

MASTERARBEIT

Studentische Inhalte in institutionellen Repositorien: Welche Faktoren beeinflussen die Bereitschaft zur Selbstarchivierung?

vorgelegt im Mai 2025 von
Silja Wiegmann

1. Prüfer: Prof. Dr. Hardy Gundlach
 2. Prüfer: Prof. Dr. Ulrich Hofmann
-

**HOCHSCHULE FÜR ANGEWANDTE
WISSENSCHAFTEN HAMBURG**

Department Information und Medienkommunikation

Studiengang Digitale Transformation der Informations- und Medienwirtschaft

**HOCHSCHULE FÜR ANGEWANDTE
WISSENSCHAFTEN HAMBURG**
Hamburg University of Applied Sciences

**Studentische Inhalte in institutionellen
Repositorien: Welche Faktoren beeinflussen die
Bereitschaft zur Selbstarchivierung?**

Masterarbeit vorgelegt von
Silja Wiegmann

Abstract

Hochschulbibliotheken sehen sich mit geringen Partizipationsraten bei der Produktion von Inhalten für institutionelle Repositorien (IR) konfrontiert. Eine lange vernachlässigte Gruppe potenzieller Stakeholder sind Studierende. In dieser Arbeit wird der Frage nachgegangen, welche Faktoren die Bereitschaft von Studierenden zur Selbstarchivierung in IR beeinflussen. Hierfür wird zunächst eine literaturbasierte empirische Modellierung der studentischen Selbstarchivierungsbereitschaft in Abhängigkeit von Gewinn-, Kosten- und Kontextfaktoren vorgenommen. Anschließend werden die Hypothesen mittels einer linearen Regressionsanalyse nach der Methode der kleinsten Quadrate empirisch überprüft, indem eine Stichprobe ($n=94$) der genannten Zielgruppe im Rahmen eines Online-Fragebogens befragt wird. Im Ergebnis werden zwei signifikante Gewinnfaktoren (Streben nach persönlicher Anerkennung sowie das Unterstützen einer Open-Access-Philosophie) und ein signifikanter Kostenfaktor (wahrgenommener Zeit- und Arbeitsaufwand) identifiziert, aus denen Handlungsempfehlungen für Hochschulbibliotheken zur Akquise studentischer Autor:innen abgeleitet werden. Hierzu zählen das Veranstellen von Preisverleihungen und Schulungsformaten, die Ernennung von Verbindungsbibliothekar:innen sowie das Angebot eines Upload-Services.

Schlagwörter: Wissenschaftskommunikation, Open Access, Institutionelle Repositorien, Selbstarchivierung, Studentische Beteiligung, Regressionsanalyse, Methode der kleinsten Quadrate

Abstract (English)

University libraries are confronted with low participation rates in the production of content for institutional repositories (IR). Students are a long-neglected group of potential stakeholders. This paper explores the question of which factors influence students' willingness to self-archive in IRs. To this end, a literature-based empirical modeling of students' willingness to self-archive is carried out as a function of benefit, cost and contextual factors. The hypotheses are then empirically tested by conducting an online survey among the target group ($n=94$) and applying ordinary least squares linear regression. As a result, two significant benefit factors (personal recognition and supporting a philosophy of open access) and one significant cost factor (perceived time and effort) are identified, from which recommended actions for university libraries to recruit student authors are derived. These include organizing award ceremonies and training formats, appointing liaison librarians and offering an upload service.

Keywords: Scholarly communication, Open Access, Institutional repository, Self-archiving, Student contribution, Regression analysis, Ordinary least squares

Inhaltsverzeichnis

Abstract	i
Abbildungsverzeichnis.....	iv
Tabellenverzeichnis	v
1 Einleitung: Die digitale Transformation des wissenschaftlichen Publikationswesens durch Open Access	1
2 Problemstellung: Studentische Inhalte in institutionellen Repositorien ..	5
2.1 Beschaffenheit und Zielsetzung von institutionellen Repositorien	6
2.2 Studierende als Stakeholder bei institutionellen Repositorien	12
2.2.1 Absteckung des Sammlungsbereiches.....	13
2.2.2 Vor- und Nachteile der Selbstarchivierung für Studierende	14
3 Stand der Forschung zur Publikationsbereitschaft von Studierenden in institutionellen Repositorien.....	18
4 Forschungsfrage und Zielsetzung.....	22
5 Modellierung der Bereitschaft zur Selbstarchivierung	23
5.1 Theoretischer Hintergrund: <i>Socio-Technical Network Model</i> (STIN) und Austauschtheorie (<i>Social Exchange Theory</i>)	23
5.2 Entwicklung und kritische Einordnung des Kim-Modells.....	26
5.2.1 Pilotstudie (2007).....	26
5.2.2 Nachfolgestudie (2010).....	29
5.2.3 Abschlussstudie (2011)	34
5.3 Auswahl der Variablen	37
5.3.1 Gewinnfaktoren.....	38
(a) Altruismus.....	38
(b) Schulischer Fleiß.....	39
(c) Karrierestreben	40
(d) Persönliche Anerkennung	41
(e) Vertrauenswürdigkeit	41
5.3.2 Kostenfaktoren.....	42
(a) Zeit- und Aufwandsinvestition	43
(b) Plagiatsfurcht.....	43
5.3.3 Kontextfaktoren	43
(a) Identifikation.....	44

5.3.4	Individualfaktoren	45
5.3.5	Abhängige Variable	45
6	Hypothesen und empirisches Modell.....	46
7	Operationalisierung der Konstrukte.....	48
7.1	Frageblock 1: Anerkennung.....	49
7.2	Frageblock 2: Hochschule	53
7.3	Frageblock 3: Upload-Prozess	55
7.4	Frageblock 4: Selbstarchivierungsbereitschaft	58
7.5	Frageblock 5: Demografische Daten.....	60
8	Empirische Studie	61
8.1	Konstruktion des Fragebogens und qualitativer Pretest	62
8.2	Teilnehmendenakquise.....	63
9	Datenanalyse	67
9.1	Reliabilitätsanalyse: Spearman-Brown-Koeffizient und Cronbachs Alpha	67
9.2	Deskriptive Statistik.....	69
9.2.1	Stichprobenkontrolle	69
9.2.2	Statistische Maßzahlen der Verteilung.....	74
9.3	Regressionsanalyse	76
9.3.1	Einfache lineare Regression: Überprüfung der Einzelhypothesen	76
9.3.2	Moderationsanalyse zu H7b.....	82
9.3.3	Multiple lineare Regression: Überprüfung des Gesamtmodells.....	82
9.4	Ergebnisse	85
10	Diskussion und Limitationen	87
11	Handlungsempfehlungen	97
12	Desiderata	100
13	Fazit	102
14	Literaturverzeichnis	104
Anhang A: Fragebogen		115
Anhang B: R-Notebook.....		123

Abbildungsverzeichnis

Abb. 1	Version 1 des Kim-Modells	27
Abb. 2	Version 2 des Kim-Modells	30
Abb. 3	Version 3 des Kim-Modells	35
Abb. 4	Modell der Selbstarchivierungsbereitschaft von Studierenden in IR	48
Abb. 5	Mockup des Upload-Formulars	56
Abb. 6	Rekrutierungsflyer für die Fachbibliothek DMI	65
Abb. 7	Rekrutierungsflyer für die Fachbereichsbibliothek Kulturwissenschaften	66
Abb. 8	Anteile der Geschlechter in % ($n=94$)	70
Abb. 9	Altersverteilung ($n=92$)	71
Abb. 10	Anteile der angestrebten Studienabschlüsse in % ($n=94$)	73
Abb. 11	Verteilung der Variable <i>prestige</i> ($n=94$)	75
Abb. 12	Einfache lineare Regression der wahrgenommenen Zeit- und Aufwandsinvestition	77
Abb. 13	Einfache lineare Regression der Plagiatsfurcht	77
Abb. 14	Einfache lineare Regression des Altruismus	78
Abb. 15	Einfache lineare Regression des Wertes auf schulische Leistung	78
Abb. 16	Einfache lineare Regression des Strebens nach persönlicher Anerkennung	79
Abb. 17	Einfache lineare Regression des Vertrauens in qualitätssichernde Prozesse	79
Abb. 18	Einfache lineare Regression des Karrierestrebens im Studium	80
Abb. 19	Einfache lineare Regression der Identifikation mit der Hochschule	80
Abb. 20	Einfache lineare Regression des ‚OA-Spirits‘	81
Abb. 21	Heteroskedastizität des Gesamtmodells	84
Abb. 22	Selbstarchivierungsbereitschaft nach Prestigestreben	89
Abb. 23	Selbstarchivierungsbereitschaft nach ‚OA-Spirit‘	91
Abb. 24	Selbstarchivierungsbereitschaft nach Aufwandswahrnehmung	94

Tabellenverzeichnis

Tab. 1	Zur Selbstarchivierung motivierende Faktoren von Studierenden der Loughborough University	19
Tab. 2	Von der Selbstarchivierung abhaltende Faktoren von Studierenden der Loughborough University	20
Tab. 3	Relevante Faktoren aus Kims Pilotstudie	28
Tab. 4	Signifikante Faktoren aus Kims Nachfolgestudie	32
Tab. 5	Signifikante Faktoren aus Kims Abschlussstudie	36
Tab. 6	Hypothesen	46
Tab. 7	Operationalisierung der persönlichen Anerkennung	51
Tab. 8	Operationalisierung des Karrierestrebens	52
Tab. 9	Operationalisierung des schulischen Fleißes	53
Tab. 10	Operationalisierung der Identifikation	54
Tab. 11	Operationalisierung der Vertrauenswürdigkeit	54
Tab. 12	Operationalisierung der Zeit- und Aufwandsinvestition	57
Tab. 13	Operationalisierung der Plagiatsfurcht	58
Tab. 14	Operationalisierung des Altruismus	59
Tab. 15	Operationalisierung der Selbstarchivierungsbereitschaft	60
Tab. 16	Operationalisierung der wissenschaftlichen Karriere	60
Tab. 17	Operationalisierung der demografischen Daten	61
Tab. 18	Ergebnisse der Reliabilitätsanalyse mittels Cronbachs Alpha	68
Tab. 19	Ergebnisse der Reliabilitätsanalyse mittels des Spearman-Brown-Koeffizienten	68
Tab. 20	Vertretene Studienfächer ($n=92$)	72
Tab. 21	Mediane und Schiefen der Variablen	74
Tab. 22	Ergebnisse der einfachen linearen Regressionen	81
Tab. 23	Ergebnisse der Moderationsanalyse zu H7b	82
Tab. 24	Ergebnisse der Multikollinearitätsdiagnose	83
Tab. 25	Ergebnisse der multiplen linearen Regression	84

1 Einleitung: Die digitale Transformation des wissenschaftlichen Publikationswesens durch Open Access

Das Internet hat die Art und Weise verändert, auf die Wissen generiert und verbreitet wird. Die kollaborative Anwendung digitaler Tools im wissenschaftlichen Forschungsprozess lässt sich mit dem Begriff *E-Research* beschreiben. Während das E-Präfix traditionell für ‚elektronisch‘ steht (z. B. E-Mail, E-Commerce), nimmt es im vorliegenden Kontext zusätzlich die Bedeutung *enhanced* (d. h. erweitert, verbessert) an (Borgman, 2007, S. 20; Meyer & Schroeder, 2015, S. 219). So ermöglicht die Erweiterung bestehender wissenschaftlicher Kommunikationspraktiken durch technische Komponenten ein digitales Kommunikationssystem, das nicht nur die globale Vernetzung von Forschenden, sondern auch die Verbreitung des verstärkt kollaborativ erzeugten Wissens verbessert (Meyer & Schroeder, 2015, S. 4-5; Taubert & Weingart, 2016, S. 7).

Weil die beschriebenen Veränderungen auf der innovativen Weiterentwicklung von Informations- und Kommunikationstechnologien beruhen, findet an dieser Stelle ein Prozess der *Digitalisierung* statt, die sich durch ihre Unabgeschlossenheit auszeichnet: In Abwesenheit eines Ausgangspunktes oder Zielzustandes veranlassen aufeinanderfolgende Innovationswellen einen steten digitalen Wandel (Taubert & Weingart, 2016, S. 10-11). Dieser Veränderungsdruck, der von technologischen Potenzialen ausgeübt wird (*Technology Push*), führt nachfrageseitig zu neuen Anforderungen (*Market Pull*; Guo et al., 2020, S. 352-353; Picot & Fischer, 2006, S. 7-8). Während digitale Transformationsprozesse im Bereich der Wirtschaft bestehende Geschäftsmodelle infrage stellen (Hüther, 2017, S. 179), verändern sie in der Wissenschaft die Kommunikations- und Forschungspraktiken von Natur-, Sozial- und Geisteswissenschaftler:innen (Meyer & Schroeder, 2015, S. 1).

Wissenschaftliche Kommunikation lässt sich als *soziotechnisches System* begreifen (s. Kap. 5.1), dessen Grundannahme darin besteht, dass Technologie und Gesellschaft eng miteinander verwoben sind und sich gegenseitig beeinflussen. E-Research bringt daher nicht nur eine digitale Infrastruktur mit sich, sondern verändert auch die Alltagsgewohnheiten von Forschenden und die Beziehungen zwischen den Stakeholdern wissenschaftlicher Kommunikation, zu denen neben Autor:innen und Universitäten auch Bibliotheken, Verlage und Stiftungen gehören (Borgman, 2007, S. 38;55; Meyer & Schroeder, 2015, S. 15). Transformiert werden sowohl die Produktionsprozesse und Verbreitungswege als auch die Rezeption von

Wissen in Form von Publikationen (Meyer & Schroeder, 2015, S. 1;6;18-19; Taubert & Weingart, 2016, S. 11-12). Insbesondere der Wandel von wissenschaftlichen Print-Zeitschriften hin zu digitalen Formaten hat eine nachhaltige Veränderung im Lese-, Recherche- und Publikationsverhalten von Wissenschaftler:innen hervorgerufen. Um diese veränderten Nachfragen zu bedienen, durchlaufen Bibliotheken einen strukturellen Wandel, der ihren Fokus von der Pflege eines Lokalbestandes hin zur Bereitstellung virtueller Zugänge verschiebt (Sathe, 2002, S. 242; Schallehn & Schimmer, 2014, S. 311; Smith, 2003, S. 162).

Der Begriff der Verbreitung bezeichnet „die Herstellung von Reichweite innerhalb einer wissenschaftlichen (Kommunikations-)Gemeinschaft“. Erst die Zirkulation von Forschungsergebnissen erlaubt das Ankurbeln weiterer Forschungsprozesse und die Anerkennung von Forschungsleistungen (Taubert & Weingart, 2016, S. 6). Das freie Teilen von Wissen und Ideen fungiert somit als „sozialer Klebstoff“, der Forschungsgemeinschaften zusammenhält (Borgman, 2007, S. 56). Entsprechend groß sind die Auswirkungen, welche die digitale Transformation auf das wissenschaftliche Publikationswesen ausübt (Kim, 2010, S. 1909; Meyer & Schroeder, 2009, S. 218; Swan et al., 2014, S. 2).

Im späten 20. und frühen 21. Jahrhundert wurde der Publikationsmarkt von kommerziellen Verlagen dominiert, die wissenschaftliche Beiträge in fachspezifischen Zeitschriften sammeln. Neben der Qualitätssicherung in Form des Peer-Reviews – d. i. die förmliche Prüfung von Wahrheitsansprüchen durch Fachkolleg:innen – gehören auch das Lektorat und die Formatierung der Texte zu ihren Dienstleistungen. Hierfür treten die Autor:innen in der Regel das ausschließliche Nutzungsrecht ihrer Werke an die Verlage ab. Profit generieren die Verlage durch den Verkauf von Abonnements, die Bibliotheken und Forschenden den Zugang zu den Artikeln ermöglichen. Dieses klassische Publikationsmodell wird auch als *Toll Access* bzw. *Closed Access* bezeichnet (Herb, 2016, S. 163; Söllner, 2017, S. 4; Swan et al., 2014, S. 4; Taubert & Weingart, 2016, S. 5).

Die Transformation der Verbreitungswege wissenschaftlicher Inhalte stellt sich insbesondere in der Veröffentlichung und Archivierung im Open Access dar (Kim, 2010, S. 1909). Open Access bedeutet

freie[n] Zugang zu wissenschaftlichen Publikationen und anderen Materialien im Internet. Ein unter Open-Access-Bedingungen publiziertes wissenschaftliches Dokument kann jede und jeder lesen, herunterladen, speichern, verlinken, drucken und damit entgeltfrei nutzen (DFG, 2022).

Damit ist Open Access wichtiger Bestandteil der Open-Science-Bewegung, deren Kernidee aussagt, dass die freie Zirkulation von Forschungsergebnissen nicht nur den auf diesen aufbauenden Forschenden, sondern auch der Gesamtgesellschaft zugutekommt, welche Einsicht in den wissenschaftlichen Kommunikationsfluss erhält (Borgman, 2007, S. 65; Miedema, 2022, S. 186).

Aus der Abwesenheit von Paywalls, die sich durch den Erwerb von Abonnements überwinden lassen, ergibt sich die Notwendigkeit neuer Geschäftsmodelle. Finanziert wird die Veröffentlichung im Open Access durch die Erhebung von Publikationsgebühren bzw. *Article Processing Charges (APCs)*, die entweder von den Autor:innen selbst oder den sie beschäftigenden Forschungsorganisationen beglichen werden (Taubert & Weingart, 2016, S. 20). Indem die Autor:innen den Verlagen lediglich ein einfaches Nutzungsrecht einräumen, steht es ihnen frei, ihre Werke unter eine freie Lizenz (zumeist *Creative Commons*) zu stellen, welche die breite Zugänglichkeit und Wiederverwendbarkeit der Inhalte gewährleistet und definiert (Carroll, 2011, S. 2; Deppe & Beucke, 2017, S. 16; Söllner, 2017, S. 4-5).

Die Verbreitung wissenschaftlicher Inhalte wird somit einerseits durch E-Research begünstigt, auf der anderen Seite jedoch auch durch Toll-Access-Geschäftsmodelle eingeschränkt (Borgman, 2007, S. 36;109). Weil Verlage als Medienunternehmen dazu angehalten sind, neben den Kommunikationsinteressen der Wissenschaft auch ökonomische Gesichtspunkte miteinzubeziehen, wird der Zugang zu qualitätsgesicherten Publikationen vor dem Hintergrund der Profitgenerierung einer künstlichen Verknappung unterworfen (Bergstrom, 2001, S. 191; Schimmer & Geschuhn, 2017, S. 174; Taubert & Weingart, 2016, S. 9). Der Begriff der Künstlichkeit unterstreicht dabei, dass diese Knappheit nicht auf die Natur des Gutes zurückzuführen ist: Die Immaterialität von wissenschaftlichen Informationen bedeutet, dass diese endlos und sogar gleichzeitig genutzt werden können, ohne sich zu verbrauchen oder an Wert zu verlieren. Diese Kriterien der Nicht-Ausschließbarkeit von Konsument:innen und der Nicht-Rivalität im Konsum sind hinreichend dafür, Informationen unter die *öffentlichen Güter* zu subsumieren. Gleichzeitig spielt bei öffentlichen Gütern, deren Bereitstellung für privatwirtschaftliche Anbietende häufig nicht profitabel ist, auch der Aspekt einer öffentlichen Be-

reitstellung durch den Staat und entsprechender Finanzierung aus Steuergeldern eine Rolle. Anstatt finanzieller Interessen steht hier die öffentliche Verfügbarkeit und Förderung des Angebotes im Vordergrund (Borgman, 2007, S. 109;193; Dewenter & Rösch, 2015, S. 15-16).

