Ausfallquote ein variables Maß ist, das durch viele Faktoren bestimmt wird und auch im Laufe einer Untersuchung noch beeinflusst werden kann. Um sie zu optimieren, empfiehlt es sich, sich mit der geplanten Stichprobe und den Vor- und Nachteilen des Untersuchungsdesigns schon vorab intensiv auseinanderzusetzen und Überlegungen anzustellen, wie die Rekrutierungsmaßnahmen optimal gesetzt und interessant gestaltet werden können. Zusammen mit ansprechenden Anreizen kann so die bestmögliche Wirkung auf die Ausfallquote entfaltet werden, um effiziente Forschung zu ermöglichen.

Quellenverzeichnis

- Aerny-Perreten, N., Domínguez-Berjón, F., Esteban-Vasallo, M. D., & García-Riolobos, C. (2015). Participation and factors associated with late or non-response to an online survey in primary care. *Journal of Evaluation in Clinical Practice*, *21*, 688–693.
- Altman, N. (2005). Manic society: Toward the depressive position. *Psychoanalytic Dialogues*, *15*, 321–346.
- Aronson, E., Wilson, T. D. & Akert, R. M. (2008). *Sozialpsychologie*. München: Pearson Studium.
- Baruch, Y. (1999). Response Rate in Academic Studies – A Comparative Analysis. *Human Relations*, *52*(4), 421-438.
- Baur, N., Florian, M. J. (2009). Stichprobenprobleme bei Online-Umfragen. In N. Jakob, H. Schoen, T. Zerback (Hrsg.), *Sozialforschung im Internet*. Wiesbaden: VS Verlag für Sozialwissenschaften,.
- Beller, S. (2016). *Empirisch forschen lernen: Konzepte, Methoden, Fallbeispiele, Tipps*. Bern: Hogrefe Verlag.
- Berkowitz, L., & Daniels, L. R. (1964). Affecting the salience of the social responsibility norm: effects of past help on the response to dependency relationships. *The Journal of Abnormal and Social Psychology*, *68*(3), 275–281.
- Bierhoff, H.-W. (2012). SV. Skala der sozialen Verantwortung. [Verfahrensdokumentation aus PSYNDEX Tests-Nr. 9004061 und Fragebogen]. In Leibniz-Zentrum für Psychologische Information und Dokumentation (ZPID) (Hrsg.), *Elektronisches Testarchiv*. Trier: ZPID. Abgerufen am 18. April 2022 von <https://www.testarchiv.eu/de/test/9004061>
- Blumenberg, C., Menezes, A. M. B., Gonçalves, H., Formoso Assunção, M. C., Wehrmeister, F. C., Barros, F. C., & Barros, A. J. D. (2019). The role of questionnaire length and reminders frequency on response rates to a web-based epidemiologic study: a randomised trial. *International Journal of Social Research Methodology*, *22*(6), 625-635.
- Boland, B., & Burnett, F. (2014). Optimising outpatient efficiency – Development of an innovative ‘Did Not Attend’ management approach. *International Journal of Psychiatry in Clinical Practice*, *18*(3), 217-219.

- Borella, E., Ludwig, C., Dirk, J., & de Ribaupierre, A. (2011). The Influence of Time of Testing on Interference, Working Memory, Processing Speed, and Vocabulary: Age Differences in Adulthood. *Experimental Aging Research*, 37(1), 76-107.
- Boulianne, S. (2012). Examining the Gender Effects of Different Incentive Amounts in a Web Survey. *Field Methods*, 25(1), 91-104.
- Brennan, M., & Hoek, J. (1992). The Behavior of Respondents, Nonrespondents, and Refusers Across Mail Surveys. *The Public Opinion Quarterly*, 56(4), 530-535.
- Burholt, V., Windle, G., Ferring, D., Balducci, C., Fagerström, C., Thissen, F., Weber, G., & Wenger, G. C. (2007). Reliability and Validity of the Older Americans Resources and Services (OARS) Social Resources Scale in Six European Countries. *The Journals of Gerontology: Series B: Psychological Sciences and Social Sciences*, 62(6), 371–379.
- Butt, G., McGuinness, L., Buller-Taylor, T. & Mitchell, S. (2013). Reasons for Nonattendance across the Hepatitis C Disease Course. *International Scholarly Research Notices*, 2013. Abgerufen am 18. April 2022 von <https://downloads.hindawi.com/archive/2013/579529.pdf>
- Clough, B. & Casey, L. (2014). Using SMS reminders in psychology clinics: A cautionary tale. *Behavioural and Cognitive Psychotherapy*, 42, 257-268.
- Cosgrove, J. A. (2018). Using a Small Cash Incentive to Increase Survey Response. *Administration and Policy in Mental Health and Mental Health Services Research*, 45, 813–819.
- Costa, P. T., Jr., McCrae, R. R., & Dye, D. A. (1991). Facet scales for Agreeableness and Conscientiousness: A revision of the NEO Personality Inventory. *Personality and Individual Differences*, 12, 887–898.
- Cough, H. G., McClosky, H. & Meehl, P. E. (1952). A personality scale for social responsibility. *The Journal of Abnormal and Social Psychology*, 47(1), 73–80.
- Crawford, S. D., Cooper, M. P. & Lamias, M. J. (2001). Web Surveys: Perceptions of Burden. *Social Science Computer Review*, 19(2), 146-162.
- Davern, M. (2013). Nonresponse Rates are a Problematic Indicator of Nonresponse Bias in Survey Research. *Health services research*, 48(3), 905–912.
- de Heer, W. (1999). International Response Trends: Results of an International Survey. *Journal of Official Statistics*, 15(2), 129-142.

- Digman, J. M. (1990). Personality Structure: Emergence of the Five-Factor Model. *Annual Review of Psychology*, 41(1), 417-440.
- Dillman, D. A., Sinclair M. D. & Clark, J. R. (1993). Effects of questionnaire length, respondent-friendly design, and a difficult question on response rates for occupant-addressed census mail surveys, *Public Opinion Quarterly*, 57(3), 289–304.
- Dudley, N.M., Orvis, K. A., Lebiecki, J. E. & Cortina, J. M. (2006). A Meta-Analytic Investigation of Conscientiousness in the Prediction of Job Performance: Examining the Intercorrelations and the Incremental Validity of Narrow Traits. *Journal of Applied Psychology*, 91(1), 40-57.
- Dykema, J., Stevenson, J., Klein, L., Kim, Y. & Day B. (2012). Effects of E-Mailed Versus Mailed Invitations and Incentives on Response Rates, Data Quality, and Costs in a Web Survey of University Faculty. *Social Science Computer Review*, 31(3), 359-370.
- Erwin, W. J. & Wheelright, L. A. (2002). Improving Mail Survey Response Rates Through the Use of a Monetary Incentive. *Journal of Mental Health Counseling*, 24(3), 247-255.
- Fang, J., Wen, C. & Pavur, R. (2012). Participation Willingness in Web Surveys: Exploring Effect of Sponsoring Corporation's and Survey Provider's Reputation. *Cyberpsychology, Behavior, and Social Networking*, 15(4), 195-199.
- Faught, K. S., Whitten, D. & Green, K. W. Jr. (2004). Doing Survey Research on the Internet: Yes, Timing Does Matter. *Journal of Computer Information Systems*, 44 (3), 26-34.
- Faulkner, K. K. & Cogan, R. (1990). Measures of shame and conflict tactics: effects of questionnaire order. *Psychological Reports*, 66, 1217-1218.
- Festinger, L. & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *The Journal of Abnormal and Social Psychology*, 58(2), 203–210.
- Fienup, D. M. & Dixon, M. R. (2003). Attendance probabilities of participants in a human experiment. *Psychological Reports*, 92, 445-448.
- Furse, D. H. & Stewart, D. W. (1984). Monetary Incentives Versus Promised Contribution to Charity: New Evidence on Mail Survey Response. *Journal of Marketing Research*, 19(3), 375-380.
- Galesic, M. (2006). Dropouts on the Web: Effects of Interest and Burden Experienced During an Online Survey. *Journal of Official Statistics*, 22(2), 313–328.