Open Science basiert auf der Annahme, dass wissenschaftliche Informationen gemäß ihrer Natur als öffentliches Gut behandelt und der Zugang zu ihnen nicht aus profitorientierten Motiven eingeschränkt werden sollte (Borgman, 2007, S. 35; Madalli, 2015, S. 5). Eine wichtige Argumentationsgrundlage hierfür ist die Ökonomisierung im Verlagswesen, die mit der Bildung eines Angebotsoligopols der Großverlage Elsevier, Springer und Wiley-Blackwell einhergegangen ist. Seit den 1990er-Jahren erschweren überinflationäre Preissteigerungen den Bibliotheken eine adäquate Literaturversorgung der Forschenden, was mit dem Begriff der *Zeitschriftenkrise* beschrieben wird (Deppe & Beucke, 2017, S. 13; Söllner, 2017, S. 6; Swan et al., 2014, S. 4; Taubert & Weingart, 2016, S. 12). Durch Paywalls wird das ursprünglich öffentliche Gut wissenschaftlicher Informationen zu einem *Klubgut*, welches sich durch die Ausschließbarkeit von Konsument:innen und die Nicht-Rivalität im Konsum auszeichnet (Dewenter & Rösch, 2015, S. 16).

Gemäß der eingangs erläuterten Dynamik digitaler Transformationsprozesse wird die Veränderung des Publikationsverhaltens von Forschenden durch ein Zusammenspiel neuer Technologien mit neuen Nachfragen vorangetrieben: Angesichts der von den Verlagen vorgenommenen Beschränkung der Verbreitungswege wissenschaftlicher Inhalte nutzen Forschende verstärkt die Möglichkeiten zur Selbstverbreitung, die sich durch digitale Technologien eröffnen. So fungierte die Zeitschriftenkrise als Katalysator der sich um die Jahrtausendwende formierenden Open-Access-Bewegung (Borgman, 2007, S. 76;101; Deppe & Beucke, 2017, S. 13-14; Taubert & Weingart, 2016, S. 12).

Internationale Vertreter:innen aus Wissenschaft, Wissenschaftsorganisationen und Verlagen entwickelten die programmatischen Ziele von Open Access im Rahmen dreier Konferenzen, der *Budapest Open Access Initiative* (BOAI) vom Februar 2002, dem *Bethesda Statement on Open Access Publishing* vom Juni 2003 und der *Berlin Declaration on Open Access to Knowledge* vom Oktober 2003 (Ball, 2016, S. 180; Schallehn & Schimmer, 2014, S. 312). Die als Geburtsstunde der Open-Access-Bewegung geltende BOAI prägte den Open-Access-Begriff und benannte zwei Strategien zu seiner praktischen Umsetzung, die später als grüner bzw. goldener Weg zu Open Access bezeichnet wurden. Während der goldene Weg die Erstveröffentli-

chung einer Publikation im Open Access meint, beschreibt der grüne Weg die Praxis, einen in einer Toll-Access-Zeitschrift veröffentlichten Artikel zusätzlich in einem frei zugänglichen digitalen Repositorium abzulegen (BOAI, 2002; Deppe & Beucke, 2017, S. 14; Schallehn & Schimmer, 2014, S. 312-313;316). Die Bethesda-Erklärung differenzierte diese Definition aus, indem sie eine passive Rezeption von der Möglichkeit schöpferischer Komponenten unterschied, was später mit dem Begriffspaar ‚gratis‘ bzw. ‚libre‘ beschrieben wurde. Gratis-Publikationen bieten einen freien lesenden Zugriff, wohingegen Libre-Publikationen darüber hinaus auch Weiterbearbeitungsrechte gewähren (Brown et al., 2003; Schallehn & Schimmer, 2014, S. 313; Swan et al., 2014, S. 6). In der Berliner Erklärung wurde diese erweiterte Definition auch auf Forschungs- und Metadaten sowie multimediale Formate übertragen (Deppe & Beucke, 2017, S. 14; Max-Planck-Gesellschaft, o. J.).

Eine mögliche Publikationsplattform offener Inhalte stellen institutionelle Repositorien dar, die im Fokus des folgenden Kapitels stehen. Der potenzielle Sammlungsbereich dieser Plattformen ist breit gefächert und kann u. a. graue Literatur beinhalten, zu der auch studentische Arbeiten gehören, die im Kontext von Lehrveranstaltungen entstanden sind. Mit dem Ziel, Handlungsempfehlungen für Hochschulbibliotheken zur Akquise studentischer Autor:innen zu formulieren, geht diese Arbeit der Frage nach, welche Faktoren die Bereitschaft von Studierenden zur Selbstarchivierung in institutionellen Repositorien beeinflussen. Ausgehend von dem empirischen Modell der Informationswissenschaftlerin Kim (2007; 2010; 2011) zur Publikationsbereitschaft wissenschaftlichen Personals in institutionellen Repositorien wird dabei die Hypothese zugrunde gelegt, dass sich anhand von Gewinn-, Kosten- und Kontextfaktoren die Selbstarchivierungsbereitschaft der Studierenden erklären lässt.

2 Problemstellung: Studentische Inhalte in institutionellen Repositorien

Obwohl sich die Open-Access-Bewegung traditionell darauf fokussiert, begutachtete Zeitschriftenartikel digital verfügbar zu machen, entsteht im Zuge der Forschung und Lehre eine weite Bandbreite an digitalen Objekten, die für Forschungsarbeiten von Interesse sein können. Hierzu gehören u. a. Konferenzschriften, Präsentationen, Abschlussarbeiten oder andere Lehr- und Lernmaterialien. Weil derartige Publikationen nicht auf dem Buchmarkt erhältlich sind, fallen sie in das Spektrum

der *grauen Literatur*, die vorrangig von Institutionen, wie z. B. Forschungseinrichtungen oder Hochschulen, herausgegeben wird. Als Veröffentlichungsplattform dieser Inhalte dienen oftmals Repositorien (Bashir et al., 2022, S. 524; Borgman, 2007, S. 238; Gantert, 2016, S. 78-79; SUB Göttingen, 2022), definiert als „an Universitäten oder Forschungseinrichtungen betriebene Dokumentenserver, auf denen wissenschaftliche Materialien archiviert und weltweit entgeltfrei zugänglich gemacht werden“ (SUB Göttingen, 2022). So erhöhen Repositorien die Sichtbarkeit wissenschaftlicher Inhalte, die von kommerziellen Verlagen keine Aufmerksamkeit erfahren (Bankier & Smith, 2010, S. 256; Bashir et al., 2022, S. 524-525).

Die Entscheidung über die dargebotenen Inhalte fällt für jeden Dokumentenserver anders aus. Während einige Repositorien jegliche Art von grauer Literatur auf goldenem Wege publizieren, beschränken sich andere auf die grüne Zweitveröffentlichung von begutachteten Zeitschriftenartikeln (Bankier & Smith, 2010, S. 245; SUB Göttingen, 2022). Erstere bieten Inhalten eine Plattform, die ohne Repositorien mit breiten Publikationsrichtlinien weitgehend unsichtbar bleiben würden. Hierzu zählen u. a. die Arbeiten von Studierenden, die im Rahmen der Lehre oder studentischen Forschung entstanden sind (Bankier & Smith, 2010, S. 253;256; Demetres et al., 2020, S. 177). Welche Relevanz die Publikation von studentischen Inhalten auf dem goldenen Weg des Open Access birgt, wird im Verlauf dieses Kapitels dargestellt. Hierfür wird zunächst die Beschaffenheit und Zielsetzung von institutionellen Repositorien definiert und von anderen Repositoriumsarten abgegrenzt. Nachdem die Vorteile von breiten Publikationsrichtlinien aufgezeigt wurden, wird anschließend erläutert, welchen Platz studentische Inhalte in institutionellen Repositorien einnehmen können.

2.1 Beschaffenheit und Zielsetzung von institutionellen Repositorien

Open Access hat die Art und Weise transformiert, auf die wissenschaftliche Literatur und Forschungsergebnisse genutzt werden können. Die Vorteile der Digitalisierung werden insbesondere in Repositorien offenbar, die Bibliothekssoftware mit der Möglichkeit des direkten Publizierens verknüpfen. Die Masse an digital aggregierten Daten und Inhalten lässt eine Reihe an Mehrzwecken entstehen, die von Data Mining über Zeitreihenanalysen bis hin zur kooperativen Inhaltserschließung mittels Crosskonkordanzen reichen (Borgman, 2007, S. 104; Salganik, 2018, S. 2). Grob lässt sich zwischen drei Arten von Repositorien unterscheiden: Während (a) *allgemeine Repositorien* Dokumente verschiedener Fächer sammeln (z. B. Zeno-

do), stellen (b) *disziplinäre Repositorien* Dokumente einer bestimmten Fachrichtung zusammen (z. B. PubMed für die Lebenswissenschaften). Beide Repositoriumsarten sind institutionsübergreifend und publizieren somit die Dokumente von Forschenden jeglicher institutionellen Affiliation (Harris & Issakson, 2019; SUB Göttingen, 2022). Diese Eigenschaft unterscheidet allgemeine und disziplinäre Repositorien von institutionellen Repositorien (IR), die im Folgenden näher beschrieben und definiert werden.

Mit der Veröffentlichung der freien Repositoriumssoftware *DSpace* im Jahr 2002 am *Massachusetts Institute of Technology* (MIT) war die für das Betreiben von IR erforderliche Infrastruktur erstmals weiträumig verfügbar. Nach dem Vorbild des MIT richteten weitere renommierte Universitäten, darunter die *University of California*, institutionsgebundene Dokumentenserver ein, wodurch sich diese rasch als wertvolles Element der digitalen Wissenschaftskommunikation etablierten (Kim, 2007, S. [1]; Ware, 2004, S. 115). Wegweisend für diese Entwicklung war ein Positionspapier von Crow (2002), das von der *Scholarly Publishing and Academic Resources Coalition* (SPARC) herausgegeben wurde und als Grundlagentext zu IR gilt (Borgman, 2007, S. 102; Callicot et al., 2016, xvi; Ware, 2004, S. 116).

Einleitend beschreibt Crow IR im weiteren Sinne als Dokumentenserver, eingerichtet von „any organization that wishes to capture and openly disseminate its intellectual product, thus contributing to scientific/scholarly discourse and benefiting from the resulting organizational visibility“. Hierunter fallen u. a. Behörden, NGOs, Forschungseinrichtungen oder Museen (2002, S. 17). Anschließend verengt Crow seine Sichtweise und definiert IR im engeren Sinne stipulativ als „digital archive[s] of the intellectual product created by the faculty, research staff, and students of an institution and accessible to end users both within and outside of the institution, with few if any barriers to access“ (2002, S. 16). Diese enge Definition hat das Konzept langfristig geprägt, sodass der Begriff IR einen von einer Hochschule oder anderen Forschungseinrichtung betriebenen Dokumentenserver bezeichnet. Die übrigen Repositorien, die im Zuge von Crows einleitender Konzeptbeschreibung Erwähnung finden, erhalten hingegen spezifischere Bezeichnungen, wie z. B. *staatliches Repository* (Bashir et al., 2022, S. 522; Lynch, 2003, S. 328; SUB Göttingen, 2022). In seinem Positionspapier verwendet Crow den Begriff der Institution synonym mit dem Begriff der Universität bzw. Hochschule (s. z. B. 2002, S. 4;6), was bereits durch die bewusste Erwähnung von wissenschaftlichem Personal und Studierenden in der engen Definition festgelegt wird. Dieses synonyme Begriffsverständnis wird in dieser Arbeit übernommen.

Nach Crow verfügen IR über folgende vier Eigenschaften (2002, S. 16-19):

- (a) *Institutionell definiert*: IR präsentieren den intellektuellen Output, der von sämtlichen Mitgliedern einer Hochschule generiert wird.
- (b) *Akademische Inhalte*: IR enthalten jegliche Materialien, welche die Hochschule für eine digitale Aufbewahrung vorgesehen hat. Je nach Zielsetzung können diese von studentischen Portfolios über Fotografien bis hin zu Jahresberichten der Hochschule reichen.
- (c) *Anwachsend und fortwährend*: IR gewährleisten sowohl für die Mitglieder der Hochschule als auch für Außenstehende eine digitale Langzeitarchivierung der gesammelten Objekte, welche nur in Ausnahmefällen wieder entfernt werden können.
- (d) *Offen und interoperabel*: Um die intellektuellen Erzeugnisse der Hochschule in der Forschungsgemeinschaft zu verbreiten, unterliegen die Inhalte von IR keinerlei Zugangsbeschränkungen. Eine leichte Auffindbarkeit wird u. a. durch das Bereitstellen von Metadaten gewährleistet.

Für Crow dienen IR zwei unterschiedlichen, sich dabei jedoch ergänzenden Zwecken: (i) Auf *institutionsübergreifender Ebene* treibt ihre Etablierung die Transformation des wissenschaftlichen Publikationswesens von monopolartigen Strukturen hin zu einer Disaggregation des Marktes voran, was angesichts der Zeitschriftenkrise (s. Kap. 1) zu einer finanziellen Entlastung der Hochschulbibliotheken führt. Die so gewonnene Entscheidungshoheit über den Zugang zu den eigenen Forschungsergebnissen lässt die Institutionen selbst an Relevanz gewinnen, was gleichzeitig den interuniversitären Wettbewerb ankurbelt. Vor diesem Hintergrund dienen IR (ii) auf *institutionsspezifischer Ebene* dem Zweck, die Qualität der sie betreibenden Hochschule zu demonstrieren. Inhalte von wissenschaftlicher, gesellschaftlicher und ökonomischer Bedeutung erhöhen die Reputation einer Institution, was sich konkret in einem Anstieg an öffentlichen und privaten Fördergeldern niederschlagen kann (Crow, 2002, S. 2;4;6-7).

Wie Crows Definition konnte sich auch sein Verständnis einer dualen Funktion von IR langfristig durchsetzen. So beschreiben Bankier und Smith IR nicht nur als Antwort auf die Zeitschriftenkrise, sondern auch als Schaufenster der intellektuellen Arbeit einer Institution (2010, S. 245). Ein weiteres Jahrzehnt später greifen auch Bashier et al. dieses Verständnis auf, indem sie IR einerseits aufgrund der

Abwesenheit von Zugangsbeschränkungen eine demokratisierende Funktion zuzusprechen und andererseits die Langzeitarchivierung des universitären Outputs mit einem Zuwachs an Prestige für die Institution verknüpfen (2022, S. 518).

Gleichzeitig besteht in der Literatur kein Konsens darüber, ob Crows institutionenübergreifende Funktion als Primärfunktion von IR gehandelt werden sollte. Abgelehnt wird diese Sichtweise von Lynch (2003), der als geschäftsführender Direktor der *Coalition for Networked Information* (CNI) den neben Crow wirkungsmächtigsten Grundlagentext zu IR verfasste (Plutchak & Moore, 2017, S. 28). Anstatt zugunsten der Open-Access-Transformation mit kommerziellen Verlagen in Konkurrenz zu treten, liegt die eigentliche Verantwortung von IR für Lynch darin, die wissenschaftlichen Aktivitäten der Hochschulen langfristig zu dokumentieren und öffentlich zugänglich zu machen (2003, S. 328-329; 2016, xi). Diese Sichtweise steht nicht in Widerspruch zu Crow, sondern kommt vielmehr einer Schwerpunktsetzung auf Crows institutionsspezifische Funktion gleich. Während Crows Priorität auf einer Transformation des wissenschaftlichen Publikationsmarktes liegt, betrachtet Lynch IR primär als ergänzende Publikationsplattformen für jene Inhalte, die von traditionellen Verlagen keine Beachtung finden (Plutchak & Moore, 2017, S. 29). Lynchs reduzierte Formulierung der Zielsetzung von IR ist hinreichend für die vorliegende Arbeit und wird daher an dieser Stelle übernommen.

Die Einrichtung von IR birgt nicht nur Vorteile für Hochschulen und Autor:innen, sondern auch für die betreuenden Bibliotheken sowie die Lehre. Für Hochschulen fungieren IR als Hochschulbibliografie, die einerseits die Forschungsaktivitäten der Institution nach außen hin für potenzielle Studierende, Lehrende und Geldgeber:innen präsentiert und andererseits ein internes Monitoring des Outputs ermöglicht (Bashir et al., 2022, S. 523; Swan et al., 2014, S. 19). Autor:innen profitieren von einer größeren Reichweite und erhöhten Zitationsraten ihrer Publikationen, was wiederum der Reputation der sie beschäftigenden Hochschule zugutekommt (Demetres et al., 2020, S. 182; Ukwoma & Dike, 2017, S. 19). Lehrende können für die Planung ihrer Veranstaltungen auf archivierte Kursmaterial, wie z. B. Visualisierungen oder Videos, zurückgreifen (Crow, 2002, S. 23-24). Bibliotheken sichern ihren administrativen Platz in der Wissenschaftskommunikation, indem sie über ihr traditionelles Tätigkeitsfeld hinauswachsen und die Autor:innen mit ihrer Fachkompetenz im Dokumenten- und Datenmanagement unterstützen, wodurch das an der Hochschule generierte Wissen geteilt und nutzbar gemacht wird (Crow, 2002, S. 20; Swan et al., 2014, S. 2).

Um die Inhalte von digitalen Bibliotheken zu beschreiben, greift Borgman auf Bucklands (1991, S. 351) Konzept von „information-as-thing“ zurück, definiert als „an object such as a document that is informative or may impart knowledge“ (Borgman, 2007, S. 41). Anhand von Fallstudien identifizieren Bankier und Smith (2010, S. 249;251;253-254) vier Kategorien von Wissensobjekten, die in IR enthalten sein können. Hierzu gehören (a) *Publikationen des wissenschaftlichen Personals*, wie Zeitschriftenartikel oder Forschungsberichte, (b) *administrative Dokumente*, wie Jahres- oder Spendenberichte der Hochschule, die entweder der historischen Dokumentation oder der Mittelbeschaffung dienen, (c) *Inhalte externer Verfasser:innen*, wie Reden oder Konferenzschriften, die unter der Schirmherrschaft der Hochschule entstanden sind, sowie (d) *studentische Inhalte*, wie Abschluss- oder Hausarbeiten.

Lediglich die Kategorien (a) und (d) fallen unter Crows Definition von IR, sodass in der Praxis weitere Inhaltskategorien zutage treten. Welche Inhalte in ein IR gehören, ist dementsprechend umstritten. Borgman vertritt die Position, dass ein zu breiter Gegenstandsbereich von IR zwar den sie betreibenden Bibliotheken nützt, gleichzeitig aber auch die Zielbestimmung von IR verwässert. Die Bibliotheken, so Borgman, sind dazu angehalten, eine weite Bandbreite an Dokumenten zu archivieren, die mit der Hochschule in Verbindung stehen. Hierfür bietet sich die bereits existierende Infrastruktur von IR an, was jedoch von deren eigentlicher Aufgabe ablenkt. Im Gegensatz zu Crow sieht Borgman die Zielsetzung von IR in der direkten Verfügbarmachung von begutachteten Zeitschriftenartikeln (2007, S. 102-103). Damit entspricht Borgmans Verständnis von IR der ursprünglichen Zielsetzung der Open-Access-Bewegung (Kim, 2010, S. 1909).

Neben Crow hinterfragen auch Plutchak und Moore die Effektivität einer derartigen Fokussierung auf den grünen Weg des Open Access, da dieser notwendigerweise von der Existenz traditioneller Zeitschriften abhängt, was der Open-Access-Transformation des wissenschaftlichen Publikationsmarktes langfristig im Wege stehen könnte (2017, S. 30). Gleichzeitig argumentieren Bankier und Smith, dass der enge Ansatz einer Einschränkung auf begutachtete Zeitschriftenartikel in den ersten zehn Jahren nach der Entstehung von IR bereits erprobt wurde und aufgrund der geringen Beteiligung vonseiten des wissenschaftlichen Personals als gescheitert erachtet werden kann: „The singular focus on post-prints“, so die Autor:innen, „was ultimately a failure on the part of the library and the community to recognise and articulate the strengths of the IR in terms that resonated with faculty“ (2010, S. 246). Der Erfolg eines IR hängt maßgeblich davon ab, inwieweit die von

den Bibliotheken dargebotene Möglichkeit zur Selbstarchivierung von den Mitgliedern einer Hochschule ergriffen wird. Nur ein aktiv genutztes und mit Inhalten befülltes IR kann zum festen Bestandteil der intellektuellen Infrastruktur einer Institution werden (Abrizah et al., 2015, S. 731; Kim, 2007, S. [1]; Shearer, 2003, S. 256). Trotz europäischer Initiativen, wie z. B. der *Berlin Declaration* (s. Kap. 1), und dem engagierten Aufbau der erforderlichen technischen Infrastruktur durch die Hochschulbibliotheken bleibt die Partizipationsrate von potenziellen Autor:innen jedoch niedrig (Demetres et al., 2020, S. 177; Pickton & McKnight, 2006, S. 204; Scherer, 2016, S. 159), wie auch Borgman als Befürworterin strenger Publikationsrichtlinien eingestuft (2007, S. 102;149).

Bankier und Smith vertreten die Auffassung, dass sich der Herausforderung von geringen Partizipationsraten begegnen lässt, indem der Gegenstandsbereich von IR über begutachtete Zeitschriftenartikel hinaus erweitert wird (2010, S. 247). Eine bedarfsgerechte Lockerung der Publikationsrichtlinien vergrößert die Anzahl an Stakeholdern, d. h. potenziellen Autor:innen und Lesenden, was den Wert des IR für die Hochschule erhöht. Zu den Stakeholdern gehören dabei nicht nur Forschende, Lehrende und Promovierende, sondern auch Bachelor- sowie Masterstudierende (Abrizah et al., 2015, S. 731; Bankier & Smith, 2010, S. 247). Diese Position lässt sich zurückführen auf Lynch, der ebenfalls gegen „gate keeping“ argumentiert und die Übernahme traditioneller Richtlinien, wie etwa eine fachliche Begutachtung, als kontraproduktiv ablehnt. Um den Bedürfnissen aller Stakeholder gerecht zu werden, sollte die Zugehörigkeit zur Hochschule als hinreichend dafür erachtet werden, im IR publizieren zu dürfen (2003, S. 333). Bedenken, dass eine derartige Öffnung des IR das wissenschaftliche Personal vom dortigen Publizieren abhalten könnte, haben sich in der Praxis nicht bestätigt (Bankier & Smith, 2010, S. 247).

Im Verlauf dieses Kapitels hat sich gezeigt, dass Studierende schon seit der Anfangszeit von IR als wichtige Gruppe von Stakeholdern identifiziert wurden, deren Partizipation maßgeblich zu einer erfolgreichen Integration von IR in die intellektuelle Infrastruktur von Hochschulen beitragen kann. Das folgende Kapitel legt den Fokus daher auf studentische Partizipation und untersucht, welchen Platz studentische Inhalte in IR einnehmen können. Dabei wird sowohl eine nähere Beschreibung potenzieller studentischer Inhalte als auch eine Abwägung der Vor- und Nachteile des Selbstarchivierens für Studierende vorgenommen.

2.2 Studierende als Stakeholder bei institutionellen Repositorien

In seinem Grundlagentext formuliert Lynch die Vision, dass ein erfolgreich etabliertes IR die intellektuellen Arbeiten – d. h. Forschungs- und Lehrmaterialien – von sowohl dem wissenschaftlichen Personal als auch den Studierenden einer Hochschule enthalten wird (2003, S. 328). Trotzdem wurde zu Beginn der Implementierung von IR in den frühen 2000er-Jahren die Gruppe des wissenschaftlichen Personals als Stakeholder in den Mittelpunkt gerückt, indem vorrangig begutachtete Zeitschriftenartikel gesammelt wurden (Davis-Kahl, 2021, S. 87). Studentische Arbeiten hingegen wurden als lernunterstützende Übung betrachtet, deren Wert sich auf eine bestimmte Lehrveranstaltung beschränkt und in einer abschließenden Note ausgedrückt wird (Levin et al., 2005, S. 39;45).

Mit den Jahren ist das Interesse an studentischen Inhalten jedoch gewachsen, wie Ammarukleart und Kim im Rahmen einer Verlaufsstudie feststellen, welche die Forschungstrends zu IR anhand von bibliometrischen Daten und Text Mining untersucht. Um Trends zu identifizieren, wurde eine frühe Zeitperiode (2005-2010) mit einer späten Zeitperiode (2011-2015) verglichen. Neben Forschungsdatenmanagement und Linked Open Data traten dabei auch Studierende als neues Themenfeld hervor, was sich u. a. dadurch ausdrückt, dass studentische Autor:innengruppen, wie z. B. „student“, „undergraduate student“, „graduate student“ oder „postgraduate student“, in der späten Zeitperiode verstärkt benannt wurden. Während sich IR in der frühen Zeitperiode nahezu ausschließlich mit den Arbeiten des wissenschaftlichen Personals befassten, wurde in der späten Zeitperiode zunehmend Raum für studentische Inhalte geschaffen, die seitdem ein bedeutendes und rasch anwachsendes Inhaltssegment darstellen (2017, S. 264;274). Dieses Ergebnis reflektiert die zunehmende Anerkennung, die Studierende in ihrer neuen Rolle als Inhaltsproduzierende von den Hochschulen erfahren (Caprio, 2014, S. 147).

Die beschriebene Entwicklung wurde erst durch das Engagement der Bibliotheken ermöglicht, welche vermehrt eine Chance darin sahen, die neu aufgebaute Infrastruktur von IR zusätzlich für das Archivieren von grauer Literatur zu nutzen. Hierfür wurden verstärkt studentische Inhalte in den Blick genommen (Watkinson, 2014, S. 181;183). Wie die bereits in Kapitel 2.1 erwähnten Fallstudien von Bankier und Smith zeigen, beschränken sich die meisten IR mit Beginn der 2010er-Jahre nicht länger auf das Sammeln von begutachteten Zeitschriftenartikeln des wissenschaftlichen Personals, sondern akzeptieren nunmehr eine Vielzahl an Dokument-

arten verschiedener Autor:innengruppen, was die Partizipation vonseiten der Hochschulgemeinschaft erhöht und dementsprechend zu einer erfolgreichen Etablierung von IR beiträgt (2010, S. 256-257). Wie die Bibliothekarin Davis-Kahl beobachtet, haben diese Bemühungen zur Folge, dass die IR-Landschaft mit Beginn der 2020er-Jahre nahe davorsteht, die zu Anfang dieses Kapitels wiedergegebene Vision von Lynch zu verwirklichen. Hierzu gehört, dass die Aufnahme von studentischen Inhalten in IR inzwischen eine fest etablierte Praxis darstellt (2021, S. 88;91). Um welche Inhalte es sich dabei genau handelt, wird im folgenden Kapitel untersucht.

2.2.1 Absteckung des Sammlungsbereiches

Die Arbeitsergebnisse, die Studierende am häufigsten in IR publizieren, sind Abschlussarbeiten, welche sich unter dem Begriff *Electronic Theses and Dissertations* (ETD) zusammenfassen lassen. Auch Zeitschriftenartikel oder Arbeiten, die im Rahmen von Veranstaltungen oder Konferenzen präsentiert wurden, werden häufig in IR aufgenommen (Bankier & Smith, 2010, S. 253; Hertenstein, 2014, S. 10). Der Fokus liegt somit auf Leuchtturmarbeiten, die in Zusammenarbeit mit einer betreuenden Lehrperson entstanden sind, was keine vollständige Repräsentation studentischer Arbeitsaktivitäten darstellt (Barandiaran et al., 2014, S. 546; Davis-Kahl, 2021, S. 88; Passehl-Stoddart & Monge, 2014, S. 8; Rozum & Thoms, 2016, S. 311). Insbesondere die Arbeiten von Bachelorstudierenden sind häufig unterrepräsentiert und werden nicht standardmäßig in IR aufgenommen (Davis-Kahl, 2021, S. 88; Rozum et al., 2015, S. 804-805). Seit Mitte der 2010er-Jahre entwickelt sich der Trend jedoch dahingehend, dass verstärkt eine erweiterte Bandbreite an studentischen Arbeiten publiziert wird, die über ETD hinausgehen (Hertenstein, 2014, S. 10-11).

Wie eine derart erweiterte Sammlung studentischer Inhalte aussehen kann, demonstriert das Open-Access-Publikationsportal *SAIL* der *Hochschule für Angewandte Wissenschaften (HAW) Hamburg*, dessen Sammlungsbereich „alle studentischen Arbeiten [... umfasst], die während des Studiums an der HAW Hamburg als Teil einer Lehrveranstaltung oder eines Hochschulprojektes angefertigt wurden“ (HAW Hamburg, o. J.a). Die von SAIL verwendete Definition wird an dieser Stelle übernommen, sodass mit dem Begriff des studentischen Inhalts alle Arbeitsergebnisse von Studierenden gemeint sind, die im Rahmen der Hochschullehre und des lernenden Forschens entstanden sind.

Um welche konkreten Objekte es sich dabei handelt, hängt vom jeweiligen Studiengang ab: Neben fachübergreifend relevanten Objekten, wie Hausarbeiten oder Präsentationen, spielen in kreativen Studiengängen auch Zeichnungen oder Musikkompositionen eine Rolle, während in technischen Fächern Laborberichte oder Konstruktionsarbeiten verstärkt von Relevanz sind. Sowohl in technischen als auch in sozialwissenschaftlichen Studiengängen ist der Übergang von Dokumenten zu Daten oftmals fließend, wenn etwa Experimente oder Interviews durchgeführt werden. Vonseiten der IR-Betreibenden ist dabei zu beachten, dass das Publizieren von Daten sowohl datenschutzrechtliche als auch die Katalogisierung betreffende Besonderheiten aufweist, die auf Dokumente nicht zutreffen (Borgman, 2007, S. 182;205; Davis-Kahl, 2021, S. 88;90). Gleichwohl Forschungsdaten an dieser Stelle nicht als potenzielle studentische Sammlungsobjekte ausgeschlossen werden sollen, liegt der Fokus dieser Arbeit dennoch auf Dokumenten, da die Publikationsbereitschaft für Dokumente und Daten von unterschiedlichen Faktoren beeinflusst werden könnte, die zu untersuchen den Umfang dieser Arbeit übersteigen würde.

Für Bibliotheken und Hochschulen birgt die Aufnahme von studentischen Inhalten in IR eine Vielzahl an Vorteilen, welche die hierfür aufgewendeten Arbeitsressourcen lohnenswert machen: Während Bibliotheken innerhalb der Hochschulgemeinschaft an Relevanz gewinnen und die Chancen des freien Publizierens direkt an die Studierenden herantragen können, profitieren Hochschulen von der öffentlichen Demonstration ihrer Lehrqualität, die nicht nur die Beziehungen zu Geldgeber:innen und Alumni stärkt, sondern auch der Rekrutierung von neuem wissenschaftlichen Personal sowie von Studierenden zugutekommt (Bankier & Smith, 2010, S. 254; Davis-Kahl, 2021, S. 91;93; Exline, 2016, S. 25-26; Rozum et al., 2015, S. 809;811). Strittig ist hingegen, inwieweit das freie Publizieren vor- oder nachteilhaft für die Studierenden selbst ist – insbesondere dann, wenn die Arbeiten von Bachelorstudierenden zur Diskussion stehen (Caprio, 2014, S. 148-149). Dementsprechend findet im folgenden Kapitel eine Abwägung der Vor- und Nachteile der Selbstarchivierung aus Studierendenperspektive statt.

2.2.2 Vor- und Nachteile der Selbstarchivierung für Studierende

Bei der Untersuchung der Vorteile von Selbstarchivierungspraktiken für Studierende lässt sich unterscheiden zwischen jenen Vorteilen, die allen Studierenden zugutekommen – unabhängig davon, ob sie selbst publizieren oder nicht –, sowie

jenen Vorteilen, die konkret auf die Produzent:innen studentischer Inhalte zutreffen. Folgende zwei Vorteile eröffnen sich allen Konsument:innen:

- (a) *Hochschulwahl*: Nicht nur die Hochschulen, sondern auch angehende Studierende profitieren von der öffentlichen Demonstration der Lehrinhalte, welche die Forschungsaktivitäten an den jeweiligen Institutionen vor Augen führt und dadurch als Entscheidungshilfe für die Wahl einer Hochschule dienen kann (Davis-Kahl, 2021, S. 90; Rozum & Thoms, 2016, S. 316; Passehl-Stoddart & Monge, 2014, S. 7).
- (b) *Anschauungsbeispiele*: Das Veröffentlichen von studentischen Best Practices illustriert die Qualitätsstandards der jeweiligen Fachbereiche und ermöglicht allen Mitgliedern der akademischen Gemeinschaft, auf den Arbeitsergebnissen von Studierenden aufzubauen (Davis-Kahl, 2021, S. 90-91; Heinrich, 2004, S. 87; Rozum et al., 2015, S. 809;811).

Darüber hinaus profitieren die Produzent:innen studentischer Inhalte von folgenden drei Vorteilen:

- (c) *Bewerbungschancen*: Indem sie Bezug auf ihre Publikationen nehmen, können Studierende ihre Wettbewerbschancen auf dem Arbeitsmarkt oder bei Graduiertenschulen erhöhen (Bankier & Smith, 2010, S. 253; Borgman, 2007, S. 238; Caprio, 2014, S. 149; Exline, 2016, S. 26; Rozum et al., 2015, S. 810). Persistente Identifikatoren (z. B. URN) ermöglichen eine langfristige Referenz der Arbeiten auf Lebensläufen. Zudem geht die Selbstarchivierung in IR mit der Generierung eines Autor:innenprofils einher, in dem sämtliche Arbeiten und akademischen Fortschritte gebündelt dokumentiert werden können (Barandiaran et al., 2014, S. 547; Passehl-Stoddart & Monge, 2014, S. 8; Rozum & Thoms, 2016, S. 316).
- (d) *Feedback*: Das öffentliche Formulieren von Ideen und der damit einhergehende Umgang mit Feedback stellt eine entscheidende Stufe im Forschungsprozess dar und eröffnet einen neuen Raum des Dialogs, in dem die Studierenden ihre Ideen gemeinsam mit ihren Kommiliton:innen und Lehrenden weiterentwickeln können (Caprio, 2014, S. 148;153; Walkington, 2012, S. 558).

(e) *Wissenschaftliche Karriere*: Beim Publizieren trainieren die Studierenden nicht nur ihre Forschungs-, Medien- und Informationskompetenz, sondern kommen auch mit rechtlichen Themenfelder der wissenschaftlichen Publikationsökonomie, wie z. B. Copyright, intellektuellem Eigentum oder Open Access, in Berührung. In Kombination mit der in Vorteil (d) beschriebenen Verbesserung der Kommunikationskompetenz wird das akademische Selbstbewusstsein der Studierenden gestärkt, was sie dazu motivieren könnte, auch nach Abschluss ihres Studiums in der Forschung tätig zu sein (Caprio, 2014, S. 148; Davis-Kahl, 2021, S. 89; Passehl-Stoddart & Monge, 2014, S. 2;6; Rozum et al., 2015, S. 811).

Neben den beschriebenen Vorteilen gibt es zwei potenzielle Nachteile, die sich negativ auf selbstarchivierende Studierende auswirken könnten und im Folgenden dargestellt werden. Aufgeführt werden dabei auch die in der Literatur genannten Lösungsvorschläge dafür, wie sich den genannten Einwänden begegnen ließe.