- Galesic, M. & Bosnjak, M. (2009). Effects of questionnaire length on participation and indicators of response quality in a web survey. *Public Opinion Quarterly*, 73(2), 349–360.
- Gheorghe, N. & Deaconu, D. (2018). When the Outside Strikes a Chord Inside: The Role of Identity in Taking Social Responsibility. *Transactional Analysis Journal*, 48(2), 152-165.
- Groves, R. M., Presser, S. & Dipko, S. (2004). The role of topic interest in survey participation decisions. *Public Opinion Quarterly*, 68(1), 2-31.
- Groves, R. M. (2006). Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly*, 70(5), 646-675.
- Groves, R. M. & Peytcheva, E. (2008). The impact of nonresponse rates on nonresponse bias. *Public Opinion Quarterly*, 72(2), 167-189.
- Gullo, M. J., Irvine, K., Feeney, G. F. X. & Connor, J. P. (2018). Short Message Service (SMS) reminders improve treatment attendance in alcohol dependence, but are less effective for patients high in impulsivity, *Addictive Behaviors*, 87, 97-100.
- Göritz, A. S. (2014) Determinants of the starting rate and the completion rate in online panel studies. In: Callegaro M, Baker R, Bethlehem J, Göritz AS, Krosnick JA, Lavrakas PJ (Hrsg.) *Online panel research: a data quality perspective* (S. 154–170). Chichester: Wiley.
- Han, V., Albaum, G., Wiley, J. B. & Thirkell, P. (2009). Applying theory to structure respondents' stated motivations for participating in web surveys. *Qualitative Market Research: An International Journal*, 12(4), 428-442.
- Hardigan, P. C., Popovici, I. & Carvajal, M. J. (2016). Response rate, response time, and economic costs of survey research: A randomized trial of practicing pharmacists. *Research in Social and Administrative Pharmacy*, 12(1), 141-148.
- Hasher, L., Chung, C., May, C. P. & Foong, N. (2002). Age, Time of Testing, and Proactive Interference. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 56(3), 200–207.
- Hawker, D. S. J. (2007). Increasing initial attendance at mental health out-patient clinics: opt-in systems and other interventions. *Psychiatric Bulletin*, 31(5), 179-182.
- Hazel, N. & Clark, A. (2013). Negotiating doorstep access: door-to-door survey researchers' strategies to obtain participation. *International Journal of Social Research Methodology*, 16(4), 307–321.

- Hinrichs-Krapels, S. & Grant, J. (2016). Exploring the effectiveness, efficiency and equity (3e's) of research and research impact assessment. *Palgrave Commun* 2. Abgerufen am 18. April 2022 von <https://www.nature.com/articles/palcomms201690#t2>
- Hippler, H.-J., Schwarz, N. & Singer, E. (1990). Der Einfluß von Datenschutzzusagen auf die Teilnahmebereitschaft an Umfragen. *ZUMA Nachrichten*, 14(27), 54-67.
- Hoerger, M. (2010). Participant Dropout as a Function of Survey Length in Internet-Mediated University Studies: Implications for Study Design and Voluntary Participation in Psychological Research. *Cyberpsychology, Behavior, and Social Networking*, 13(6), 697-700.
- Hough, L.M. (1992). The “Big Five” personality variables—construct confusion: Description versus prediction. *Human Performance*, 5, 139–155.
- Hough, L. M., Ones D. S. (2001). The structure, measurement, validity, and use of personality variables in industrial work, and organizational psychology. In Anderson N., Ones, D. S., Sinangil, H. K., Viswesvaran, C. (Hrsg.), *Handbook of industrial, work and organizational psychology* (Vol. 1, pp. 233–277). Sage.
- Jin, L. (2011). Improving response rates in web surveys with default setting: The effects of default on web survey participation and permission. *International Journal of Market Research*, 53(1), 75-94.
- Kannisto, K. A., Koivunen, M. H. & Välimäki, M. A. (2014). Use of Mobile Phone Text Message Reminders in Health Care Services: A Narrative Literature Review. *Journal of Medical Internet Research*, 16(10). Abgerufen am 18. April 2022 von <https://www.jmir.org/2014/10/e222>
- Keusch, F. (2015). Why do people participate in Web surveys? Applying survey participation theory to Internet survey data collection. *Management Review Quarterly*, 65, 183-216.
- Keusch, F., Batinic B. & Mayerhofer, W. (2014). Motives for joining nonprobability online panels and their association with survey participation behaviour. In M. Callegaro, R. Baker, J. Bethlehem, A.S. Göritz, J.A. Krosnick and P.J. Lavrakas (Hrsg.), *Online Panel Research: A Data Quality Perspective* (S. 171-191). Chichester: Wiley.
- Koch, T., Peter, C. & Müller, P. (2019). *Das Experiment in der Kommunikations- und Medienwissenschaft: Grundlagen, Durchführung und Auswertung experimenteller Forschung*. Wiesbaden: Springer VS.

- Konradt, U. & Fary, Y. (2006). Determinanten der Motivation und der Bereitschaft zur Teilnahme an Fragebogenstudien. *Zeitschrift für Psychologie*, 214(2), 87-96.
- Jenkins, P. E. (2017). Reducing Non-Attendance Rates for Assessment at an Eating Disorders Service: A Quality Improvement Initiative. *Community Mental Health Journal*, 53, 878–882.
- Koloski, N. A., Talley, N. J., Boyce, P. M. & Morris-yates, A. D. (2001). The effects of questionnaire length and lottery ticket inducement on the response rate in mail surveys. *Psychology and Health*, 16, 67-75.
- Laguilles, J. S., Williams, E. A. & Saunders, D. B. (2011). Can Lottery Incentives Boost Web Survey Response Rates? Findings from Four Experiments. *Research in Higher Education*, 52, 537-553.
- Lewandowska, K., Wachowicz, B., Marek, T., Oginska, H. & Fafrowicz, M. (2018). Would you say “yes” in the evening? Time-of-day effect on response bias in four types of working memory recognition tasks. *Chronobiology International*, 35(1), 80-89.
- Liu, M., Kuriakose, N., Cohen, J. & Cho, S. (2015). Impact of Web Survey Invitation Design on Survey Participation, Respondents, and Survey Responses. *Social Science Computer Review*, 34(5), 631-644.
- Lozar Manfreda, K., Bosnjak, M., Berzelak, J., Haas, I. & Vehovar, V. (2008). Web surveys versus other survey modes: a meta-analysis comparing response rates. *International Journal of Market Research*, 50(1), 79-104.
- Lynn, P. (2014). Longer interviews may not affect subsequent survey participation propensity. *Public Opinion Quarterly*, 78(2), 500-509.
- Martin, C. L. (1994). The impact of topic interest on mail survey response behaviour. *Journal of the Market Research Society*. 36(4), 327-338.
- Mayhorn, C. B., Fisk, A. D. & Whittle, J. D. (2002). Decisions, Decisions: Analysis of Age, Cohort, and Time of Testing on Framing of Risky Decision Options. *Human Factors*, 44(4), 515-521.
- Maynard, D. W., Freese, J. & Schaeffer, N. C. (2010). Calling for Participation: Requests, Blocking Moves, and Rational (Inter)action in Survey Introductions. *American Sociological Review* 75(5) 791-814.

- Maynard, D. W. & Hollander, M. M. (2014). Asking to Speak to Another: A Skill for Soliciting Survey Participation. *Research on Language and Social Interaction*, 47(1), 28-48.
- McCrae, R. R. & John, O. P. (1992). An Introduction to the Five-Factor Model and Its Applications. *Journal of Personality*, 60(2), 175-215.
- Mercer, A., Caporaso, A., Cantor, D. & Townsend, R. (2015). How much gets you how much? Monetary incentives and response rates in household surveys. *Public Opinion Quarterly*, 79(1), 105-129.
- Milward, J., Lynskey, M. & Strang, J. (2014). Solving the problem of non-attendance in substance abuse services. *Drug and Alcohol Review*, 33, 625–636.
- Olaru, G., Burrus, J., MacCann, C., Zaromb, F. M., Wilhelm, O. & Roberts, R. D. (2019). Situational Judgment Tests as a method for measuring personality: Development and validity evidence for a test of Dependability. *PLoS ONE*, 14(2), art. no. e0211884.
- Patrick, M. E., Singer, E., Boyd, C. J., Cranford, J. A. & McCabe, S. E. (2013). Incentives for college student participation in web-based substance use surveys. *Addictive Behaviors*, 38, 1710–1714.
- Pfarr, K., Blohm, M., Blom, A. G., Erdel, B., Felderer, B., Fräßdorf, M., Hajek, K., Helmschrott, S., Kleinert, C., Koch, A., Krieger, U., Kroh, M., Martin, S., Saßenroth, D., Schmiedeberg, C., Trüdinger, E.-M. & Rammstedt, B. (2015). Are Incentive Effects on Response Rates and Nonresponse Bias in Large-scale, Face-to-face Surveys Generalizable to Germany? Evidence from Ten Experiments. *Public Opinion Quarterly*, 79(3), 740–768.
- Poll, R., Allmark, P. & Tod, A. M. (2017). Reasons for missed appointments with a hepatitis C outreach clinic: A qualitative study. *International Journal of Drug Policy*, 39, 130-137.
- Poon, P. S., Albaum, G. & Evangelista, F. U. (2003). Why People Respond to Surveys. *Journal of International Consumer Marketing*, 16(2), 75-90.
- Porter, S. R. & Whitcomb, M. E. (2005). E-mail Subject Lines and Their Effect on Web Survey Viewing and Response. *Social Science Computer Review*, 23(3), 380-387.
- Porter, S. R., Whitcomb, M. E. & Weitzer, W. H. (2004). Multiple Surveys of Students and Survey Fatigue. *New Directions for Institutional Research*, 124, 63-73.

- Robb, K. A., Gattling, L. & Wardle, J. (2017). What impact do questionnaire length and monetary incentives have on mailed health psychology survey response?. *British Journal of Health Psychology*, 22, 671–685.
- Roberts, B. W., Chernyshenko, O. S., Stark, S. & Goldberg, L. R. (2005). The structure of conscientiousness: An empirical investigation based on seven major personality questionnaires. *Personnel Psychology*, 58, 103-139.
- Roszkowski, M. J. & Bean, A. G. (1990). Believe It or Not! Longer Questionnaires Have Lower Response Rates. *Journal of Business and Psychology*, 4(4), 495-509.
- Reiß, S. & Sarris, V. (2012). *Experimentelle Psychologie: Von der Theorie zur Praxis*. München: Pearson Studium.
- Sanchez, M. E. (1992). Effects of questionnaire design on the quality of survey data. *Public Opinion Quarterly*, 56(2), 206-217.
- Satow, L. (2011). B5T. Psychomeda Big-Five-Persönlichkeitstest. Skalendokumentation und Normen sowie Fragebogen mit Instruktion [PSYNDEX Tests-Nr. 9006357]. In Leibniz-Zentrum für Psychologische Information und Dokumentation (ZPID) (Hrsg.), *Elektronisches Testarchiv*. Trier: ZPID. Abgerufen am 18. April 2022 von <https://www.testarchiv.eu/de/test/9006357>
- Satow, L. (2012). *Big-Five-Persönlichkeitstest (B5T): Test- und Skalendokumentation*. Abgerufen am am 18. April 2022 von <https://www.drSATOW.de/tests/persoenlichkeitstest/>
- Sauermann, H. & Roach, M. (2013). Increasing web survey response rates in innovation research: An experimental study of static and dynamic contact design features. *Research Policy*, 42(1), 273-286.
- Scandell, D. J., Wlazelek, B., Bentelspacher, C. E., Rees, K. S. & Thomas, S. L. (2003). Effects of Questionnaire Order on Self-Reported Sexual Behavior, Risk Perceptions, and Ratings of HIV and STD Protection Strategies. *Journal of Psychology & Human Sexuality*, 15(1), 53-67.
- Schnauber, A. & Daschmann, G. (2008). States oder Traits? : was beeinflusst die Teilnahmebereitschaft an telefonischen Interviews?. *Methoden, Daten, Analysen*, 2(2), 97-123.
- Shih, T. & Fan, X. (2008). Comparing Response Rates from Web and Mail Surveys: A Meta-Analysis. *Field Methods*, 20 (3), 249-271.

- Singer, E. (1978). Informed Consent: Consequences for Response Rate and Response Quality in Social Surveys. *American Sociological Review*, 43(2), 144-162.
- Singer, E., Groves, R. M. & Corning, A. D. (1999). Differential incentives: Beliefs about practices, perceptions of equity, and effects on survey participation. *Public Opinion Quarterly*, 63(2), 251-260.
- SoSci Panel. (n.d.). *SoSci Panel für Wissenschaftlerinnen und Wissenschaftlern*. Abgerufen am 19. September 2020 von <https://www.soscipanel.de/researchers.php>
- Stewart, G. L. (1999). Trait bandwidth and stages of job performance: Assessing differential effects for conscientiousness and its subtraits. *Journal of Applied Psychology*, 84(6), 959-968.
- Strauss, A. & Corbin, J. (1996). *Grounded Theory: Grundlagen Qualitativer Sozialforschung*. Weinheim: Beltz, Psychologie Verlags Union.
- Strübing, J. (2014). *Grounded Theory: Zur sozialtheoretischen und epistemologischen Fundierung eines pragmatischen Forschungsstils*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Sturgis, P., Williams, J., Brunton-Smith, I. & Moore, J. (2017). Fieldwork Effort, Response Rate, and the Distribution of Survey Outcomes: A Multilevel Meta-analysis, *Public Opinion Quarterly*, 81(2), 523-542.
- SurveyCircle. (n.d.). SurveyCircle. Abgerufen am 19. September 2020 von <https://www.surveycircle.com/de/>
- Tolonen, H., Ahonen, S., Jentoft, S., Kuulasmaa, K. & Heldal, J. (2015). Differences in participation rates and lessons learned about recruitment of participants – The European Health Examination Survey Pilot Project. *Scandinavian Journal of Public Health*, 43(2), 212-219.
- Tourangeau, R. (2017). Presidential Address: Paradoxes of Nonresponse. *Public Opinion Quarterly*, 81(3), 803-814.
- Tourangeau, R., Groves, R. M., Kennedy, C. & Yan, T. (2009). The Presentation of a Web Survey, Nonresponse and Measurement Error among Members of Web Panel. *Journal of Official Statistics*, 25(3), 299–321.