(a) *Kompromittierung der Online-Identität*: Sowohl die IR-Betreibenden als auch die Studierenden selbst befürchten, dass das Publizieren von Arbeiten mangelnder Qualität den Autor:innen langfristig anhaft (Borgman, 2007, S. 239; Exline, 2016, S. 17;26; Zeller & Stenberg, 2016, S. 44). Gleichzeitig hat das wissenschaftliche Personal die Sorge, dass unausgereifte Publikationen negativ auf die Hochschule und ihre Lehrveranstaltungen zurückfallen könnten (O’Gara & Davis, 2016, S. 158; Passehl-Stoddart & Monge, 2014, S. 9). Für diesen Kritikpunkt existieren folgende Gegeneinwände:

- i. *Qualitätssicherung*: Um die Publikation von Arbeiten mangelnder Qualität zu vermeiden, verfügen die meisten IR bereits über Richtlinien zur Qualitätssicherung. Auch die Zitierfähigkeit der studentischen Inhalte wird durch das Durchführen einer Qualitätskontrolle erhöht. In der Regel spielt das wissenschaftliche Personal eine wichtige Rolle bei der Formulierung der Inklusionskriterien (Davis-Kahl, 2021, S. 88; Gilbert, 2004, S. 23; Hertenstein, 2014, S. 10), sodass die Bibliotheken als staatliche Einrichtungen nicht eigenmächtig über die Publikationswürdigkeit von Arbeiten entscheiden.
- ii. *Identifikation*: Wenn studentische Autor:innen bei ihren Publikationen als solche identifiziert werden, erscheinen ihre Arbeiten im Kontext des IR als Arbeitsergebnisse eines klar benannten Ausbildungs-

stadiums. Die Zuordnung der Dokumente zu einer spezifischen Lehrveranstaltung stellt eine weitere Möglichkeit dafür dar, die Publikationen mit einem Zeitstempel zu versehen (Rozum et al., 2015, S. 810-811; Zeller & Stenberg, 2016, S. 45).

iii. *Löschoption*: Obwohl die Langzeitarchivierung nach Crow zu den Kerneigenschaften eines IR gehört (s. Kap. 2.1), erlauben manche Bibliotheken, wie z. B. die der *University of New Hampshire*, im Falle studentischer Inhalte ein nachträgliches Löschen bereits veröffentlichter Dokumente (Exline, 2016, S. 26). Andere IR-Betreibende hingegen verbieten das Zurückziehen von Publikation auch dann, wenn es sich hierbei um studentische Arbeiten handelt. Zu letzteren zählt u. a. auch das bereits in Kapitel 2.2.1 erwähnte Open-Access-Publikationsportal SAIL der HAW Hamburg, welches auf die Unwiderrufbarkeit der dort verwendeten *Creative Commons*-Lizenzen verweist (HAW Hamburg, o. J.a). Crow gibt zu bedenken, dass eine derartige immerwährende Zugänglichkeit die Partizipationsbereitschaft potenzieller Autor:innen schmälern könnte, und empfiehlt daher das Entwickeln ausgewogener Publikationsrichtlinien, welche die Zielsetzung von IR und die Interessen der Autor:innen gleichermaßen berücksichtigen (2002, S. 18). Für Davis-Kahl haben Studierende sogar ein Recht darauf, ihre Dokumente auch noch Jahre später zu löschen (2021, S. 92). Das Gewähren einer Löschoption im besonderen Falle studentischer Inhalte würde verhindern, dass die Autor:innen langfristig mit Werken in Verbindung gebracht werden, deren Qualität sie nicht länger zufriedenstellt.

(b) *Plagiate*: In einem IR verfügbare studentische Arbeiten laufen Gefahr, plagiiert zu werden. Insbesondere Lehrende befürchten, dass andere Studierende die Arbeiten weitenutzen und in verwandten Lehrveranstaltungen als ihre eigenen ausgeben könnten (Davis-Kahl, 2021, S. 92; O’Gara & Davis, 2016, S. 158; Passehl-Stoddart & Monge, 2014, S. 9). Diesem Einwand lässt sich wie folgt begegnen:

i. *Allgemeingültigkeit*: Alle publizierten Forschungsergebnisse können auf unethische Weise genutzt werden, unabhängig davon, ob sie digital oder in Printform erschienen sind. Die Gefahr für Studierende ist dabei nicht größer als für andere Autor:innen, weshalb die oben

aufgeführten Vorteile des Publizierens schwerer wiegen als die Gefahr möglicher Plagiate (Davis-Kahl, 2021, S. 92; Pickton & McKnight, 2006, S. 214).

- ii. *Plagiatsoftware*: Studentische Arbeiten, die in einem IR zur Verfügung stehen, können mithilfe von Plagiatsoftware geprüft werden. In Streitfällen gibt der Zeitstempel Aufschluss darüber, welche Arbeit zuerst erstellt wurde (Exline, 2016, S. 26).

Nachdem die Vorteile studentischer Selbstarchivierung dargestellt und die wiederkehrend in der Literatur genannten Einwände relativiert wurden, beschäftigt sich das folgende Kapitel mit existierenden Studien zur Publikationsbereitschaft von Studierenden in IR.

3 Stand der Forschung zur Publikationsbereitschaft von Studierenden in institutionellen Repositorien

Es existieren nur wenige Studien, die sich mit studentischen Inhalten in IR befassen (Hertenstein, 2014, S. 2; Rozum et al., 2015, S. 805). Ein Leuchtturmbeispiel sind strukturierte Interviews mit Studierenden der *Loughborough University*, die im Jahr 2006 von der Bibliothekarin Pickton und dem Informationswissenschaftler McKnight durchgeführt wurden. Ziel der Untersuchung war es, den Wissensstand und die Einstellung der Studierenden gegenüber dem Publizieren im Open Access und digitalen Repositorien im Allgemeinen sowie dem IR ihrer Hochschule im Speziellen festzustellen (2006, S. 205). Insgesamt wurden 34 Studierende interviewt, von denen 13 der technischen, neun der naturwissenschaftlichen und zwölf der sozial- und geisteswissenschaftlichen Fakultät angehören (2006, S. 207, Tab. 1). Hervorzuheben ist, dass es sich hierbei um Forschungs- bzw. Promotionsstudierende handelt, die auf einen Dokortitel hinarbeiten (2006, S. 203;214).

Die Ergebnisse von Pickton und McKnight lassen sich in die drei Segmente (a) Dokumenttypen, (b) motivierende Faktoren sowie (c) abschreckende Faktoren einteilen:

- (a) *Dokumenttypen*: Mehr als drei Viertel der Befragten stimmen zu, dass sie Konferenzschriften (91,2%), am Department entstandene Arbeiten (82,4%)

sowie ihre vollständige Abschlussarbeit (79,4%) im IR einstellen würden. Darüber hinaus erklären sie sich u. a. zur Publikation von Softwarecode, Bibliografien, Abstracts, Bildern und Kunstwerken sowie Audiopräsentationen für sehbehinderte Menschen bereit (Pickton & McKnight, 2006, S. 209).

(b) *Motivierende Faktoren*: Tabelle 1 zeigt jene Faktoren, die die Studierenden zur Selbstarchivierung im IR motivieren. Dabei enthält die Spalte „Priorisierung in %“ die relative Anzahl der Studierenden, die den fraglichen Faktor in ihre „Top Fünf“ der wichtigsten Motivatoren gewählt haben (Pickton & McKnight, 2006, S. 212). Die Tabellenansicht wird an dieser Stelle auf die nach Einschätzung der Autor:innen wichtigsten Faktoren beschränkt.

Tab. 1: Zur Selbstarchivierung motivierende Faktoren von Studierenden der Loughborough University (Quelle: Pickton & McKnight, 2006, S. 212)

Motivierende Faktoren	Priorisierung in %
Dissemination der eigenen Arbeit in der Forschungsgemeinschaft und über diese hinaus	58,8%
Erhalten von Feedback und Anmerkungen	50,0%
Unterstützung der Prinzipien von Open Access	44,1%
Teilen von Forschungsmaterialien mit Mitarbeiter:innen	32,4%
Ermutigung durch Betreuer:innen	29,4%
Verfügbarmachung der eigenen Arbeit für andere Studierende	23,5%
Sammeln von Informationen zu Karrierezielen	23,5%

Neben einer weitläufigen Kommunikation der eigenen Forschungsergebnisse spielen für die Befragten auch Feedback zu ihrer Arbeit sowie ideelle Gründe bezüglich der Sinnhaftigkeit von Open Access eine Rolle. In einem verwandten Interviewsegment, in dem die Studierenden mittels geschlossener Fragen nach ihren allgemeinen Gründen für das Publizieren – ob im IR oder anderswo – befragt wurden, treten zwei weitere potenzielle Motivatoren hervor: 64,7% der Befragten stimmen zu, dass sie publizieren, um ihre *Aussicht auf Fördergelder* zu erhöhen. Darüber hinaus stimmen 55,9% der Studierenden zu, dass das Mehrere *persönlichen Prestiges* einen Publikationsgrund für sie darstellt (Pickton & McKnight, 2006, S. 205;208, Tab. 2). Weil diese beiden Gründe nicht Teil des Motivatoren-Segments und der

dort vorgenommenen Top-Fünf-Priorisierung waren, bleibt unklar, wie wichtig sie in Relation zu den in Tab. 1 aufgeführten Motivatoren sind.

(c) *Abschreckende Faktoren:* Tabelle 2 zeigt die nach Einschätzung der Autor:innen relevantesten Faktoren, die die Studierenden vom Publizieren ihrer Arbeiten im IR abhalten. Analog zur Darstellung der Motivatoren enthält die Spalte „Priorisierung in %“ die relative Anzahl der Studierenden, die den fraglichen Faktor in ihre „Top Fünf“ der wichtigsten abschreckenden Faktoren gewählt haben (Pickton & McKnight, 2006, S. 212). Unklar ist an dieser Stelle, wie die Autor:innen den Unterschied zwischen unerlaubtem Kopieren und der Anfertigung von Plagiaten definieren.

Tab. 2: Von der Selbstarchivierung abhaltende Faktoren von Studierenden der Loughborough University (Quelle: Pickton & McKnight, 2006, S. 212)

Abschreckende Faktoren	Priorisierung in %
Verwirken zukünftiger Veröffentlichungsmöglichkeiten in anderen Publikationsoutlets	55,0%
Unerlaubtes Kopieren der Inhalte durch andere	32,4%
Copyrightstreitigkeiten mit Verlegern bereits veröffentlichter Inhalte	29,4%
Gefahr von Plagiaten	29,4%
Vertraulichkeit der Inhalte	26,5%

Insbesondere die Ergebnisse der abschreckenden Faktoren führen die Limitation der Studie von Pickton und McKnight vor Augen, indem die Fokussierung auf Copyrightfragen die Befürchtungen von Promotionsstudierenden reflektiert, die eine wissenschaftliche Karriere anstreben und bereits Publikationserfahrung gesammelt haben. Die Autor:innen selbst identifizieren die geringe Stichprobengröße ebenfalls als Limitation ihrer Arbeit und weisen auf die Notwendigkeit weiterer Untersuchungen hin (2006, S. 218). Im Fazit resümieren Pickton und McKnight, dass die überwältigende Anzahl der interviewten Studierenden einer Selbstarchivierung positiv gegenübersteht: 33 von 34 Befragten wären zur Publikation im IR ihrer Hochschule bereit (Pickton & McKnight, 2006, S. 211;213).

Vereinzelte Studien weisen darauf hin, dass nicht nur Promotionsstudierende offen für eine Selbstarchivierung in IR sind. Im Zeitraum vom 19. April bis zum 24. Juni 2006 gaben im Rahmen des von einer Gruppe von Informationswissenschaftler:innen der *University of Michigan* initiierten *MIRACLE-Projekts (Making Institutional Repositories a Collaborative Learning Environment)* insgesamt 446 Bibliothekslei-

ter:innen US-amerikanischer Hochschulen Auskunft über den Stand der IR-Implementierung an ihren jeweiligen Institutionen. 289 Teilnehmende (64,8%) gehören Hochschulen an, die Bachelor- und Masterprogramme anbieten (*Master's and Baccalaureate Institutions*, M&BIs), während die übrigen 157 Teilnehmenden Forschungsuniversitäten angehören. Die Untersuchung ergab, dass an M&BIs Bachelorstudierende zu gleichen Anteilen im IR publizieren wie das wissenschaftliche Personal (Markey et al., 2008, S. 157;167;171-172). Hochschulübergreifend identifizieren die Bibliotheksleitungen drei hauptsächliche Gründe für das Publizieren im IR: Während die ersten beiden Motivatoren – die Dissemination von Forschungsergebnissen und das Mehreren persönlichen Prestiges – auch bei Pickton und McKnight (2006) von Relevanz sind, kommt mit der *Formatvielfalt* von IR, die neben Textdokumenten auch z. B. Daten oder Video- und Audiomaterial archivieren können, ein weiterer Aspekt hinzu (Markey et al., 2008, S. 167).

Eine Limitation der Untersuchung von Markey et al. (2008) ist, dass – im Gegensatz zur Studie von Pickton und McKnight (2006) – keine direkte Befragung der Studierenden stattgefunden hat, sodass alle Ergebnisse durch die Wahrnehmung der Bibliotheksleitungen gefiltert werden. Da manche Leiter:innen das Beantworten des Fragebogens an andere Personen delegiert haben, die stärker in den IR-Implementierungsprozess eingebunden sind (Markey et al., 2008, S. 160), ist zudem nicht immer klar, welche Funktion die antwortgebenden Personen genau innehalten.

Sowohl im Ergebnis als auch in der Limitation deckt sich die Studie von Markey et al. (2008) mit einer Untersuchung, die Hertenstein (2014) als Bibliothekarin an der *Bowling Green State University* im Zeitraum vom 8. Mai bis zum 7. Juni 2013 durchgeführt hat. Insgesamt nahmen 44 Bibliothekar:innen an der Befragung teil, die 23 Items zum Stand der IR-Implementierung an den jeweiligen Institutionen enthielt, wobei ein besonderer Fokus auf die Inklusion studentischer Inhalte gelegt wurde. Die Untersuchung ergab, dass Studierende zur freiwilligen Publikation in IR bereit sind, sofern sie zuvor über die Vorteile einer Selbstarchivierung aufgeklärt wurden. Wie die Autorin selbst anmerkt, wird die Studie nicht nur durch die geringe Stichprobengröße limitiert, sondern auch durch den Mangel an studentischen Perspektiven und Erfahrungsberichten (Hertenstein, 2014, S. 1;3;10-11;). Wie im folgenden Kapitel näher ausgeführt wird, verbleiben aktuelle Daten zur Selbstarchivierungsbereitschaft von Studierenden daher als Forschungsdesideratum.

4 Forschungsfrage und Zielsetzung

Indem die Vorteile eines freien Zugangs zu wissenschaftlichen Forschungsergebnissen betont werden (s. Kap. 1), argumentiert die Open-Access-Bewegung primär aus Nutzendensperspektive. Weniger beleuchtet ist hingegen die Perspektive der Inhaltsproduzent:innen, die ihre Dokumente und Daten frei zur Verfügung stellen (Borgman, 2007, S. 193). Dass Diskrepanzen zwischen diesen beiden Perspektiven bestehen, deutet sich auch in der in Kapitel 3 vorgestellten Studie von Pickton und McKnight an, die beobachten, dass sich die interviewten Promotionsstudierenden vom IR ihrer Universität lesenden Zugriff auf Dokumentarten erhoffen, die sie selbst nicht zu publizieren bereit sind (2006, S. 215). Weil der Erfolg eines IR – wie in Kapitel 2.1 erläutert – maßgeblich von der Publikationsbereitschaft der Hochschulmitglieder abhängt, stehen Hochschulbibliotheken angesichts geringer Partizipationsraten vor der Herausforderung, Autor:innen für ihre IR zu akquirieren.

Eine Gruppe von Stakeholdern, die nach der Implementierung von IR lange Zeit als potenzielle Konsument:innen und Produzent:innen vernachlässigt wurde, sind Studierende (s. Kap. 2.2). Diese Zielgruppe näher einzubinden und als Autor:innen zu gewinnen, stellt eine Möglichkeit dafür dar, IR zu beleben und stärker in der intellektuellen Infrastruktur ihrer Hochschulen zu verankern. Hierfür ist es erforderlich, die Publikationsbereitschaft von Studierenden differenziert zu untersuchen. Wie in Kapitel 3 gezeigt wurde, weist die existierende Forschung zur studentischen Publikationsbereitschaft die Limitation auf, dass die Studierenden zu selten direkt befragt wurden. Dieses Desideratum stellt die Forschungslücke dar, an der die vorliegende Arbeit ansetzt. Wie bereits in Kapitel 1 skizziert, wird hierfür konkret der Frage nachgegangen, welche Faktoren die Bereitschaft von Studierenden zur Selbstarchivierung in IR beeinflussen. Ziel ist dabei die Ableitung von Handlungsempfehlungen für Hochschulbibliotheken zur Akquise studentischer Autor:innen. Um die reine Publikationsbereitschaft – d. h. das Ermöglichen eines freien lesenden Zugriffs – ohne Beeinflussung durch Einstellungen bezüglich spezifischer Bearbeitungs- und Weiternutzungsrechte untersuchen zu können, wird der Begriff des Open Access in dieser Arbeit im Sinne des Gratis-Gold-Modells verwendet.

Im weiteren Verlauf der Arbeit wird zunächst eine literaturbasierte Modellierung der Publikationsbereitschaft von Studierenden in Abhängigkeit von mehreren Prädiktoren vorgenommen. Anschließend werden die aufgestellten Hypothesen mittels einer linearen Regressionsanalyse nach der Methode der kleinsten Quadrate

empirisch überprüft, indem eine Stichprobe der genannten Zielgruppe im Rahmen eines Online-Fragebogens befragt wird.

5 Modellierung der Bereitschaft zur Selbstarchivierung

Grundlage der in dieser Arbeit vorgenommenen Regressionsanalyse ist das empirische Modell der Informationswissenschaftlerin Kim zur Publikationsbereitschaft von wissenschaftlichem Personal in IR. Kapitel 5.1 stellt die theoretischen Grundlagen des Kim-Modells in Form des *Socio-Technical Network Models* (STIN) sowie der Austauschtheorie (*Social Exchange Theory*) vor. Anschließend wird in Kapitel 5.2 dargestellt, wie Kim ihr Modell im Verlauf einer Pilotstudie (2007) sowie einer Nachfolgestudie (2010) und einer Abschlussstudie (2011) entwickelt hat. In diesem Kontext findet auch eine kritische Beurteilung ihres Vorgehens statt. Um das Modell für die gesetzte Fragestellung dieser Arbeit nutzen zu können, werden Kims Variablen in Kapitel 5.3 auf Basis bestehender Forschung zu studentischen Inhalten in IR (s. Kap. 2.2 und 3) angepasst sowie ergänzt bzw. verworfen. Hierbei entsteht eine empirische Neumodellierung, welche die im Rahmen der vorliegenden Arbeit zu überprüfenden Hypothesen zur Publikationsbereitschaft von Studierenden in IR ausdrückt (s. Kap. 6).

5.1 Theoretischer Hintergrund: *Socio-Technical Network Model* (STIN) und Austauschtheorie (*Social Exchange Theory*)

Das Kim-Modell ist eine empirische Modellierung der motivierenden bzw. abschreckenden Faktoren, die Einfluss auf die Selbstarchivierungsaktivitäten wissenschaftlichen Personals in IR ausüben (Kim, 2007, S. [1]). Nach Kim lassen sich IR als dreierlei Entitäten konzeptualisieren (2007, S. [2]):

- (a) *Electronic Scholarly Communication Forums* (e-SCF): Der auf Kling et al. zurückgehende Sammelbegriff e-SCF bezeichnet eine Gruppe von Foren zur computervermittelten Kommunikation, zu der u. a. auch serverbasierte Repositorien zählen. Die kommunikative Funktion eines e-SCF reicht von der Bereitstellung von Forschungsdaten über Lehrmaterialien bis hin zu wissenschaftlichen Arbeiten in verschiedenen Begutachtungsstadien, was eine Veränderung der Wissenschaftskommunikation nach sich zieht (Kim,

2007, S. [2]; Kling et al., 2003, S. 47). Dementsprechend lassen sich e-SCF als Katalysator von E-Research begreifen, in deren Zuge die elektronischen Foren kollaborativ genutzt werden (s. Kap. 1).