- Universität Graz Institut für Psychologie. (n.d.). *Untersuchungsscheine*. Abgerufen am 19. September 2020 von <https://psychologie.uni-graz.at/de/service-studierende/studieren/untersuchungsscheine/>
- Vercruyssen, A., Roose, H., Carton, A. & Van De Putte, B. (2014). The effect of busyness on survey participation: being too busy or feeling too busy to cooperate?. *International Journal of Social Research Methodology*, 17(4), 357-371.
- Vicente, P. & Reis, E. (2010). Using Questionnaire Design to Fight Nonresponse Bias in Web Surveys. *Social Science Computer Review*, 28(2), 251-267.
- Virtanen, V., Sirkiä, T. & Jokiranta, V. (2007). Reducing Nonresponse by SMS Reminders in Mail Surveys. *Social Science Computer Review*, 25(3), 384-395.
- Wagner, J. (2012). A Comparison of Alternative Indicators for the Risk of Nonresponse Bias, *Public Opinion Quarterly*, 76(3), 555–575.
- Wolfe, E. W., Converse, P. D., Airen, O. & Bodenhorn, N. (2009). Unit and Item Nonresponses and Ancillary Information in Web- and Paper-Based Questionnaires Administered to School Counselors. *Measurement and Evaluation in Counseling and Development*, 42(2), 92-103.
- Wright, G. (2015). An empirical examination of the relationship between nonresponse rate and nonresponse bias. *Statistical Journal of the IAOS*, 31(2), 305-315.

Anhang

Im Anhang befinden sich die Dokumentation der Literaturrecherche, ergänzende Daten zur Hauptuntersuchung sowie Abbildungen dazu, wie die Voruntersuchung und die Hauptuntersuchung gestaltet wurden.

Bei der Abgabe wurde dieser Masterarbeit ein USB-Stick mit den erhobenen Daten beigelegt. Auf Anfrage werden die Daten dieser Untersuchung von Em.Univ.-Prof. Dr.rer.nat. Dietrich Albert oder vom Autor dieser Arbeit gerne zur Verfügung gestellt.

Dokumentation der Literaturrecherche

Um wissenschaftliche Literatur zu finden, wurde eine Suche in den wissenschaftlichen Datenbanken PsychINFO, PSYINDEXplus und PubMed durchgeführt. Zusätzlich wurde nach weiterführender Literatur gesucht, auf die in den so gefundenen Studien verwiesen wurde. Für die Recherche wurden folgende Schlüsselwörter verwendet: Appointments, attendance, attendance rate, absence, big five, behavior, behaviour, conscientiousness, dependability, dependable, dropout, experiment, five factor model, motivation, non-attendance, nonresponse bias, nonresponse rate, participation, participation rate, personality, participant, probability, prosocial, questionnaire, questionnaire length, recruitment, reliability, reliable behavior, reliable behaviour, reliableness, research, response, responsibility, scientific, social responsibility, studies, study, time of day, time of testing, university studies, volunteer, willingness, Fragebogenuntersuchung, Gewissenhaftigkeit, Pflichtbewusstsein, Selektionsverzerrung, Teilnahmebereitschaft, Zuverlässigkeit.

Die Literaturrecherche wurde im März 2019 beendet. Publikationen die später veröffentlicht wurden, wurden in der vorliegenden Arbeit daher nicht mehr berücksichtigt.

Ergänzende Daten zur Hauptuntersuchung

Häufigkeitstabelle zur Teilnahmewahrscheinlichkeit

Die Wahrscheinlichkeit tatsächlich teilzunehmen...	sinkt stark	sinkt ein wenig	bleibt gleich	steigt ein wenig	steigt stark	keine Anmeldung
Wenn ich am Vortag eine Erinnerung an den Termin erhalte...	3	4	50	75	44	4
Wenn ich von dem/der Versuchsleiter/in keine Erinnerung an den Termin erhalte...	5	53	94	19	6	2
Wenn ich den Untersuchungstermin frei wählen kann...	2	5	46	73	51	3
Wenn ich bei der Untersuchung einen Leistungstest bearbeiten soll...	3	36	120	10	5	3
Wenn bei der Untersuchung aufwendige technische Hilfsmittel zum Einsatz kommen (EEG, fMRT, etc.)...	2	34	96	25	12	11
Wenn ich bei der Untersuchung Fragebögen bearbeiten soll...	2	5	147	14	9	2
Wenn bei der Untersuchung meine aktive Teilnahme besonders gefordert ist...	4	27	127	14	7	0
Wenn die Untersuchung weniger als 30 Minuten dauert...	2	11	79	57	29	2
Wenn die Untersuchung zwischen 30 und 90 Minuten dauert...	2	18	134	20	3	2
Wenn die Untersuchung mehr als 1 1/2 Stunden dauert...	20	53	82	5	7	13

<i>Wenn ich den/die Versuchsleiter/in vor der Untersuchung bereits persönlich getroffen und mit ihr/ihm gesprochen habe...</i>	0	2	52	90	33	2
<i>Wenn ich vor der Untersuchung bereits per Mail oder SMS mit dem/r Versuchsleiter/in kommuniziert habe...</i>	2	2	53	83	37	3
<i>Wenn ich den Untersuchungstermin ohne persönlichen Kontakt mit dem/der Versuchsleiter/in vereinbart haben... (Z.B. durch einen Eintrag in eine Terminliste)</i>	4	33	122	14	5	1
<i>Wenn der Untersuchungstermin frühestens einen Monat nach der Anmeldung stattfindet...</i>	12	51	102	7	2	6
<i>Wenn ich mit einem/r Freund/in oder meinem/r Partner/in gemeinsam zur Untersuchung erscheinen soll...</i>	6	22	70	53	23	5
<i>Wenn ich in der Ausschreibung zur Untersuchung mit „Sie“ angesprochen wurde...</i>	4	13	156	4	2	1
<i>Wenn ich in der Ausschreibung zur Untersuchung mit „Du“ angesprochen werde...</i>	1	9	155	9	3	2
<i>Wenn bei der Untersuchung aufwendige technische Hilfsmittel zum Einsatz kommen (EEG, fMRT, etc.)...</i>	2	34	96	25	12	11

Wenn die Untersuchung aus mehreren Terminen besteht...	25	88	40	13	0	13
Wenn ich nach einer Voruntersuchung zur Hauptuntersuchung eingeladen werde...	3	26	82	46	18	4
Wenn bei der Untersuchung meine aktive Teilnahme besonders gefordert ist...	4	27	127	14	7	0
Wenn ich 5€ für die Teilnahme erhalte...	4	5	75	67	27	2
Wenn ich 10€/Stunde für die Teilnahme erhalte...	1	3	39	67	65	5
Wenn ich 30€ für die Teilnahme an der Untersuchung erhalte...	0	1	32	52	86	8
Wenn ich nach der Untersuchungsteilnahme bei einem Gewinnspiel mitmachen darf...	1	3	114	50	11	1
Wenn ich das Thema der Untersuchung nicht interessant finde...	27	58	50	14	11	20
Wenn ich das Thema der Untersuchung besonders interessant finde...	0	1	28	86	62	2
Wenn ich Süßigkeiten für die Teilnahme bekomme...	4	7	146	20	1	1
Wenn ich nach der Untersuchungsteilnahme bei einem Gewinnspiel mitmachen darf...	1	3	114	50	11	1
Wenn ich 5€ für die Teilnahme erhalte...	4	5	75	67	27	2
Wenn ich 10€/Stunde für die Teilnahme erhalte...	1	3	39	67	65	5
Wenn ich 30€ für die Teilnahme an der Untersuchung erhalte...	0	1	32	52	86	8