- (b) *Digitale Bibliotheken*: Anstatt lediglich als passive Ressourcen für die Informationsbedürfnisse einer bestimmten Zielgruppe bereitzustehen, werden digitale Bibliotheken aktiv durch die informationsbezogenen Aktivitäten ihrer Nutzenden gestaltet. In diesem Sinne können auch IR als digitale Bibliotheken betrachtet werden, die durch die dort archivierten Beiträge bzw. Wissensobjekte (s. Kap. 2.1) der wissenschaftlichen Gemeinschaft konstituiert werden (Kim, 2007, S. [2]).
- (c) *Knowledge Management Systems (KMS)*: Innerhalb von Organisationen, wie z. B. Hochschulen, unterstützen IT-basierte Systeme beim Management des gewachsenen Wissensschatzes, wobei sämtliche Prozesse von der Generierung über die Archivierung und Abrufung bis hin zum Transfer und zur Anwendung des akkumulierten Wissens abgedeckt werden (Alavi & Leidner, 2001, S. 114; Kim, 2007, S. [2]).

Nachdem Kim die Beschaffenheit von IR näher bestimmt hat, fährt sie fort, theoretische Implikationen aus den genannten drei Eigenschaften abzuleiten. Die Charakterisierung von IR als e-SCF, was ihre kommunikative Funktion betont, sowie als digitale Bibliotheken, was ihre gemeinschaftliche Konstituierung hervorhebt, impliziert, dass sich IR vor dem Hintergrund des *Socio-Technical Network Models (STIN)* untersuchen lassen (Kim, 2007, S. [2]). Das STIN ist ein auf Kling et al. (2003) zurückgehendes Modell, dessen Grundannahme darin besteht, dass technologische Systeme nicht isoliert, sondern nur im Zusammenspiel mit sozialen Faktoren sinnvoll betrachtet werden können. Der Grund hierfür liegt in der Notwendigkeit einer zufriedenstellenden *Usability* (Gebrauchstauglichkeit) der Systeme, die nicht zum Selbstzweck erschaffen wurden, sondern immer eine konkrete Funktion für eine ausgewählte Zielgruppe zu erfüllen haben. Technische und soziale Elemente konstituieren sich gegenseitig, sodass das STIN an der Schnittstelle zwischen Technologie und sozialem Verhalten ansetzt (Borgman, 2007, S. 3;37;173; Kim, 2010, S. 1910; Kling et al., 2003, S. 54;56).

Dabei weist das STIN jedoch, wie Kim bemerkt, ein entscheidendes Desideratum auf: In Ermangelung konkreter Variablen stellt das Modell keine hinreichende theoretische Grundlage zur empirischen Untersuchung der Anreizstruktur dar, welche

die Selbstarchivierungsbereitschaft in IR beeinflusst. Um motivierende sowie abschreckende Faktoren operationalisieren zu können, kombiniert Kim das STIN daher mit der Austauschtheorie, deren Anwendbarkeit sie aus der Charakterisierung von IR als KMS ableitet, welche den Austausch von Wissen hervorhebt. Die Vereinbarkeit dieser Theorien, so Kim, ergibt sich daraus, dass beide den Austausch von materiellen sowie immateriellen Ressourcen durch Netzwerkbeziehungen zwischen Individuen oder Gruppen beschreiben (2007, S. [2]; 2010, S. 1910).

Zu den immateriellen Ressourcen, die gemäß der Austauschtheorie geteilt werden können, gehört u. a. Wissen in Form von Informationen. Inwieweit dieser Wissensaustausch wiederholt stattfindet, hängt dabei von den bestehenden Beziehungen zwischen den sich austauschenden Individuen bzw. Gruppen ab. Wiederkehrende Interaktionen etablieren feste Verhaltensmuster und stärken das Netzwerk langfristig. Im akademischen Kontext wird Wissen vorrangig durch das Publizieren von Forschungsergebnissen geteilt, weshalb die Austauschtheorie zur Untersuchung der Selbstarchivierungsbereitschaft in IR geeignet ist (Kim, 2007, S. [2]; 2010, S. 1910-1911).

Die konkreten Variablen zur Erfassung der Anreizstruktur lassen sich in vier Gruppen aufteilen: Während (a) *Gewinnfaktoren* positive Auswirkungen des Austauschs beschreiben, die zum Teilen des eigenen Wissens motivieren, umfassen (b) *Kostenfaktoren* negative Folgen bzw. Bedenken, die potenzielle Autor:innen vom Publizieren abschrecken. Darüber hinaus beeinflussen (c) lokale *Kontextfaktoren*, wie z. B. herrschendes Vertrauen oder geltende Normen, die Entscheidung zur Selbstarchivierung. Obwohl sie nicht unmittelbarer Bestandteil der Anreizstruktur sind, können auch (d) *individuelle Faktoren* bzw. Eigenschaften der potenziellen Autor:innen, wie z. B. demografische Daten oder akademische Erfahrung, Einfluss auf die Publikationsbereitschaft nehmen (Kim, 2007, S. [2]; 2010, S. 1910-1911; 2011, S. 247).

Im gegebenen Kontext beruht die Anwendung der Austauschtheorie auf der Annahme, dass potenzielle Autor:innen eine mindestens implizite Kosten-Nutzen-Abwägung der Selbstarchivierung in IR vornehmen (Abrizah et al., 2015, S. 734; Kim, 2010, S. 1911). Eine ähnliche, häufig unterbewusst stattfindende Abwägung findet im Bereich der Technologieadaptation statt, wenn User:innen den Nutzen einer neuen Technologie höher bewerten als die Aufwandskosten des Umstiegs (Borgman, 2007, S. 3). Analog hierzu müssen auch für potenzielle Autor:innen die Vorteile des Publizierens schwerer als seine Nachteile wiegen. In dem Versuch,

diesen Abwägungsprozess quantifizierbar zu machen, hat Kim (2007; 2010; 2011) ein eigenes empirisches Modell aller vier Faktorengruppen aufgestellt, welches sie im Zuge dreier Studien überprüft und auf Grundlage ihrer Ergebnisse sukzessiv überarbeitet hat. Kims Vorgehensweise wird im Folgenden dargestellt und evaluiert.

5.2 Entwicklung und kritische Einordnung des Kim-Modells

Dieses Kapitel bietet einen zusammenfassenden Überblick sowie eine kritische Beurteilung von Kims Entwicklungsprozess einer Modellierung jener Faktoren, die Einfluss auf die Selbstarchivierungsaktivitäten wissenschaftlichen Personals in IR nehmen. Das Modell wurde im Verlauf dreier Studien erarbeitet, die im Rahmen des bereits in Kapitel 3 erwähnten MIRACLE-Projekts im Zeitraum zwischen April 2006 und Mai 2007 durchgeführt wurden (Kim, 2007, S. [3]; 2010, S. 1913; 2011, S. 248n33). Während die Pilotstudie (2007) sowie die Nachfolgestudie (2010) das Erheben empirischer Fragebogen- und Interviewdaten beinhalten, nutzt die Abschlussstudie das Datenmaterial der Nachfolgestudie nach (2011, S. 248-49).

5.2.1 Pilotstudie (2007)

Im April 2006 führte Kim eine Pilotstudie anhand einer nicht-repräsentativen Stichprobe von 31 wissenschaftlichen Mitarbeitenden verschiedener Fachrichtungen einer US-amerikanischen Universität durch, die ein DSpace-basiertes IR betreibt (Kim, 2007, S. [3]). Abbildung 1 zeigt Kims erste Modellierung der Selbstarchivierungsaktivitäten wissenschaftlichen Personals in IR:

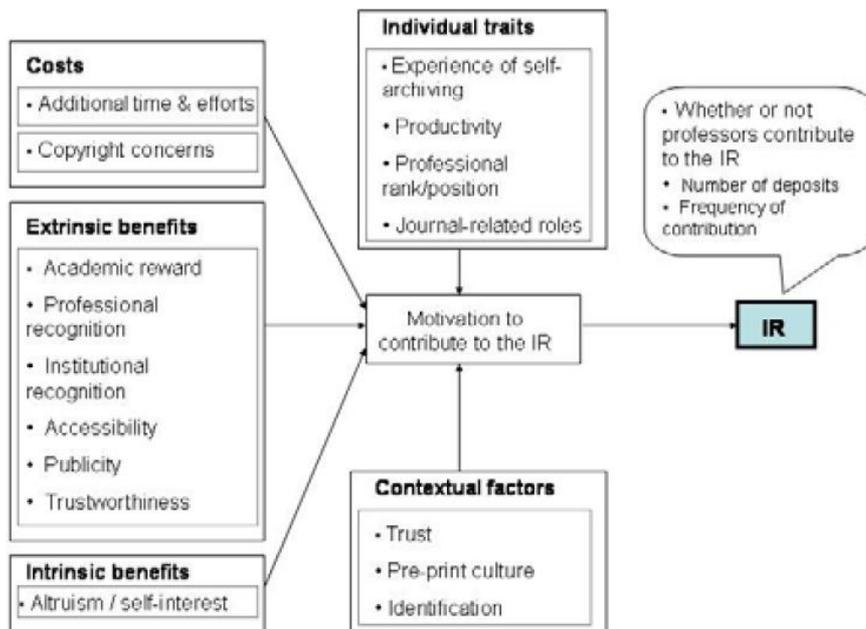


Abb. 1: Version 1 des Kim-Modells (Quelle: Kim, 2007, S. [3], Abb. 1)

Die Gruppe der Gewinnfaktoren wird aufgespalten in extrinsische und intrinsische Vorteile, die gemeinsam mit den Kosten- und Kontextfaktoren sowie den individuellen Eigenschaften Einfluss auf die Beitragsmotivation nehmen. Dieses Konstrukt wird in Form eines dichotomen Items (ja/nein) operationalisiert, das nach der Wahrscheinlichkeit einer zukünftigen Selbstarchivierung im IR fragt. Die Motivationsvariable nimmt wiederum Einfluss auf die tatsächliche Publikationsaktivität der Befragten, die anhand der Anzahl und Häufigkeit ihrer Publikationen im IR gemessen wird. Aufgrund der geringen Stichprobe wurden für die Auswertung Permutationstests eingesetzt, wobei die Gruppen der Befragten mit bzw. ohne Selbstarchivierungserfahrung sowie mit bzw. ohne zukünftige Selbstarchivierungsmotivation miteinander verglichen wurden. Darüber hinaus bewertete die publikationsmotivierte Gruppe eine Reihe potenzieller Motivatoren anhand einer fünfstufigen Skala („sehr unwichtig“ bis „sehr wichtig“). Durch Mittelwertbildungen wurde festgestellt, welche Motivatoren am stärksten wiegen (Kim, 2007, S. [3;6-7]).

Tabelle 3 bietet eine Zusammenschau aller signifikanten Ergebnisse der Permutationstests sowie relevanter Auffälligkeiten in der deskriptiven Statistik:

Tab. 3: Relevante Faktoren aus Kims Pilotstudie (Quelle: Kim, 2007, S. [6-8])

Gewinnfaktoren	Kostenfaktoren	Kontextfaktoren
Kommunikation der eigenen Forschung	Copyrightbedenken	Preprint-Kultur am Department
Erhöhter Impact der eigenen Forschung	Bedenken bzgl. Pflege und Wartung des IR	Angst vor Plagiaten
Vergrößerte Leserschaft		Erhalten von Fördergeldern
Altruismus		
Langzeitarchivierung		
Nutzungsstatistiken		

Einige Faktoren reflektieren die Ergebnisse der in Kapitel 3 vorgestellten Studie an der Loughborough University von Pickton und McKnight (2006). Aufseiten der Gewinnfaktoren zählen hierzu eine verbesserte Dissemination der eigenen Forschungsergebnisse, die Aussicht auf Fördergelder sowie Altruismus, definiert als „the idea of providing OA benefits for users“ (Kim, 2010, S. 1909). Kosten- und Kontextfaktoren, die in beiden Studien relevant sind, umfassen Copyrightbedenken und die Gefahr von Plagiaten. Während für die Promotionsstudierenden der Loughborough University das Erhalten von Feedback einen hohen Stellenwert einnimmt, legen die wissenschaftlichen Mitarbeitenden aus Kims Pilotstudie verstärkt Wert auf die Wirkkraft (*impact*) und nachhaltige Archivierung ihrer Arbeit. Obwohl Pickton und McKnight sicherstellen, dass ihre Stichprobe verschiedene Fakultäten repräsentiert, wird der fachliche Hintergrund der Studierenden nicht als potenzieller Faktor evaluiert. Kim hingegen nimmt diese Variable in Form einer subjektiv in der eigenen Fachdisziplin bzw. am eigenen Department wahrgenommenen Preprint-Kultur in ihr Modell auf, wobei der Begriff *Preprint* einen wissenschaftlichen Beitrag bezeichnet, der sich im Stadium zwischen einem Manuskript und einem formell publizierten Artikel befindet (Borgman, 2007, S. 51). In manchen Disziplinen, wie z. B. Physik, Mathematik oder den Wirtschaftswissenschaften, ist es gebräuchlich, diese Preprints über ein Repositorium frei zugänglich zu machen. In anderen Fächern, wie z. B. Psychologie, Soziologie oder den Geisteswissenschaften, ist das freie Publizieren von Preprints hingegen unüblich, was die Motivation zur Selbstarchivierung – wie Kim vermutet – beeinflussen könnte (Borgman, 2007, S. 52; Kim, 2007, S. [6]; 2010, S. 1913).

Obwohl Kim nicht explizit Bezug auf existierende Modelle der Technologieakzeptanz nimmt, erinnert die zweistufige AV ihres Modells (Selbstarchivierungsmotivation und -aktivität) an die *Unified Theory of Acceptance and Use of Technology* (UTAUT) von Venkatesh et al., in der die AV der Nutzungsabsicht die tatsächliche Nutzung einer Technologie beeinflusst (2003, S. 447, Abb. 3). Wie sich im Verlauf dieses Kapitels zeigt, wird Kim ihre AV jedoch mit der Nachfolgestudie neu konzipieren. Die Pilotstudie greift noch nicht so weit, das vorgeschlagene Modell mittels einer Regressionsanalyse zu überprüfen: Einerseits werden nicht alle untersuchten Faktoren explizit im Modell aufgelistet – so werden z. B. Bedenken bezüglich der Langzeitarchivierung nicht als Kostenfaktor aufgeführt. Andererseits werden nicht alle im Modell aufgeführten Faktoren tatsächlich untersucht – so wird z. B. der Kontextfaktor der Identifikation, definiert als „faculty members’ attitudes towards membership and loyalty towards their universities“ (Kim, 2011, S. 247), unbegründet ausgespart. Diese Inkonsistenzen weisen darauf hin, dass die Pilotstudie ähnlich einem Pretest dem Zweck dient, anhand von deskriptiver Statistik und Permutationstests erste grobe Auffälligkeiten zu identifizieren, um sowohl das Modell als auch den korrespondierenden Fragebogen vor dem Durchführen einer größer angelegten Studie verbessern zu können. Anstatt finale Konklusionen zu ziehen, weist Kim auf interessante Auffälligkeiten hin, auf deren Grundlage Vermutungen über potenzielle Einflussbeziehungen geäußert werden, die zu überprüfen als Desiderata für die Nachfolge- und Abschlussstudie verbleiben.

5.2.2 Nachfolgestudie (2010)

In der Nachfolgestudie erweitert Kim ihr Forschungsinteresse von der Publikationsbereitschaft wissenschaftlichen Personals in IR im Speziellen auf die Bereitschaft derselben Zielgruppe zur Selbstarchivierung im Allgemeinen, was neben IR auch disziplinäre Repositorien (s. Kap. 2.1) sowie Websites beinhaltet, die von Forschungsgruppen, Departments oder den Wissenschaftler:innen selbst betrieben werden. Die Studie besteht aus zwei Komponenten: (a) einer im Oktober 2006 durchgeführten *Online-Befragung* von 684 wissenschaftlichen Mitarbeitenden von insgesamt 17 verschiedenen Universitäten mit DSpace-basierten IR sowie (b) 41 *Nachfolgeinterviews*, die per Telefon im Zeitraum zwischen März und Mai 2007 stattfanden. Die Stichprobe der Online-Befragung ist repräsentativ in Hinblick auf vertretene Universitäten, Fachdisziplinen und Dienstgrade (Kim, 2010, S. 1909;1913-1914).

Abbildung 2 zeigt Kims modifiziertes Modell der Selbstarchivierungsbereitschaft wissenschaftlichen Personals:

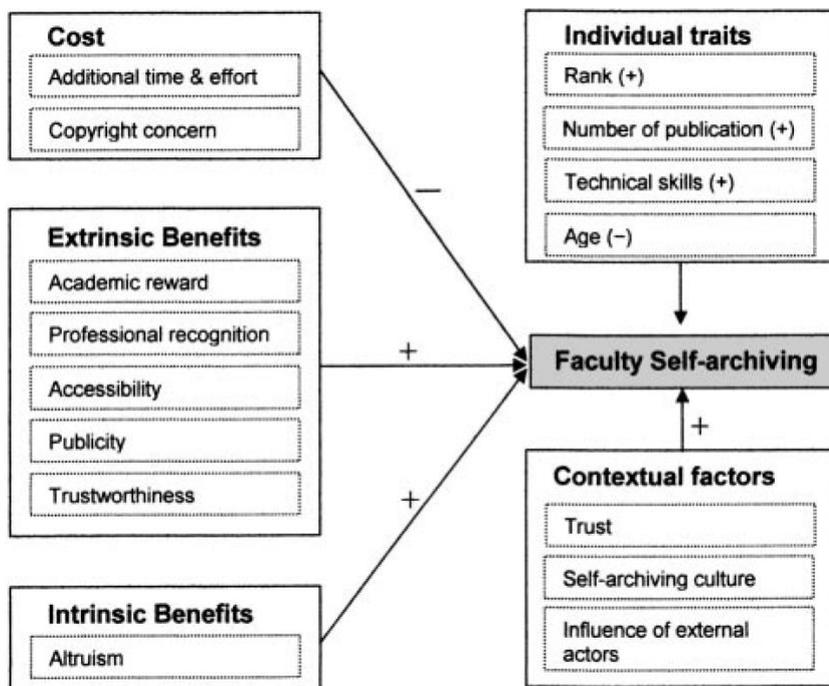


Abb. 2: Version 2 des Kim-Modells (Quelle: Kim, 2010, S. 1912, Abb. 1)

Angenommene positive bzw. negative Zusammenhänge zwischen den Variablen werden durch die Verwendung der Symbole „+“ bzw. „-“ ausgedrückt (Kim, 2010, S. 1911). Neben dem Wegfallens einer Fokussierung auf IR fällt beim Betrachten der AV auf, dass auch die zwischengeschaltete AV der Selbstarchivierungsmotivation fehlt. Gleichwohl diese Entscheidung von der Autorin nicht begründet wird, könnte sie auf das Bestreben zurückzuführen sein, das Modell vor der Analyse zu vereinfachen. Darüber hinaus ist auffällig, dass der Faktor der institutionellen Anerkennung bei den extrinsischen Gewinnfaktoren fehlt, obwohl er in der Pilotstudie als drittichtigster Faktor für das zukünftige Publizieren in IR identifiziert wurde (Kim, 2007, S. [7], Tab. 3). Diese Entscheidung könnte darin begründet sein, dass mit den Variablen der professionellen und institutionellen Anerkennung sehr ähnliche Konstrukte bzw. möglicherweise sogar dasselbe Konstrukt gemessen wird, sodass sich nicht sinnvoll zwischen den beiden Faktoren unterscheiden lässt. Auch der Zusatz des Selbstinteresses beim intrinsischen Gewinnfaktor des Altruismus wurde entfernt – möglicherweise, um die Benennung zu verkürzen, da der Begriff des Selbstinteresses in diesem Kontext lediglich die Abwesenheit von altruistischem Verhalten bezeichnet und nicht durch ein gesondertes Item abgefragt wird.

Auch die Individualfaktoren, die in der Pilotstudie (mit Ausnahme der demografischen Daten) keinerlei Auswertung erfuhren, werden sowohl verschlankt als auch konkretisiert: Während die Anzahl der Publikationen pro Jahr, die in der Pilotstudie noch Teil der AV-Operationalisierung war, eine nähere Bestimmung der Produktivitätsvariable darstellt, wird der Faktor einer editorischen Verantwortung bei der Herausgabe wissenschaftlicher Zeitschriften (Kim, 2007, S. [3]), der in der Pilotstudie nicht näher untersucht wurde, in der zweiten Version des Modells vollständig entfernt – vermutlich, weil das Charakteristikum zu spezifisch ist und nur von wenigen Teilnehmenden erfüllt wird. Neu hinzugefügt werden die Individualfaktoren des Alters sowie der technischen Fähigkeiten.

Die Variable der Selbstarchivierungserfahrung, die in der Pilotstudie unter die Individualfaktoren fiel, wird in der Nachfolgestudie zur AV umfunktioniert, neu operationalisiert als ein kontinuierliches fünfstufiges Likert-Item, das nach dem prozentualen Anteil selbstarchivierter Arbeiten (nichts, 1–25%, 26–50%, 51–75%, und 76–100%) fragt (Kim, 2010, S. 1913;1915). Die *Likert-Skala* gehört zu den psychometrischen Skalen und wird in der empirischen Sozialforschung zum Erfassen von Einstellungen oder Vorlieben auf Intervallskalenniveau eingesetzt. Die Verwendung einer Intervallskala zum Messen von Einstellungen ist umstritten, da die Befragten den Abstand zwischen den Skalenpunkten möglicherweise uneinheitlich wahrnehmen könnten. In den Sozialwissenschaften findet die Likert-Skala trotzdem häufig Anwendung, weil die Vorgabe des Benennens der Endpunkte eine für die Auswertung hinreichende Gleichabständigkeit erzeugt (Döring, 2023, S. 270-271; Porst, 2014, S. 76).

Weitere Auffälligkeiten von Kims neuer Modellierung beinhalten, dass der Kontextfaktor der Preprint-Kultur zu Selbstarchivierungskultur umbenannt wurde. Hinter dieser Entscheidung könnte die Motivation stehen, die Menge an potenziell zu archivierenden Objekten sowie deren Begutachtungsstatus weniger stark einzugrenzen. So nehmen die Items der Nachfolgestudie neben klassischen Preprints u. a. auch Bezug auf begutachtete Artikel, studentische Arbeiten oder technische Berichte (Kim, 2010, S. 1922). Gleichzeitig fällt der Kontextfaktor der institutionellen Identifikation, der in der Pilot-Studie nicht näher analysiert wurde, vollständig weg.

Obwohl die Kombination quantitativer und qualitativer Methoden (*Mixed-Methods*) wissenschaftstheoretisch diskutiert wird (Döring, 2023, S. 72), arbeitet Kim mit trianguliertem Datenmaterial, wobei der Fokus der Auswertung auf den

quantitativen Ergebnissen des Fragebogens liegt und die qualitativen Interviews lediglich der Kontextualisierung signifikanter Faktoren dienen. Die elf Kosten-, Gewinn- und Kontextfaktoren werden mit insgesamt 33 Likert-Items erfasst. Ehe die Items mithilfe von Summenbildungen zu Skalen zusammengefasst werden, nimmt Kim eine Reliabilitätsanalyse mittels *Cronbachs Alpha* vor, um die interne Konsistenz der Konstrukte zu überprüfen (2010, S. 1915-1916). Cronbachs Alpha ist ein Reliabilitätsmaß, das in der empirischen Wirtschafts- und Sozialforschung zum Messen der internen Konsistenz theoretischer Konstrukte verwendet wird und die durchschnittliche Korrelation zwischen den Items ausdrückt. Grundsätzlich wird eine hohe positive Korrelation erwartet, weil die Items Indikatoren desselben Konstruktes darstellen. Üblicherweise wird ein Konsistenzreliabilitätswert von >0,80 angestrebt, doch in der Literatur wird auch ein Wert von >0,70 als akzeptabel betrachtet (Döring, 2023, S. 274; Eckstein, 2016, S. 317; Venkatesh, 2003, S. 457).

Kim legt für ihre Auswertung einen Grenzwert von >0,70 fest und schließt drei Items mit niedrigen Alpha-Werten von der weiteren Analyse aus. Anschließend führt sie eine lineare Regressionsanalyse durch, um herauszufinden, inwieweit die AV durch die UV erklärt wird. Hierbei wendet sie die Methode der kleinsten Quadrate an, bei der die optimale Regressionslinie den kleinstmöglichen Abstand zu den Datenpunkten aufweist. Das Bestimmtheitsmaß R^2 , das auf die Güte der Regressionsanalyse hinweist, beträgt 0,55, sodass Kims Modell die Varianz in der AV zu 55% zu erklären vermag. An den p -Werten der Regressionskoeffizienten lässt sich ablesen, ob der gerichtete Zusammenhang zwischen den Variablen signifikant (<0,05) ist und sich die Nullhypothese zurückweisen lässt (Döring, 2023, S. 675; Kim, 2010, S. 1915; Kronthaler, 2021, S. 245;251).

Tabelle 4 zeigt die in der Regressionsanalyse signifikanten Faktoren. Positive bzw. negative Einflussbeziehungen auf die AV werden durch die Verwendung der Symbole „+“ bzw. „-“ ausgedrückt. Faktoren, die bereits in der Pilotstudie (s. Kap. 5.2.1) relevant waren, sind kursiv gesetzt.

Tab. 4: Signifikante Faktoren aus Kims Nachfolgestudie (Quelle: Kim, 2010, S. 1915, Tab. 7)

Gewinnfaktoren	Kostenfaktoren	Kontextfaktoren	Individualfaktoren
<i>Altruismus (+)</i>	<i>Copyrightbedenken (-)</i>	<i>Selbstarchivierungskultur (+)</i>	Technische Fähigkeiten (+)
Akademische Belohnung (+)	Zusätzlicher Zeit- und Arbeitsaufwand (-)		Alter (-)

Der Faktor des Altruismus übt den stärksten Einfluss auf die AV aus, während die Faktoren des zusätzlichen Zeit- und Arbeitsaufwands sowie der akademischen Belohnung (*academic reward*) den geringsten Einfluss ausüben. Neben der Altruismus-Variable waren auch Copyrightbedenken sowie der Faktor der Selbstarchivierungskultur bereits in Kims Pilotstudie (2007) von Relevanz. Wie in den Nachfolgeinterviews zutage kam, kann beim Faktor der Selbstarchivierungskultur auch wahrgenommener Gruppenzwang eine Rolle spielen, der Professor:innen aus Fachbereichen mit einer starken Selbstarchivierungskultur zum freien Publizieren bewegt (Kim, 2010, S. 1917). Um Unterschiede zwischen den Natur-, Geistes-, Sozial- und Ingenieurwissenschaften bezüglich der wahrgenommenen Selbstarchivierungskultur feststellen zu können, hat Kim zusätzlich eine einfaktorielle Varianzanalyse (ANOVA) durchgeführt, deren Ergebnis zeigt, dass der Mittelwert dieser Variable in den Geisteswissenschaften signifikant niedriger ist als in den übrigen drei Gruppen (2010, S. 1916-1917).

An Kims Operationalisierung ist kritisch anzumerken, dass ihre Konstrukte der akademischen Belohnung (*academic reward*), öffentlichen Aufmerksamkeit (*publicity*) sowie professionellen Anerkennung (*professional recognition*) inhaltlich sehr nahe beieinanderliegen und sich dabei überschneiden könnten. Eine unzureichende Trennschärfe der Konstrukte birgt die Gefahr, dass in der Studie nicht das gemessen wurde, was gemessen werden sollte (mangelnde *Konstruktvalidität*; Döring, 2023, S. 94). Auch stellt sich die Frage, inwieweit es sinnvoll ist, den Faktor der Zugänglichkeit (*accessibility*) als eigenständigen potenziellen Motivator in das Modell aufzunehmen, weil die reine Zugänglichmachung der eigenen Forschungsergebnisse ihrerseits von den Gewinnfaktoren des Altruismus oder der professionellen Anerkennung motiviert werden könnte. Sollten derartige Einflussbeziehungen vorhanden sein, bestünde die Gefahr einer *Multikollinearität*, bei der starke Korrelationen zwischen den UVs existieren, was zu einer Unterschätzung ihres Erklärungsbeitrags führen könnte (Eckstein, 2016, S. 216; Kronthaler, 2021, S. 272-273).

Auch die Operationalisierung des Konstruktes der akademischen Belohnung wirft Fragen auf, indem die nach Durchführung der Reliabilitätsanalyse verbleibenden zwei Items – entgegen Kims positiver Einordnung des Konstruktes als Gewinnfaktor – lediglich nach der Einschätzung negativer Folgen einer Selbstarchivierung für die Aussicht auf eine Festanstellung (*tenure*) oder Fördergelder fragen (Kim, 2010, S. 1922). So rührt die Signifikanz der Variable nicht etwa daher, dass die Befragten einen akademischen Vorteil in der Selbstarchivierung sehen, sondern ist

darauf zurückzuführen, dass die Teilnehmenden im Falle einer Selbstarchivierung mit keinerlei Sanktionen rechnen.

Die Operationalisierung von Kims Vertrauensvariable, die mit insgesamt drei Items erfasst wird, ist ebenfalls überarbeitungswürdig: Ein Item enthält das Statement, dass ein freies Publizieren von Forschungsmaterialien „die Integrität der Arbeit kompromittiert“ (Kim, 2010, S. 1922). Weil zu vermuten ist, dass die Teilnehmenden diese Begrifflichkeit nicht einheitlich interpretieren, ist dieses Item nicht hinreichend definiert. Die übrigen zwei Items fragen einerseits nach der Furcht der Teilnehmenden vor Plagiaten und andererseits nach einem möglicherweise mangelnden Vertrauen in die Archivierungsarchitektur öffentlich zugänglicher Websites. Hiermit werden zwei gänzlich verschiedene Konstrukte abgebildet, die dennoch in der Vertrauensvariable zusammengefasst werden. Anstatt eine gemeinsame mehrdimensionale Skala zu bilden, wäre es sinnvoller, diese beiden Items als eigenständige Faktoren in das Modell einfließen zu lassen.

5.2.3 Abschlussstudie (2011)

Wie bereits in Kapitel 5.2 erwähnt, nutzt Kim in ihrer Abschlussstudie das Datenmaterial der vorhergehenden Studie aus dem Jahr 2010 nach (2011, S. 248-249). Während die Vorgängerstudie die Publikationsbereitschaft wissenschaftlichen Personals im Allgemeinen untersucht hat, verengt die Abschlussstudie ihren Fokus auf die Bereitschaft derselben Zielgruppe zur Selbstarchivierung in IR und greift damit die Forschungsfrage aus der Pilotstudie (2007; s. Kap. 5.2.1) erneut auf.

Abbildung 3 zeigt Kims modifiziertes Modell der Selbstarchivierungsbereitschaft wissenschaftlichen Personals in IR:

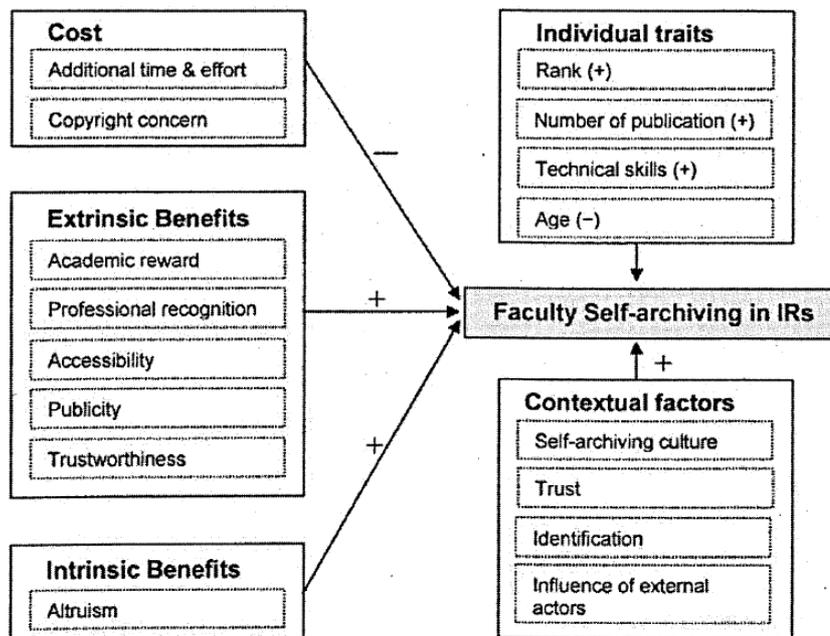


Abb. 3: Version 3 des Kim-Modells (Quelle: Kim, 2011, S. 247, Abb. 1)

Um das Modell an den verengten Forschungsgegenstand anzupassen, wurde die AV auf die Publikationsbereitschaft in IR beschränkt, operationalisiert als ein dichotomes Item (ja/nein), welches abfragt, ob die Teilnehmenden bereits Forschungsmaterialien in IR deponiert haben oder nicht (Kim, 2011, S. 249). Bei der Betrachtung der UV fällt auf, dass der Kontextfaktor der institutionellen Identifikation, der bereits Bestandteil der ersten Modellversion war und in der Nachfolgestudie unbegründet gestrichen wurde (s. Kap. 5.2.1;5.2.2), in der Abschlussstudie wieder aufgenommen wurde. Diese Entscheidung könnte darin begründet sein, dass die persönliche Identifikation mit der eigenen Institution von besonderer Relevanz für die Entscheidung zur Selbstarchivierung in IR sein könnte.

Während es sich bei der AV in der Nachfolgestudie um eine stetige Variable handelte, liegt in der Abschlussstudie eine dichotome AV vor, weshalb Kim hier eine logistische Regressionsanalyse durchführt. Ein Hosmer-Lemeshow-Test zeigt, dass Kims Modell die erhobenen Daten zufriedenstellend zu erklären vermag (hohe *Anpassungsgüte*; Kim, 2010, S. 1915; 2011, S. 249-250). Tabelle 5 zeigt die festgestellten signifikanten Faktoren, wobei positive bzw. negative Einflussbeziehungen auf die AV erneut durch die Verwendung der Symbole „+“ bzw. „-“ ausgedrückt werden.

Tab. 5: Signifikante Faktoren aus Kims Abschlussstudie (Quelle: Kim, 2011, S. 250, Tab. 4)

Gewinnfaktoren	Kostenfaktoren	Kontextfaktoren	Individualfaktoren
<i>Altruismus (+)</i>	<i>Copyrightbedenken (+)</i>	Vertrauen (+)	Akademischer Rang (+)
Zugänglichkeit (+)			

Die Variablen des Altruismus sowie der Copyrightbedenken waren nicht nur bereits in der Pilotstudie relevant, sondern auch in der Nachfolgestudie signifikant, weshalb sie in Tabelle 5 kursiv gesetzt sind. Auffällig ist, dass die Copyrightbedenken – im Gegensatz zu den Ergebnissen der Nachfolgestudie – einen positiven Einfluss auf die AV ausüben. Dies könnte laut Kim darauf zurückzuführen sein, dass die Befragten den Betreibenden von IR eine höhere Kompetenz im Rechtemanagement zusprechen als denen anderer Plattformen (2011, S. 252-253). Mit einem Quotenverhältnis (*Odds Ratio*) von 1,30 übt der Faktor einer verbesserten Zugänglichkeit zu Forschungsergebnissen den größten Einfluss auf die AV aus. Neben der Vertrauensvariable, deren Operationalisierung im vorhergehenden Kapitel kritisiert wurde, ist auch der Faktor des akademischen Ranges erstmals signifikant, was darauf hinweist, dass die Gegebenheit einer Festanstellung (*tenure*) besonderen Einfluss auf die Entscheidung wissenschaftlichen Personals zur Selbstarchivierung in IR nimmt (Kim, 2011, S. 250). Die Auswertung einer wiederholten Abfrage potenzieller Motivatoren für das Publizieren in IR reproduzierte die Ergebnisse der Pilotstudie (s. Kap. 5.2.1), welche den Wunsch nach einer Langzeitarchivierung sowie einer Bereitstellung von Nutzungsstatistiken als wichtigste Motivatoren identifizierte (Kim, 2011, S. 250, Tab. 3).

Eine Zusammenschau der Ergebnisse von Kims Nachfolge- und Abschlussstudie legt offen, dass mit den Faktoren des Altruismus sowie der Copyrightbedenken lediglich zwei Variablen konstant signifikant sind. Um zu ergründen, warum sich in der Abschlussstudie unerwartet ein positiver Zusammenhang zwischen der AV und der Copyrightvariable zeigt, wäre es vor dem Hintergrund von Kims Interpretation sinnvoll gewesen, einen möglichen Zusammenhang zwischen Copyrightbedenken und der Vertrauensvariable zu untersuchen, wovon Kim jedoch absieht. Auch einem zu vermutenden Zusammenhang zwischen der signifikant erhöhten Selbstarchivierungsbereitschaft festangestellten Personals und dem Faktor der institutionellen Identifikation geht Kim nicht nach.

Darüber hinaus ist Kims Operationalisierung der AV als dichotome Variable zu hinterfragen: Im Zuge der logistischen Regressionsanalyse wird versucht, die Wahrscheinlichkeit einer in der Vergangenheit vorgenommenen Selbstarchivierung anhand der UV vorherzusagen. Gemäß des Forschungsinteresses wird diese Selbstarchivierungserfahrung mit einer bestehenden Motivation zur Selbstarchivierung gleichgesetzt. Dementsprechend werde Teilnehmende, die noch keine Forschungsmaterialien in IR deponiert haben, als unmotiviert eingestuft. Ob eine derartige Gleichsetzung möglich ist, ist jedoch fraglich, da mangelnde Publikationserfahrung auch auf andere Faktoren als Unmotiviertheit zurückzuführen sein könnte. So war z. B. nur 40,1% der Befragten bewusst, dass ihre Institution überhaupt über ein IR verfügt (Kim, 2011, S. 249). Personen, die der Selbstarchivierung aufgeschlossen gegenüberstehen, aber nicht über die Existenz von IR informiert sind, werden auf diese Weise als unmotiviert eingestuft. In ihrem Pilotmodell ist Kim dieser Problematik ausgewichen, indem sie – möglicherweise in Anlehnung an die UTAUT – die Motivation zur Selbstarchivierung als eigenständige Variable Einfluss auf das tatsächliche Publikationsverhalten nehmen ließ (s. Kap. 5.2.1). Um ihrer gesetzten Forschungsfrage nach den Motivatoren zur Selbstarchivierung sinnvoll nachgehen zu können, wäre eine stärkere Differenzierung in der Operationalisierung vonnöten gewesen, ob durch eine zweigeteilte oder eine einzelne stetige AV.