Wenn die Untersuchung vor 8:30 Uhr startet...	23	45	75	7	3	27
Wenn die Untersuchung zwischen 9:00 Uhr und 17:00 Uhr startet...	4	8	127	27	8	5
Wenn die Untersuchung nach 18:30 Uhr startet...	16	45	88	13	3	14
Wenn mein/e Chef/in mich (kurz) vor dem Untersuchungstermin bittet bei der Arbeit einzuspringen...	74	69	21	4	6	6
Wenn ein Familienmitglied mich vor dem Untersuchungstermin um Hilfe bittet...	68	70	16	7	8	10
Wenn ein/e Freund/in mich vor dem Untersuchungstermin spontan um Hilfe bittet...	37	79	30	11	10	13
Wenn draußen besonders schönes Wetter ist...	6	24	131	5	8	6
Wenn draußen leicht bewölkttes Wetter ist...	0	4	162	8	5	1
Wenn draußen besonders schlechtes Wetter ist...	4	28	133	7	2	5
Wenn die Untersuchung von einem/r Freund/in geleitet wird...	2	2	20	22	117	17
Wenn die Untersuchung von einem/r Fremden geleitet wird...	3	23	122	8	22	2
Wenn ich mich nicht an den Termin erinnere	93	43	25	4	7	7

Tabelle zu ergänzenden Daten 1: Die Häufigkeiten der angegebenen Teilnahmewahrscheinlichkeiten unter den vorgegebenen Bedingungen. Die Versuchspersonen gaben an, ob die Wahrscheinlichkeit, dass sie zur Untersuchung erscheinen „stark sinkt“, „ein wenig sinkt“, „gleich bleibt“, „ein wenig steigt“ oder „stark steigt“. Außerdem bestand die Möglichkeit eine Anmeldung zur Untersuchung auszuschließen.

Gestaltung der Voruntersuchung

In der Folge wird die Einleitung und Datenschutzerklärung der Voruntersuchung wiedergegeben sowie der Fragebogen in Form von Screenshots abgebildet.

Einleitung und Datenschutzerklärung

Im Rahmen meiner Masterarbeit, befasse ich mich mit dem Ausfall von Versuchspersonen bei psychologischen Studien. Im Folgenden werde ich Ihnen einige Fragen dazu stellen, bitte beziehen Sie sich bei den Antworten nur auf einen von Ihnen durchgeführten Versuch, zu dem die Versuchspersonen zu einem vorab vereinbarten Termin erscheinen mussten.

Selbstverständlich werden Ihre Daten anonym und vertraulich behandelt und nicht an Dritte weitergegeben.

Vielen Dank schon im Voraus für Ihre Unterstützung!

Für Rückfragen stehe ich Ihnen gerne per Mail (pierre.pacher@edu.uni-graz.at) zur Verfügung.

EINVERSTÄNDNISERKLÄRUNG

Ich bin ausreichend über die Ziele und Methoden, die möglichen Risiken und den Nutzen der Studie informiert worden. Ich bin damit einverstanden, dass im Rahmen der Studie persönliche Daten aufgezeichnet werden, und dass diese anonymisiert (d.h. ohne Rückschlussmöglichkeit auf meine Person) gespeichert und ausgewertet werden. Alle im Rahmen der Studie erhobenen Daten werden strikt vertraulich und gemäß dem Datenschutz behandelt, ausschließlich für wissenschaftliche Zwecke verwendet und an niemanden weitergegeben. Der wissenschaftlichen Auswertung der anonymisierten Daten und einer möglichen Veröffentlichung der Studienergebnisse stimme ich hiermit zu.

Ich weiß, dass meine Studienteilnahme freiwillig ist und dass ich meine Teilnahme an der Studie jederzeit und ohne Angabe von Gründen beenden kann und ohne dass mir daraus Nachteile entstehen. Insbesondere kann ich zu jedem beliebigen Zeitpunkt die Löschung meiner Daten verlangen.

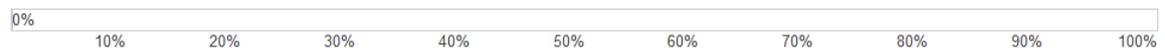
Mit einem Klick auf „Weiter“ gebe ich meine freiwillige Zustimmung zur Teilnahme an dieser Untersuchung.

Fragebogen der Voruntersuchung



Umfrage im Rahmen des Studiums

Einflussfaktoren auf zuverlässiges Verhalten am Beispiel der Teilnahme an psychologischen Untersuchungen



Thema der Untersuchung

Bitte geben Sie zunächst das Thema der Untersuchung an, auf die Sie sich bei der Beantwortung der Fragen beziehen werden. (So soll vermieden werden, dass mehrfach Daten zur selben Untersuchung erhoben werden.)

* Im Zuge welcher Forschungstätigkeit haben Sie diese Untersuchung durchgeführt?

Bitte wählen Sie eine der folgenden Antworten:

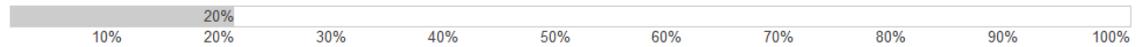
- Seminar: Praktikum für Forschungsmethodik
- Bachelorarbeit
- Masterarbeit
- Universitäres Forschungsprojekt
- Privates Forschungsprojekt
- Sonstiges:

Template © Denis Chenu, Sondages Pro

Abbildung 1: Teil 1 des Fragebogens der bei der Voruntersuchung verwendet wurde



Einflussfaktoren auf zuverlässiges Verhalten am Beispiel der Teilnahme an psychologischen Untersuchungen



Rekrutierung von Versuchspersonen

Welche Methoden haben Sie zur Rekrutierung von Versuchspersonen verwendet?

Bitte wählen Sie einen oder mehrere Punkte aus der Liste aus.

- Postings in sozialen Medien (Facebook, Twitter, etc.)
- Aushang an der Universität
- Aushang außerhalb der Universität
- Aussendung über einen E-Mail-Verteiler der Universität (z.B. student.umfrage)
- Aussendung über einen nichtuniversitären E-Mail-Verteiler
- Kontaktaufnahme mit zuvor unbekanntem Personen von einer privaten E-Mail-Adresse
- telefonische Kontaktaufnahme mit zuvor unbekanntem Personen
- persönliche Kontaktaufnahme mit zuvor unbekanntem Personen durch direktes Ansprechen
- Vorsprechen in Lehrveranstaltung zum Anwerben von Versuchspersonen
- Rekrutierung von Verwandten, Freunden oder Bekannten

Sonstiges:

* Welche Versuchsdauer (in Minuten) haben Sie für Ihre Untersuchung bei der Rekrutierung angegeben?

▲ Bitte beantworten Sie diese Frage. Bitte machen Sie mindestens ein Kreuz.

Kommentieren wenn eine Antwort gewählt wird
Bitte wählen Sie einen oder mehrere Punkte aus der Liste aus.

- Angabe in Minuten:
- Keine Angabe

🔗 (Sollten sie keine Angabe gemacht haben, tragen Sie in das Feld neben "Keine Angabe" bitte 0 ein)

Abbildung 2: Teil 2 des Fragebogens der bei der Voruntersuchung verwendet wurde

Welche zu erwartenden Untersuchungsmethoden haben Sie bei der Rekrutierung angegeben?
Bitte wählen Sie einen oder mehrere Punkte aus der Liste aus.