In den letzten drei Kapiteln wurde die Entwicklung des Kim-Modells zur Publikationsbereitschaft wissenschaftlichen Personals in IR dargestellt und evaluiert. Die abschließende Version des Modells (s. Abb. 3) stellt die Grundlage für die in dieser Arbeit vorgenommene Modellierung der Publikationsbereitschaft von Studierenden in IR dar und wird im Folgenden entsprechend weiterentwickelt.

5.3 Auswahl der Variablen

In diesem Kapitel werden Kims Variablen unter Einbezug der bestehenden Forschung zu studentischen Inhalten in IR (s. Kap. 2.2;3) angepasst sowie ergänzt bzw. verworfen. Nach Kims Vorbild findet dabei eine Gruppierung nach Gewinn-, Kosten-, Kontext- und Individualfaktoren statt.

5.3.1 Gewinnfaktoren

Kim identifiziert insgesamt sechs Gewinnfaktoren, die positiven Einfluss auf die AV ausüben und im Folgenden diskutiert werden. Die Differenzierung zwischen extrinsischen und intrinsischen Gewinnfaktoren, deren Mehrwert Kim an keiner Stelle näher erläutert, wird für die in dieser Arbeit vorgenommene Neumodellierung nicht übernommen. Weil sich die ohnehin schon nahe beieinanderliegenden Konstrukte der akademischen Belohnung und der öffentlichen Aufmerksamkeit nach einer Überarbeitung für die Zielgruppe der Studierenden nicht sinnvoll trennen lassen, werden sie für die Neumodellierung in der neuen Variable des schulischen Fleißes zusammengeführt. Der Faktor der professionellen Anerkennung wird auf die Situation von Studierenden angepasst, indem das Studium als mögliche Grundlage für künftigen beruflichen Erfolg betrachtet wird, wobei der Wunsch nach einer wissenschaftlichen Karriere als Moderator fungiert. Um dieses Konstrukt treffender zu beschreiben, wird es in Karrierestreben umbenannt. Der Faktor der Zugänglichkeit (*accessibility*), der ohnehin nur schwer von anderen Gewinnfaktoren zu trennen ist, wird ersatzlos gestrichen, um die in Kapitel 5.2.2 benannte Gefahr einer Multikollinearität zu reduzieren. Darüber hinaus wird die Variable der persönlichen Anerkennung neu hinzugefügt.

(a) Altruismus

Der intrinsische Gewinnfaktor des Altruismus ist die einzige Variable, die in Kims Forschung konstant signifikant bleibt, ohne dabei – wie der Faktor der Copyrightbedenken – ihre Einflussrichtung zu ändern. Wie in den Interviews der Nachfolgestudie zutage kam, verfügt das Altruismus-Konstrukt über zwei sich gegenseitig ergänzende Facetten:

- (i) *Forschungsergebnisse als öffentliches Gut*: Um den wissenschaftlichen Fortschritt zu beschleunigen, sehen die sich Interviewten in der moralischen Verpflichtung, ihre Forschungsergebnisse frei zu teilen. Insbesondere Menschen in Entwicklungsländern profitieren von einer gleichberechtigten Zugänglichkeit, was einen Beitrag zur Überbrückung der digitalen Spaltung¹

¹ Der Begriff der digitalen Spaltung (*digital divide*) beschreibt die Kluft zwischen jenen Menschen, die Zugang zu digitalen Medien haben und von diesen Gebrauch machen, sowie jenen Menschen, denen eine solche Nutzung aus sozioökonomischen Gründen erschwert

leistet. Diese Einstellung, von Kim bezeichnet als „the [s]pirit of OA“ (2010, S. 1916), fällt mit den Zielen der in Kapitel 1 dargestellten Open-Access-Bewegung zusammen. Auch für die von Pickton und McKnight befragten Promotionsstudierenden ist das Unterstützen einer Philosophie der Offenheit im Wissenschaftskontext ein häufiger Motivator zur Selbstarchivierung (2006, S. 212; s. Kap. 3).

- (ii) *Hilfsbereitschaft*: Einige der Interviewten gewinnen allein durch die Aussicht an Zufriedenheit, dass sie anderen Menschen durch das Teilen ihrer Materialien beim Lernen und Forschen unterstützen könnten (Kim, 2010, S. 1916). Eine weitere Komponente dieser Facette ist die Idee der Reziprozität: Gemäß der ‚goldenen Regel‘ drücken Interviewte, die bereits von den frei zugänglichen Forschungsmaterialien anderer profitiert haben, den Wunsch aus, im Gegenzug ihre eigenen Arbeiten zu teilen (Kim, 2010, S. 1916;1918). Dieser Gedanke ist, wie auch Kim feststellt, nicht mit reinem Altruismus gleichzusetzen, doch weil die beiden Ideen konzeptionell nur schwierig zu differenzieren sind, werden sie auch in der vorliegenden Arbeit unter dem Altruismus-Konstrukt subsumiert.

Beide Facetten des Altruismus-Konstruktes sind ohne weitere Anpassung auf Studierende übertragbar.

(b) Schulischer Fleiß

Die neue Variable des schulischen Fleißes stellt eine Kombination aus Kims Faktoren der akademischen Belohnung (*academic reward*) sowie der öffentlichen Aufmerksamkeit (*publicity*) dar, was in diesem Abschnitt begründet wird. Weil Kims Zielgruppe wissenschaftliches Personal umfasst, stehen bei der Variable der akademischen Belohnung für sie die Themen der Festanstellung (*tenure*) und Beförderung im Vordergrund, die im akademischen Belohnungssystem vorrangig von der Qualität und Quantität wissenschaftlicher Publikationen abhängen (Kim, 2011, S. 248). Um die Idee eines akademischen Belohnungsanreizes auf Studierende zu übertragen, wird das Konstrukt dahingehend überarbeitet, dass sich zur Selbstarchivierung motivierte Studierende hiervon einen erhöhten Erfolg in ihrem Studium erhoffen könnten. Diese Annahme stützt sich auf die Ergebnisse von Pickton und

wird. Die Problematik wird häufig vor dem Hintergrund globaler Differenzen zwischen Industrie- und Entwicklungsländern diskutiert (van Dijk, 2020, S. 1-2).

McKnight (s. Kap. 3), die den Wunsch nach konstruktivem Feedback als wichtigsten Motivator der befragten Promotionsstudierenden identifizieren. Das Teilen der eigenen Arbeit setzt – so die Hoffnung der Studierenden – einen kommunikativen Austausch in Gang, der die Qualität ihrer Arbeit steigert (Pickton & McKnight, 2006, S. 212-213).

Im Zentrum von Kims Variable der öffentlichen Aufmerksamkeit (*publicity*) steht der Wunsch des wissenschaftlichen Personals nach einer schnellen und weiten Dissemination ihrer Arbeit, was mit der Hoffnung auf eine erweiterte Leserschaft sowie eine erhöhte Wirkkraft ihrer Forschungsergebnisse verknüpft ist (Kim, 2011, S. 247). Auch für die von Pickton und McKnight interviewten Promotionsstudierenden stellt die Aussicht darauf, die eigene Arbeit innerhalb der Forschungsgemeinschaft sowie mit Mitarbeiter:innen und anderen Studierenden teilen zu können, einen wichtigen Grund für die Selbstarchivierung dar (s. Kap. 3, Tab. 1). Wie die Autor:innen vermuten, unterscheidet sich die hinter dem Wunsch nach Dissemination stehende Motivation der Studierenden jedoch von jener des wissenschaftlichen Personals: Während die Wissenschaftler:innen mit ihrer Leserschaft auch ihre Zitationsrate zu erhöhen hoffen, werden die Studierenden von der Aussicht auf Feedback motiviert (Pickton & McKnight, 2006, S. 213). Weil sich sowohl der überarbeitete Faktor der akademischen Belohnung als auch jener der öffentlichen Aufmerksamkeit für die Zielgruppe der Studierenden auf den Wunsch nach Feedback zurückführen lässt, werden die Konstrukte an dieser Stelle in der Variable des schulischen Fleißes zusammengeführt.

(c) Karrierestreben

Mit dem Faktor der professionellen Anerkennung (*professional recognition*) untersucht Kim, inwieweit die befragten Wissenschaftler:innen davon ausgehen, dass eine Selbstarchivierung ihre Sichtbarkeit in ihrer eigenen Disziplin erhöht (2011, Anhang A). Vor dem Hintergrund der in Kapitel 2.2.2 zusammengetragenen Literatur, die erhöhte Chancen auf dem Arbeitsmarkt für studentische Publizierende vermutet, wird die fortan in Karrierestreben umbenannte Variable für die Neumodellierung dahingehend angepasst, dass sie untersucht, inwiefern die Studierenden bereits im Studium die Grundlage für ihren künftigen beruflichen Erfolg legen möchten. Es lässt sich annehmen, dass die Stärke des Einflusses dieser UV davon abhängt, ob die Studierenden eine wissenschaftliche Karriere in Erwägung ziehen oder nicht. Dementsprechend wird der Neumodellierung das Anstreben einer wis-

senschaftlichen Karriere als Moderator (Hayes, 2022, S. 234) hinzugefügt, der mit der UV des Karrierestrebens interagiert: Ein stärkerer Wunsch nach einer dezidiert wissenschaftlichen Karriere verstärkt den positiven Einfluss, den das Karrierestreben auf die Bereitschaft zur Selbstarchivierung ausübt.

(d) Persönliche Anerkennung

Am Ende ihrer strukturierten Interviews gaben Pickton und McKnight den Befragten im Rahmen einer offenen Frage die Möglichkeit, weitere Motivatoren für eine Selbstarchivierung zu nennen (s. Kap. 3). Das Hinzufügen des Faktors der persönlichen Anerkennung basiert auf den Antworten der Promotionsstudierenden, die Gründe nannten wie „to show what the student’s time has been spent on, and to provide evidence for appraisals; to prove oneself; for personal satisfaction (including seeing one’s name in print); [and] for peer acceptance“ (2006, S. 208).

(e) Vertrauenswürdigkeit

Mit dem Faktor der Vertrauenswürdigkeit (*trustworthiness*) untersucht Kim jene sozialen Prozesse, welche die Qualität der veröffentlichten Dokumente anhand gemeinschaftlich festgelegter Kriterien sicherstellen. Das klassische Instrument der Qualitätssicherung ist der Peer-Review, doch auch die akademische Reputation der Autor:innen oder der mit ihnen affilierten Institutionen kann zur Evaluation herangezogen werden (Kim, 2011, S. 247-248). In seinem Grundlagentext zu IR lehnt Lynch die Übernahme solcher klassischen Richtlinien als „gate keeping“ ab (2003, S. 333; s. Kap. 2.1). Hiergegen lässt sich argumentieren, dass gewisse Publikationsrichtlinien zum Schutz der Studierenden unerlässlich sind: Wie in Kapitel 2.2.2 dargestellt, ist eine mögliche Kompromittierung der Online-Identität ein potenzieller Nachteil der Selbstarchivierung für Studierende, dem sich u. a. mit einer sorgfältigen Qualitätssicherung begegnen lässt. In Pickton und McKnights Interviews kam zutage, dass einige Promotionsstudierende besorgt über die Qualität ihrer Arbeiten sind und ihre Publikationsbereitschaft daher von einer vorherigen Begutachtung durch ihre Betreuer:innen abhängig machen (2006, S. 209;214; s. Kap. 3). Hieraus lässt sich ableiten, dass der Aspekt der Qualitätssicherung für Studierende ebenso relevant ist wie für das wissenschaftliche Personal und sich lediglich die konkreten Umsetzungsmaßnahmen voneinander unterscheiden. So wird der Peer-Review im Falle studentischer Inhalte häufig durch die Begutachtung und Empfeh-

lung durch eine Lehrperson sowie die Anforderung einer Mindestnote ersetzt (HAW Hamburg, o. J.a; Hertenstein, 2014, S. 10-11).

5.3.2 Kostenfaktoren

Kim identifiziert insgesamt zwei Kostenfaktoren, die negativen Einfluss auf die AV ausüben, was im Folgenden diskutiert wird. Ausgehend von der in Kapitel 5.2.2 geäußerten Kritik an Kims Operationalisierung des Kontextfaktors des Vertrauens (*trust*) wird die Furcht vor Plagiaten in der Neumodellierung als eigenständige UV aufgefasst, die negativen Einfluss auf die AV ausübt und unter die Kostenfaktoren subsumiert wird. Die zweite Facette von Kims Vertrauenskonstrukt, die das Vertrauen in die Archivierungsarchitektur öffentlich zugänglicher Websites untersucht, wird für die Neumodellierung hingegen nicht übernommen, weil Studierende dem Aspekt der Langzeitarchivierung eine weitaus geringere Bedeutung beimessen als Wissenschaftler:innen. Tatsächlich zeichnet sich an diesem Punkt vielmehr eine gegenteilige Anreizstruktur ab, da das wissenschaftliche Personal gerade die nachhaltige digitale Archivierung ihrer Inhalte als wichtigen Publikationsmotivator identifiziert, während Studierende die Option des nachträglichen Löschsens ihrer Arbeiten hoch priorisieren (Kim, 2011, S. 252; Pickton & McKnight, 2006, S. 215).

Obwohl sich die Einflussrichtung in Kims Abschlussstudie unerwartet ins Positive wendete, ist die Variable der Copyrightbedenken gemeinsam mit dem Altruismus-Faktor als einzige Variable konstant signifikant (s. Kap. 5.2.2;5.2.3). Die befragten Wissenschaftler:innen befürchten einerseits, bereits selbstarchivierte Inhalte nicht mehr anderweitig publizieren zu können, und andererseits, mit einer Zweitveröffentlichung in einem Repositorium gegen bestehendes Copyright zu verstoßen (Kim, 2011, S. 247;Anhang A). Beide Facetten der Copyrightbedenken lassen sich auch bei den von Pickton und McKnight befragten Promotionsstudierenden finden. Hierbei ist jedoch die Besonderheit der Stichprobe zu beachten, die allein Studierende umfasst, die auf einen Dokortitel hinarbeiten und entsprechend darum besorgt sind, sich mit einer Selbstarchivierung die Chance auf das Publizieren ihrer Ergebnisse in traditionellen Zeitschriften zu vereiteln (2006, 212;214). Abseits von Promotionsstudierenden ist hingegen nicht davon auszugehen, dass Copyrightbedenken einen nennenswerten Einfluss auf die Selbstarchivierungsbereitschaft ausüben, weshalb dieser Faktor für die Neumodellierung nicht übernommen wird.

(a) Zeit- und Aufwandsinvestition

Mit der Variable der Zeit- und Aufwandsinvestition untersucht Kim, inwieweit die zusätzliche Arbeit, die durch ein selbstständiges Deponieren von Inhalten in IR erzeugt wird, das wissenschaftliche Personal von der Selbstarchivierung abhalten könnte (2011, S. 247). Dass Pickton und McKnight diese Befürchtung teilen, tritt in ihren abschließend formulierten Handlungsempfehlungen zutage, in denen sie die Einrichtung eines vom Bibliothekspersonal angebotenen Services vorschlagen, der den Upload der Arbeiten für unsichere oder zeitlich eingeschränkte Promotionsstudierende übernimmt (2006, S. 216). Weil auch Studierende das Gefühl haben könnten, dass das Publizieren in IR einen zu hohen Aufwand für sie bedeutet, wird die Variable unverändert für die Neumodellierung übernommen.

(b) Plagiatsfurcht

Eine häufig auftretende Sorge, die Kim insbesondere unter Geisteswissenschaftler:innen beobachtet, ist die Furcht, dass ihre frei in IR publizierten Erkenntnisse von anderen Forschenden desselben Feldes ohne Quellenangabe in traditionellen Publikationen verwendet und als die eigenen ausgegeben werden könnten (2011, S. 247). Wie Pickton und McKnight feststellen, tritt dieser „professionelle Schutzinstinkt“ (2006, S. 214) auch bei Promotionsstudierenden auf, wenn sie die Furcht vor unerlaubtem Kopieren bzw. Plagiaten ihrer Arbeiten als wichtige Faktoren identifizieren, die sie von der Selbstarchivierung in IR abhalten (s. Kap. 3). Es lässt sich vermuten, dass auch Studierende, die kein traditionelles Publizieren anstreben, eine Furcht vor unerwünschtem Kopieren ihrer Arbeiten durch andere Studierende empfinden könnten, weshalb dieser Faktor als eigenständige UV in die Neumodellierung aufgenommen wird.

5.3.3 Kontextfaktoren

Kim identifiziert insgesamt vier Kontextfaktoren, die positiven Einfluss auf die AV ausüben, was im Folgenden diskutiert wird. Eine für die Neumodellierung überarbeitete Version der Vertrauensvariable wurde bereits unter die Kostenfaktoren subsumiert (s. Kap. 5.3.2). Der Faktor der externen Akteur:innen wird nicht für die Neumodellierung übernommen, weil von den von Kim diskutierten Akteur:innen (2011, S. 248) lediglich Repräsentant:innen der Hochschule für Studierende von

Relevanz sind. Wie Pickton und McKnight herausfinden, stellt die Ermutigung durch das betreuende Lehrpersonal häufig einen Publikationsmotivator für die interviewten Promotionsstudierenden dar (2006, S. 212; s. Kap. 3). Weil dieser Aspekt einer Qualitätssicherung durch Lehrende jedoch bereits im Kontext des Gewinnfaktors der Vertrauenswürdigkeit untersucht wird (s. Kap. 5.3.1), wird davon abgesehen, ihn als eigenständige UV in die Neumodellierung aufzunehmen.

Der Faktor einer wahrgenommenen Preprint- bzw. Selbstarchivierungskultur (s. Kap. 5.2.1) ist in Kims Nachfolgestudie signifikant, verliert diese Signifikanz jedoch in der Abschlussstudie, was möglicherweise auf die bereits in Kapitel 5.2.3 kritisierte Operationalisierung der AV als dichotome Variable zurückzuführen sein könnte. Pickton und McKnight stellen fest, dass Promotionsstudierende der Natur- und Ingenieurwissenschaften in einem deutlich höheren Maße dazu bereit sind, ein Selbstarchivierungsmandat ihrer Hochschule zu befolgen, als es Studierende der Sozial- und Geisteswissenschaften sind (2006, S. 211). Wie beim Copyright-Kostenfaktor (s. Kap. 5.3.2) ist jedoch auch hier die Besonderheit der Stichprobe zu beachten: Weil nicht davon auszugehen ist, dass in einer Stichprobe, die nicht ausschließlich oder vorrangig aus Promotionsstudierenden besteht, die in den jeweiligen Fachdisziplinen vorherrschende Selbstarchivierungskultur bereits bewusst wahrgenommen wird, wird davon abgesehen, diesen Faktor für die Neumodellierung zu übernehmen.