Persönlichkeitsfragebogen

andere Fragebögen

Intelligenztests

andere Leistungstests

physiologische Messungen (wie Blutdruck, Puls, etc.)

fMRT

EEG

aktive Übungen (zu kognitiven Prozessen)

aktive Teilnahme (Musik, Sport, Videos, etc.)

apparativer Versuchsaufbau (z.B. Messung der Blickbewegungen)

Computertesting

Gruppentesting

Einzeltesting

Sonstiges:

Haben Sie den Versuchspersonen bei der Rekrutierung eine Belohnung angeboten? (Wenn ja, welche?)
Bitte wählen Sie einen oder mehrere Punkte aus der Liste aus.

Versuchsscheine

Geld

Schokolade bzw. andere Süßigkeiten

Teilnahme an einer Verlosung unter den VersuchsteilnehmerInnen

Keine

Sonstiges:

*** Gab es Kriterien, die die Versuchspersonen erfüllen mussten, um an Ihrer Studie teilnehmen zu können? (Wenn ja, welche?)**

⚠ Bitte beantworten Sie diese Frage.

Bitte wählen Sie eine der folgenden Antworten:

Ja, folgende:

Es gab keine Teilnahme-kriterien

Bitte geben Sie hier Ihren Kommentar ein:

Abbildung 3: Teil 3 des Fragebogens der bei der Voruntersuchung verwendet wurde



Einflussfaktoren auf zuverlässiges Verhalten am Beispiel der Teilnahme an psychologischen Untersuchungen



Teilnahme an der Untersuchung

Haben sie die angemeldeten Personen vor der Untersuchung noch einmal an den Termin erinnert?

Bitte wählen Sie einen oder mehrere Punkte aus der Liste aus.

<input type="checkbox"/> Nein
<input type="checkbox"/> per Mail
<input type="checkbox"/> per SMS
<input type="checkbox"/> per Anruf
Sonstiges: <input type="text"/>

* Haben sie den Ausfall von Versuchspersonen bei Ihrer Untersuchung als Problem wahrgenommen?

Bitte wählen Sie eine der folgenden Antworten:

<input type="radio"/> ein großes Problem
<input type="radio"/> eher ein Problem
<input type="radio"/> eher kein Problem
<input type="radio"/> kein Problem

* Wie viele Versuchspersonen haben sich **ursprünglich** zu Ihrer Untersuchung **angemeldet**?

Kommentieren wenn eine Antwort gewählt wird

Bitte wählen Sie einen oder mehrere Punkte aus der Liste aus.

<input type="checkbox"/> Genaue Anzahl:	<input type="text"/>
<input type="checkbox"/> Schätzung:	<input type="text"/>

* Wie viele Personen haben **zugesagt** an Ihrer Untersuchung teilzunehmen, haben den vereinbarten Termin jedoch **nicht eingehalten**?

Kommentieren wenn eine Antwort gewählt wird

Bitte wählen Sie einen oder mehrere Punkte aus der Liste aus.

<input type="checkbox"/> Genaue Anzahl:	<input type="text"/>
<input type="checkbox"/> Schätzung:	<input type="text"/>

* Wie viele Personen haben **schlussendlich** an Ihrer Untersuchung **teilgenommen**?

Kommentieren wenn eine Antwort gewählt wird

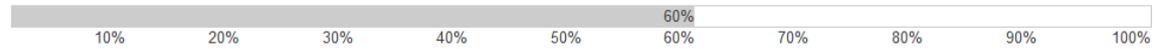
Bitte wählen Sie einen oder mehrere Punkte aus der Liste aus.

<input type="checkbox"/> Genaue Anzahl:	<input type="text"/>
<input type="checkbox"/> Schätzung:	<input type="text"/>

Abbildung 4: Teil 4 des Fragebogens der bei der Voruntersuchung verwendet wurde



Einflussfaktoren auf zuverlässiges Verhalten am Beispiel der Teilnahme an psychologischen Untersuchungen



Gründe für den Ausfall von Versuchspersonen

*
Gab es bei Ihrer Untersuchung Besonderheiten (bezüglich Rekrutierung, Ablauf, etc.) durch die sich Ihre Untersuchung von anderen Untersuchungen unterscheidet?

Bitte wählen Sie eine der folgenden Antworten:

<input type="radio"/> Ja, folgende:	Bitte geben Sie hier Ihren Kommentar ein: <input type="text"/>
<input type="radio"/> Nein	

* Haben Versuchspersonen, die trotz fixer Zusage an Ihrer Untersuchung teilzunehmen, nicht teilgenommen haben, Gründe für ihre Abwesenheit angegeben? (Wenn ja, welche Gründe?)

Bitte wählen Sie eine der folgenden Antworten:

<input type="radio"/> Ja, folgende:	Bitte geben Sie hier Ihren Kommentar ein: <input type="text"/>
<input type="radio"/> Nein, niemand hat einen Grund angegeben	

Was glauben Sie, könnten die Gründe sein, warum die ausgefallenen Versuchspersonen nicht zu Ihrer Untersuchung erschienen sind?

Haben Sie sonst noch irgendwelche Gedanken zum Thema die Sie für relevant erachten?

Abbildung 5: Teil 5 des Fragebogens der bei der Voruntersuchung verwendet wurde

Gestaltung der Hauptuntersuchung

In der Folge wird die Einleitung und Datenschutzerklärung der Hauptuntersuchung wiedergegeben sowie der Fragebogen in Form von Screenshots abgebildet. Auf die Darstellung von B5T (Satow, 2011) und SV-Skala (Bierhoff, 2012) wurde aus lizenzrechtlichen Gründen verzichtet.

Einleitung und Datenschutzerklärung

Teilnahme an der Untersuchung zur Masterarbeit „Einflussfaktoren auf zuverlässiges Verhalten am Beispiel der Teilnahme an psychologischen Studien“

Die vorliegende Studie untersucht die Zusammenhänge zwischen Persönlichkeitsmerkmalen, motivationalen Faktoren sowie Umweltfaktoren und der Zuverlässigkeit von potentiellen Versuchspersonen tatsächlich an einer psychologischen Untersuchung teilzunehmen, nachdem sie sich dafür angemeldet haben. Hierfür werden Sie gebeten verschiedene Fragebögen zu bearbeiten.

EINVERSTÄNDNISERKLÄRUNG

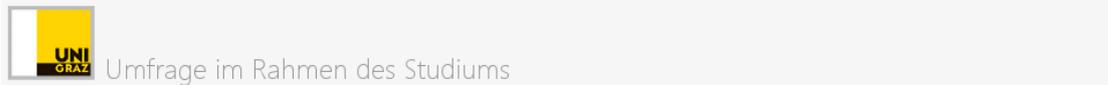
Ich bin ausreichend über die Ziele und Methoden, die möglichen Risiken und den Nutzen der Studie informiert worden. Ich hatte ausreichend Gelegenheit, die Studie mit der Versuchsleitung zu besprechen und Fragen zu stellen. Alle meine Fragen und Bedenken wurden zu meiner Zufriedenheit beantwortet bzw. geklärt.

Ich bin damit einverstanden, dass im Rahmen der Studie persönliche Daten aufgezeichnet werden, und dass diese anonymisiert (d.h. ohne Rückschlussmöglichkeit auf meine Person) gespeichert und ausgewertet werden. Alle im Rahmen der Studie erhobenen Daten werden strikt vertraulich und gemäß dem Datenschutz behandelt, ausschließlich für wissenschaftliche Zwecke verwendet und an niemanden weitergegeben. Der wissenschaftlichen Auswertung der anonymisierten Daten und einer möglichen Veröffentlichung der Studienergebnisse stimme ich hiermit zu.

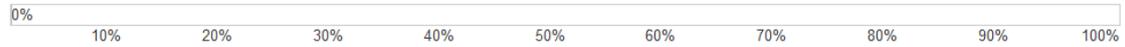
Ich weiß, dass meine Studienteilnahme freiwillig ist und dass ich meine Teilnahme an der Studie jederzeit und ohne Angabe von Gründen beenden kann und ohne dass mir daraus Nachteile entstehen. Insbesondere kann ich zu jedem beliebigen Zeitpunkt die Löschung meiner Daten verlangen.

Mit einem Klick auf „Weiter“ gebe ich meine freiwillige Zustimmung zur Teilnahme an dieser Untersuchung.

Erhebung demografischer Daten bei der Hauptuntersuchung



Effizienter Forschen



Bearbeitungsfortschritt

Demografische Daten

Geben Sie zunächst bitte folgende Daten zu Ihrer Person an:

* Alter:
In dieses Feld dürfen nur Zahlen eingegeben werden.

Geschlecht

weiblich

männlich

keine Antwort

* Welcher Tätigkeit bzw. welchen Tätigkeiten gehen Sie nach?
Bitte wählen Sie einen oder mehrere Punkte aus der Liste aus.

Schule

Studium

Geringfügig angestellt

Teilzeit angestellt

Vollzeit angestellt

Selbstständig

Sonstiges:

Abbildung 6: Teil 1 der Erhebung demografischer Daten

* Was ist Ihr höchster abgeschlossener Ausbildungsgrad?

Bitte wählen Sie eine der folgenden Antworten:

Pflichtschulabschluss

Lehre

Matura

Bachelor

Master

Magister/Magistra

Doktorat

Sonstiges:

* Studieren Sie Psychologie bzw. haben Sie Psychologie studiert?

Bitte wählen Sie einen oder mehrere Punkte aus der Liste aus.

Ja

Nein

Sonstiges:

* An wie vielen psychologischen Untersuchungen an Universitäten haben Sie (in etwa) bereits teilgenommen:

In dieses Feld dürfen nur Zahlen eingegeben werden.

* Wie viele psychologische Untersuchungen haben Sie bereits selbst durchgeführt:

In dieses Feld dürfen nur Zahlen eingegeben werden.

* Wie lange benötigen Sie für gewöhnlich von Ihrem Wohnort aus bis zur Karl-Franzens Universität Graz (in Minuten):

In dieses Feld dürfen nur Zahlen eingegeben werden.

Abbildung 7: Teil 2 der Erhebung demografischer Daten

Erhebung der Gründe für die Teilnahme und wie Versuchspersonen auf die Untersuchung aufmerksam wurden bei der Hauptuntersuchung

Die Antwortkategorien wurden bei diesen Fragen daran angepasst, ob die Versuchspersonen bei der Untersuchung an der Universität, an der Online-Umfrage oder bei der Umfrage über „SurveyCircle“ teilnehmen sollten. Die Antwortkategorien spiegeln die am häufigsten zu erwartenden Antwortmöglichkeiten wieder. Es gab aber auch die Möglichkeit Antwortkategorien frei zu ergänzen. In der Folge werden Screenshots von den drei unterschiedlichen Varianten dieses Teils des Fragebogens abgebildet.

The image shows two screenshots of a survey form. The first screenshot is for the question: '* Wie sind Sie auf diese Untersuchung aufmerksam geworden?' (How did you become aware of this survey?). Below the question is the instruction: 'Bitte wählen Sie einen oder mehrere Punkte aus der Liste aus.' (Please select one or more points from the list). The list of options includes: 'Mail von student.umfrage', 'Aushang an der Universität', 'Flyer bekommen', 'In sozialen Medien gesehen (Facebook, Instagram,...)', 'In einer Vorlesung davon gehört', 'Von FreundInnen/StudienkollegInnen davon gehört', and 'Kenne den Versuchsleiter'. There is also a 'Sonstiges:' (Other:) field with a text input box. The second screenshot is for the question: '* Was hat Sie zur Teilnahme an dieser Untersuchung bewogen?' (What motivated you to participate in this survey?). Below the question is the instruction: 'Bitte wählen Sie einen oder mehrere Punkte aus der Liste aus.' (Please select one or more points from the list). The list of options includes: 'Interesse', 'Versuchsschein benötigt', 'Terminfindung einfach', 'Langeweile', 'Versprochene Belohnung (Süßigkeiten)', 'Untersuchungsdauer passt', and 'Gewinnspielteilnahme'. There is also a 'Sonstiges:' (Other:) field with a text input box.

Abbildung 8: Erhebung der Gründe für die Teilnahme und wie Versuchspersonen auf die Untersuchung aufmerksam wurden bei der Untersuchung an der Universität

*** Wie sind Sie auf diese Untersuchung aufmerksam geworden?**

Bitte wählen Sie einen oder mehrere Punkte aus der Liste aus.

<input type="checkbox"/> Mail von student.umfrage
<input type="checkbox"/> Aushang an der Universität
<input type="checkbox"/> Flyer bekommen
<input type="checkbox"/> Link zugesendet bekommen
<input type="checkbox"/> In sozialen Medien gesehen (Facebook, Instagram,...)
<input type="checkbox"/> In einer Vorlesung davon gehört
<input type="checkbox"/> Von FreundInnen/StudienkollegInnen davon gehört
<input type="checkbox"/> Kenne den Versuchsleiter
Sonstiges: <input type="text"/>

*** Was hat Sie zur Teilnahme an dieser Untersuchung bewogen?**

Bitte wählen Sie einen oder mehrere Punkte aus der Liste aus.

<input type="checkbox"/> Interesse
<input type="checkbox"/> Versuchschein benötigt
<input type="checkbox"/> Onlineuntersuchung
<input type="checkbox"/> Langeweile
<input type="checkbox"/> Untersuchungsdauer passt
<input type="checkbox"/> Gewinnspielteilnahme
<input type="checkbox"/> Kenne den Versuchsleiter
Sonstiges: <input type="text"/>

Abbildung 9: Erhebung der Gründe für die Teilnahme und wie Versuchspersonen auf die Untersuchung aufmerksam wurden bei der Online-Umfrage

*** Wie sind Sie auf diese Untersuchung aufmerksam geworden?**
 Bitte wählen Sie einen oder mehrere Punkte aus der Liste aus.

<input type="checkbox"/> Mail von student.umfrage
<input type="checkbox"/> Aushang an der Universität
<input type="checkbox"/> Über SurveyCircle
<input type="checkbox"/> Flyer bekommen
<input type="checkbox"/> Link zugesendet bekommen
<input type="checkbox"/> In sozialen Medien gesehen (Facebook, Instagram,...)
<input type="checkbox"/> In einer Vorlesung davon gehört
<input type="checkbox"/> Von FreundInnen/StudienkollegInnen davon gehört
<input type="checkbox"/> Kenne den Versuchsleiter
Sonstiges: <input type="text"/>

*** Was hat Sie zur Teilnahme an dieser Untersuchung bewogen?**
 Bitte wählen Sie einen oder mehrere Punkte aus der Liste aus.

<input type="checkbox"/> Interesse
<input type="checkbox"/> Onlineuntersuchung
<input type="checkbox"/> Punkte für SurveyCircle
<input type="checkbox"/> Langeweile
<input type="checkbox"/> Untersuchungsdauer passt
<input type="checkbox"/> Gewinnspielteilnahme
<input type="checkbox"/> Kenne den Versuchsleiter
Sonstiges: <input type="text"/>

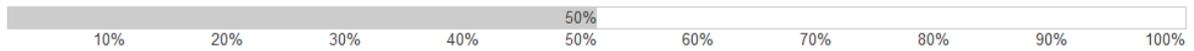
Abbildung 10: Erhebung der Gründe für die Teilnahme und wie Versuchspersonen auf die Untersuchung aufmerksam wurden bei der Umfrage über „SurveyCircle“

Fragen zum Ausfall von Versuchspersonen bei der Hauptuntersuchung



Umfrage im Rahmen des Studiums

Forschen Erforschen - online



Bearbeitungsfortschritt

Vorangegangene Untersuchungen

* Sind Sie schon einmal trotz Anmeldung nicht zu einer wissenschaftlichen psychologischen Untersuchung erschienen?

Ja
 Nein

* Wie viele Male sind Sie bereits zu psychologischen Untersuchungen nach Anmeldung nicht erschienen?
In dieses Feld dürfen nur Zahlen eingegeben werden.

Welchen Grund / welche Gründe gab es, dass Sie nicht zur Untersuchung / zu den Untersuchungen erschienen sind?

* Haben Sie sich schon einmal von einer wissenschaftlichen psychologischen Untersuchung wieder abgemeldet?

Ja
 Nein

* Wie viele Male haben Sie sich bereits von Untersuchungen nach Anmeldung wieder abgemeldet?
In dieses Feld dürfen nur Zahlen eingegeben werden.

Welchen Grund / welche Gründe gab es, dass Sie sich von der Untersuchung / den Untersuchungen wieder abgemeldet haben?

Abbildung 11: Fragen zum Ausfall und zur Abmeldung von Versuchspersonen bei wissenschaftlichen Untersuchungen

Fragebogen zu den Teilnahmewahrscheinlichkeiten bei der Hauptuntersuchung

Motivationale Faktoren

*

Inwiefern beeinflussen die folgenden Bedingungen die **Wahrscheinlichkeit**, dass Sie zu einem Versuchstermin, zu dem Sie sich angemeldet haben, **tatsächlich erscheinen**?
Geben Sie bitte an, wie sich die Wahrscheinlichkeit dafür ändern würde.

Sollte eine Bedingung eine Anmeldung zur Untersuchung Ihrerseits mit Sicherheit ausschließen, so wählen Sie bitte das Feld "würde ich mich von vornherein nicht anmelden" aus.

	sinkt die Wahrscheinlichkeit stark	sinkt die Wahrscheinlichkeit ein wenig	bleibt die Wahrscheinlichkeit gleich	erhöht sich die Wahrscheinlichkeit ein wenig	erhöht sich die Wahrscheinlichkeit stark	würde mich von vornherein nicht anmelden
Wenn draußen besonders schönes Wetter ist...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn ein/e Freund/in mich vor dem Untersuchungstermin spontan um Hilfe bittet...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn ich 10€/Stunde für die Teilnahme erhalte...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn ich am Vortag eine Erinnerung an den Termin erhalte...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn die Untersuchung mehr als 1 1/2 Stunden dauert...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn ich bei der Untersuchung einen Leistungstest bearbeiten soll...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn die Untersuchung vor 8:30 Uhr startet...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn ich vor der Untersuchung bereits per Mail oder SMS mit dem/r Versuchsleiter/in kommuniziert habe...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn die Untersuchung von einem/r Freund/in geleitet wird...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn ich 5€ für die Teilnahme erhalte...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Abbildung 12: Teil 1 des Fragebogens zu den Teilnahmewahrscheinlichkeiten

	sinkt die Wahrscheinlichkeit stark	sinkt die Wahrscheinlichkeit ein wenig	bleibt die Wahrscheinlichkeit gleich	erhöht sich die Wahrscheinlichkeit ein wenig	erhöht sich die Wahrscheinlichkeit stark	Ich würde mich nicht anmelden
Wenn ich das Thema der Untersuchung nicht interessant finde...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn ich in der Ausschreibung zur Untersuchung mit „Sie“ angesprochen wurde...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn der Untersuchungstermin frühestens einen Monat nach der Anmeldung stattfindet...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn mein/e Chef/in mich (kurz) vor dem Untersuchungstermin bittet bei der Arbeit einzuspringen...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn draußen leicht bewölktes Wetter ist...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn ich an nach der Untersuchungsteilnahme bei einem Gewinnspiel mitmachen darf...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn die Untersuchung weniger als 30 Minuten dauert...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn bei der Untersuchung aufwendige technische Hilfsmittel zum Einsatz kommen (EEG, fMRT, etc.)...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn ich den Untersuchungstermin frei wählen kann...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Abbildung 13: Teil 2 des Fragebogens zu den Teilnahmewahrscheinlichkeiten

	sinkt die Wahrscheinlichkeit stark	sinkt die Wahrscheinlichkeit ein wenig	bleibt die Wahrscheinlichkeit gleich	erhöht sich die Wahrscheinlichkeit ein wenig	erhöht sich die Wahrscheinlichkeit stark	Ich würde mich nicht anmelden
Wenn die Untersuchung von einem/r Fremden geleitet wird...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn ich das Thema der Untersuchung besonders interessant finde...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn ich nach einer Voruntersuchung zur Hauptuntersuchung eingeladen werde...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn ich von dem/der Versuchsleiter/in keine Erinnerung an den Termin erhalte...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn die Untersuchung zwischen 9:00 Uhr und 17:00 Uhr startet...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn ich in der Ausschreibung zur Untersuchung mit „Du“ angesprochen werde...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn ich den/die Versuchsleiter/in vor der Untersuchung bereits persönlich getroffen und mit ihr/ihm gesprochen habe...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn ich 30€ für die Teilnahme an der Untersuchung erhalte...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn ich mich nicht an den Termin erinnere...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn ich bei der Untersuchung Fragebögen bearbeiten soll...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Abbildung 14: Teil 3 des Fragebogens zu den Teilnahmewahrscheinlichkeiten

	sinkt die Wahrscheinlichkeit stark	sinkt die Wahrscheinlichkeit ein wenig	bleibt die Wahrscheinlichkeit gleich	erhöht sich die Wahrscheinlichkeit ein wenig	erhöht sich die Wahrscheinlichkeit stark	Ich würde mich nicht anmelden
Wenn die Untersuchung aus mehreren Terminen besteht...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn ein Familienmitglied mich vor dem Untersuchungstermin um Hilfe bittet...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn ich Süßigkeiten für die Teilnahme bekomme...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn die Untersuchung zwischen 30 und 90 Minuten dauert...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn ich den Untersuchungstermin ohne persönlichen Kontakt mit dem/der Versuchsleiter/in vereinbart haben... (Z.B. durch einen Eintrag in eine Terminliste)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn draußen besonders schlechtes Wetter ist...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn ich mit einem/r Freund/in oder meinem/r Partner/in gemeinsam zur Untersuchung erscheinen soll...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn bei der Untersuchung meine aktive Teilnahme besonders gefordert ist...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn die Untersuchung nach 18:30 Uhr startet...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

* Für wie wahrscheinlich halten Sie es, dass Sie einen Versuchstermin verschlafen, vergessen oder aus anderen Gründen übersehen?

(1 = nicht wahrscheinlich; 5 = sehr wahrscheinlich)

1

2

3

4

5