(a) Identifikation

Mit dem Faktor der Identifikation untersucht Kim Loyalitäts- und Zugehörigkeitsgefühle des wissenschaftlichen Personals gegenüber ihrer Hochschule (2011, S. 247; s. Kap. 5.2.1). Weil IR als Schaufenster der intellektuellen Arbeit einer Institution dienen (Bankier & Smith, 2010, S. 245; s. Kap. 2.1), erscheinen die dort publizierten Arbeiten klar unter dem Markenzeichen der sie betreibenden Institution. Auch die Selbstarchivierungsbereitschaft von Studierenden könnte von dem Grad ihrer Identifikation mit ihrer Hochschule beeinflusst werden, weshalb die Variable als einziger Kontextfaktor für die Neumodellierung übernommen wird.

5.3.4 Individualfaktoren

Kim untersucht insgesamt vier Individualfaktoren, die sich nicht sinnvoll auf die Zielgruppe der Studierenden übertragen lassen, weshalb dieses Faktorenssegment für die Neumodellierung vollständig gestrichen wird. Während sich die Variable des Ranges z. B. in Form des angestrebten bzw. höchsten Abschlusses oder des aktuellen Fachsemesters erfassen ließe, ist diese Information lediglich für die Stichprobenkontrolle, nicht jedoch für die gesetzte Fragestellung zur Publikationsbereitschaft in IR von Interesse. Weil anzunehmen ist, dass nur wenige Studierende bereits über Selbstarchivierungserfahrung verfügen, ist auch die Anzahl der schon erfolgten Publikationen kein sinnvoller Indikator.

Darüber hinaus hat die Variable der technischen Fähigkeiten seit dem Jahr 2011, in dem Kim ihre Abschlussstudie durchführte, an Relevanz verloren: Gleichwohl nach wie vor zu vermuten ist, dass zwischen den Studierenden Unterschiede bezüglich ihrer technischen Versiertheit bestehen, dürften die meisten von ihnen vertraut genug im allgemeinen Umgang mit digitalen Plattformen und Anwendungen sein, um die Annahme einer hervorgehobenen Einflussnahme dieses Faktors auf die AV nicht länger zu rechtfertigen. Abschließend ist auch das Alter kein sinnvoller Indikator für die Neumodellierung, weil Studierende (ungeachtet möglicher Ausreißer) eine homogen junge Zielgruppe darstellen.

5.3.5 Abhängige Variable

Weil die von Kim untersuchte Publikationsmotivation eine latente Variable ist, deren Ausprägungen sich nicht direkt beobachten lassen (Döring, 2023, S. 226), misst Kim stattdessen die Publikationsaktivitäten ihrer Teilnehmenden. Auf die Schwierigkeiten, die mit einer solchen Gleichstellung von Publikationsmotivation und -aktivität einhergehen können, wurde bereits in Kapitel 5.2.3 aufmerksam gemacht. Da nicht davon auszugehen ist, dass Studierende bereits über viel Publikationserfahrung verfügen, stellen ihre bisherigen Publikationsaktivitäten keinen sinnvollen Indikator ihrer Motivation zur Selbstarchivierung dar. Stattdessen wird für die Neumodellierung die Einstellung der Teilnehmenden zu einer zukünftigen Selbstarchivierung herangezogen – eine Thematik, die von Kim deskriptiv untersucht wird (2007, S. [3]; 2010, S. 1913; 2011, S. 250, Tab. 3), ohne dabei Eingang in ihr Regressionsmodell zu finden. So wird im Rahmen der Neumodellierung unter-

sucht, welche Faktoren die Bereitschaft der Studierenden zu einer zukünftigen Selbstarchivierung in IR beeinflussen.

Nachdem Kims Variablen begründet für die Zielgruppe der Studierenden angepasst wurden, wird nun das empirische Regressionsmodell präsentiert, wofür die im Verlauf dieses Kapitels aufgestellten Hypothesen zusammenfassend ausformuliert werden.

6 Hypothesen und empirisches Modell

Im Zuge der in Kapitel 5.3 vorgenommenen Variablenauswahl wurden insgesamt neun gerichtete Hypothesen inhaltlich erarbeitet, die in Tabelle 6 dargestellt werden:

Tab. 6: Hypothesen (Quelle: Eigene Darstellung)

H1	Je höher die Studierenden den Zeit- und Arbeitsaufwand des Selbstarchivierens einschätzen, desto weniger sind sie zur Selbstarchivierung in IR bereit. → <i>negativer Zusammenhang: Wenn time_effort steigt, dann sinkt selfarchive</i>
H2	Je mehr Furcht vor Plagiaten die Studierenden verspüren, desto weniger sind sie zur Selbstarchivierung in IR bereit. → <i>negativer Zusammenhang: Wenn plagiarism steigt, dann sinkt selfarchive</i>
H3	Je altruistischer die Studierenden veranlagt sind, desto mehr sind sie zur Selbstarchivierung in IR bereit. → <i>positiver Zusammenhang: Wenn altruism steigt, dann steigt selfarchive</i>
H4	Je mehr Wert die Studierenden auf die eigene schulische Leistung legen, desto mehr sind sie zur Selbstarchivierung in IR bereit. → <i>positiver Zusammenhang: Wenn diligence steigt, dann steigt selfarchive</i>
H5	Je mehr die Studierenden persönliche Anerkennung suchen, desto mehr sind sie zur Selbstarchivierung in IR bereit. → <i>positiver Zusammenhang: Wenn prestige steigt, dann steigt selfarchive</i>
H6	Je mehr Vertrauen die Studierenden in die qualitätssichernden Prozesse ihrer Institution haben, desto mehr sind sie zur Selbstarchivierung in IR bereit. → <i>positiver Zusammenhang: Wenn trustworthiness steigt, dann steigt selfarchive</i>

Tab. 6: Hypothesen (Fortsetzung)

H7a	<p>Je mehr die Studierenden bereits im Studium die Grundlage für ihren künftigen beruflichen Erfolg legen möchten, desto mehr sind sie zur Selbstarchivierung in IR bereit.</p> <p>→ <i>positiver Zusammenhang: Wenn career steigt, dann steigt selfarchive</i></p>
H7b	<p>Ein stärkerer Wunsch nach einer wissenschaftlichen Karriere (<i>acad_career</i>) verstärkt den positiven Einfluss, den das Karrierestreben (<i>career</i>) auf die Bereitschaft zur Selbstarchivierung in IR (<i>selfarchive</i>) ausübt.</p>
H8	<p>Je mehr sich die Studierenden mit ihrer Hochschule identifizieren, desto mehr sind sie zur Selbstarchivierung in IR bereit.</p> <p><i>positiver Zusammenhang: Wenn identification steigt, dann steigt selfarchive</i></p>

Die in den Hypothesen beschriebene Beziehungsstruktur zwischen den Variablen lässt sich anhand eines Diagramms grafisch visualisieren, bei dem die auch als *Knoten* bezeichneten Variablen mittels *Kanten*, d. h. ungerichteter Linien oder gerichteter Pfeile, miteinander in Beziehung gesetzt werden. Sind alle Kanten gerichtet, so wird auch das Gesamtmodell als *gerichtetes Modell* bezeichnet. Zudem ist das Modell *azyklisch*, wenn keine von einem Knoten ausgehenden Pfade (d. h. Sequenzen gerichteter Kanten) zu ihrem Anfangsknoten zurückkehren. In Anlehnung an Kausalmodelle werden auch Regressionsmodelle in Form von gerichteten azyklischen Diagrammen visualisiert, wobei zu beachten ist, dass sich Kausalwirkungen anhand statistischer Modelle nicht nachweisen lassen. So informieren existierende Kanten hier lediglich darüber, inwieweit sich die AV verändert, wenn der Wert der UV um einen bestimmten Betrag sinkt oder steigt. Ebenso drücken fehlende Kanten nicht etwa die Abwesenheit einer Kausalverbindung, sondern lediglich eine bedingte Unabhängigkeit aus, bei der die UV keine zusätzlichen Informationen über die AV zu liefern vermag (Arzheimer, 2016, S. 41-42; Kincaid, 2012, S. 51; Pearl 2000, S. 11-13;141; Urban & Mayerl, 2018, S. 21).

Das folgende gerichtete azyklische Diagramm visualisiert die in Tabelle 6 dargestellten Hypothesen über die vermuteten Einflussfaktoren auf die AV der Selbstarchivierungsbereitschaft von Studierenden in IR (Abb. 4):

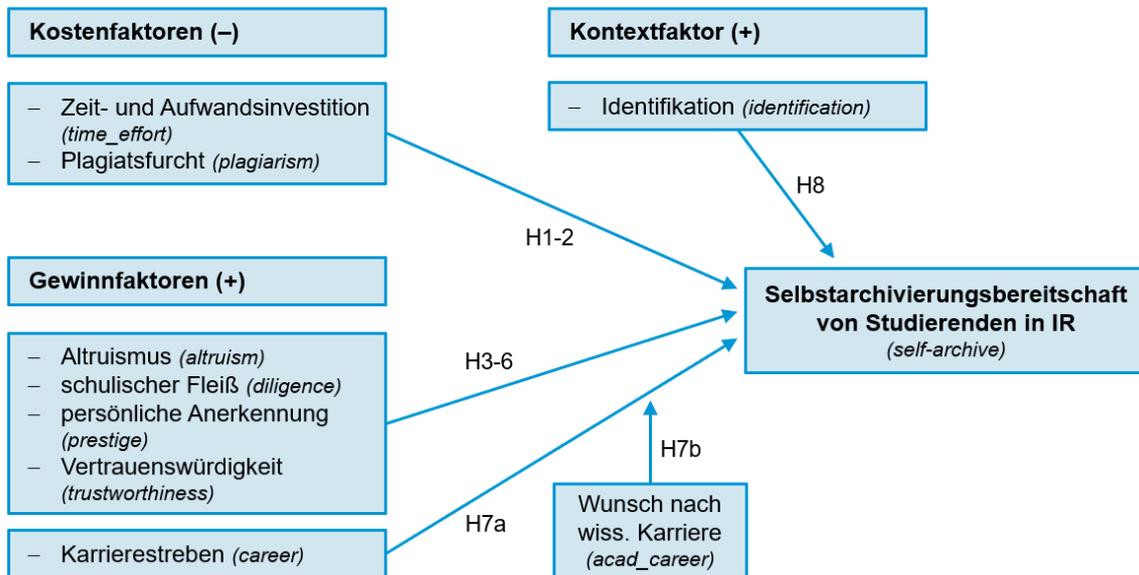


Abb. 4: Modell der Selbstarchivierungsbereitschaft von Studierenden in IR (Quelle: Eigene Darstellung, teilweise basierend auf Kim, 2011, S. 247, Abb. 1)

Nachdem die Hypothesen ausformuliert und visualisiert wurden, wird im nächsten Kapitel die Operationalisierung der Konstrukte vorgenommen.

7 Operationalisierung der Konstrukte

Bei der quantitativen Fragebogenmethode sollten Skalen nur dann selbst konstruiert werden, wenn die Fachliteratur keine geeigneten Instrumente hergibt (Döring, 2023, S. 270). Entsprechend wird auch bei der Operationalisierung der hier zu untersuchenden Konstrukte die Arbeit von Kim (2011, Anhang A) zugrunde gelegt, die bereits als Modellierungsvorlage diente (s. Kap. 5). Wie bei der Auswahl der Variablen ist auch bei der Auswahl und Formulierung der Likert-Items die studentische Zielgruppe zu beachten, die sich von Kims Zielgruppe wissenschaftlichen Personals unterscheidet, weshalb eine unveränderte Übernahme ihrer Items nicht möglich ist. Entsprechend werden Kims Items für die vollstrukturierte Befragung abgeändert und durch neue Items ergänzt, die auf den Ergebnissen von Pickton und McKnight (2006; s. Kap. 3) beruhen. Um die Antwortgenauigkeit zu erhöhen, werden alle Konstrukte – mit Ausnahme der AV und des Moderators – anhand von zwei bis drei Items erfasst, die vorbereitend für die Regressionsanalyse (s. Kap. 9.3) durch Mittelwertbildung zu einer gemeinsamen Skala zusammengefasst werden. Obwohl in der Psychometrie die Verwendung von fünfstufigen Skalen üblich ist, werden bewusst nur vier Stufen angesetzt, um den Befragten eine Positionierung zu entlo-

cken. Um Antworttendenzen zu vermeiden, enthält der Fragebogen sowohl positiv als auch negativ formulierte Items, wobei letztere im Rahmen der Datenaufbereitung umzupolen sind (Döring, 2023, S. 271).

Die folgenden Unterkapitel beschäftigen sich mit den fünf Frageblöcken des Fragebogens, wobei neben der Formulierung der Items auch ihre logische Anordnung sowie die standardisierte Ausfüllinstruktion begründet werden. Erläutert wird die finale Fassung des Fragebogens, sodass die Anpassungen, die im Rahmen des qualitativen Pretests vorgenommen wurden, bereits enthalten sind. Um welche Anpassungen es sich dabei genau handelt, wird in Kapitel 8.1 gebündelt aufgeführt und begründet. Jeder Frageblock entspricht einer eigenständigen Seite im Fragebogen. Die Tabellen 7 bis 17 bieten einen Überblick der mit den jeweiligen Konstrukten korrespondierenden Items inklusive ihrer Codierung. Der vollständig formatierte und visuell für die Zielgruppe aufbereitete Fragebogen ist Anhang A zu entnehmen.

7.1 Frageblock 1: Anerkennung

Zu Beginn ihrer strukturierten Interviews führten Pickton und McKnight die Studierenden inhaltlich in die Thematiken des OA-Publizierens sowie digitaler Repositorien ein (2006, S. 205). Hiervon wird bei der in dieser Arbeit durchgeführten Studie abgesehen, da spätere Forschung zu IR diesen ein Image-Problem attestiert, das insbesondere auf die Verwendung von Fachjargon, wie z. B. ‚Open Access‘ oder ‚institutionelles Repository‘, zurückzuführen ist. Auch die häufig mangelnde Bekanntheit von IR unter den Institutionsangehörigen könnte auf diese terminologische Schwierigkeit zurückzuführen sein, welche die Zielsetzung eines IR nicht adäquat zu vermitteln vermag. Entsprechend vermeiden viele IR es, sich nach außen hin als solche zu bezeichnen (Passehl-Stoddart & Monge, 2014, S. 8-9; Scherer, 2016, S. 164). Auch in dem hier erarbeiteten Fragebogen wird bewusst auf die Verwendung des unter Studierenden vermutlich weitgehend unbekanntem IR-Begriffes verzichtet. Stattdessen wird das Forschungsthema alltagssprachlich umschrieben, indem nach der Einstellung der Studierenden zum Veröffentlichen von Arbeiten, die sie im Rahmen ihres Studiums erstellt haben, gefragt wird. In Anlehnung an Kapitel 2.2.1 wird dabei beispielhaft aufgeführt, um welche Dokumentarten es sich handeln kann. Diese auf die Zielgruppe der Studierenden abgestimmte alltagssprachliche Tonalität wird konsequent sowohl in den Ausfüllinstruktionen als auch in den im Rahmen der Teilnehmendenakquise verschickten Ansprache-

texten verwendet, um die Teilnahme am Fragebogen so niedrigschwellig wie möglich zu gestalten und Studierende jeder Fachrichtung auch ohne Hintergrundwissen anzusprechen.

Nachdem das Forschungsthema skizziert und die Befragung formal als Bestandteil eines Masterprojekts an der HAW Hamburg ausgewiesen wurde, macht der Einleitungstext des Fragebogens darauf aufmerksam, dass es keine richtigen oder falschen Antworten gibt und alle Angaben anonym erfolgen sowie rein wissenschaftlichen Zwecken dienen. Als freiwillige Befragung ist das Projekt vollständig auf das Wohlwollen der angesprochenen Zielgruppe angewiesen, weshalb davon abgesehen wird, Pflichtfelder zu verwenden. Stattdessen werden die Teilnehmenden einleitend darüber informiert, dass das vollständige Ausfüllen des Fragebogens entscheidend für den Erfolg der Auswertung ist. Der Bearbeitungsumfang wird sowohl zeitlich mit ca. sieben Minuten als auch quantitativ mit 20 geschlossenen Ankreuzfragen angegeben (s. Kap. 8.1). Auf die Beschreibung dieses geringen Arbeitsaufwandes folgt unmittelbar die Incentivierung zur Teilnahme in Form einer Verlosung von drei Amazon-Gutscheinen im Wert von je 15 Euro. Abschließend wird eine persönliche Kontaktadresse genannt. Kernbegriffe werden fett gedruckt, damit auch beim Überfliegen des Textes alle wichtigen Informationen aufgenommen werden können.

Auf den Einleitungstext folgt eine Ausfüllinstruktion, welche die vier Stufen der Likert-Skala (überhaupt nicht, weniger, weitgehend, voll und ganz) benennt. Damit eine geringe Zustimmung einem geringen Wert entspricht, verläuft die Skala von „stimme überhaupt nicht zu“ (1) bis „stimme voll und ganz zu“ (4). Die Phrase „stimme [...] zu“ wird dabei konsequent mit aufgeführt, um Verwirrung bei den negativen Items zu vermeiden (s. Kap. 8.1). Der Frageblock enthält insgesamt drei Itembatterien, die mit allgemeinen Fragen zum Streben nach persönlicher Anerkennung beginnen (Tab. 7):

Tab. 7: Operationalisierung der persönlichen Anerkennung (Quelle: Eigene Darstellung, basierend auf Pickton & McKnight, 2006, S. 208)

Konstrukt	Korrespondierende Items		Code
Persönliche Anerkennung (prestige)	prestige1	Ich möchte zeigen, dass ich erfolgreich bin.	
		• stimme überhaupt nicht zu	1
		• stimme weniger zu	2
		• stimme weitgehend zu	3
		• stimme voll und ganz zu	4
	prestige2	Ich mag es nicht, wenn mein Name im Internet auftaucht.	
		• stimme überhaupt nicht zu	4
		• stimme weniger zu	3
		• stimme weitgehend zu	2
		• stimme voll und ganz zu	1

Die zwei Items basieren auf den Ergebnissen von Pickton und McKnight, die den Studierenden im Rahmen einer offenen Frage die Möglichkeit boten, Gründe für das Publizieren ihrer Arbeit zu nennen, wobei u. a. die Gründe „to prove oneself“ sowie „for personal satisfaction (including seeing one’s name in print)“ genannt wurden (2006, S. 208; s. Kap. 5.3.1). Während das Item *prestige1* eine sinngemäße deutsche Übersetzung des ersten Grundes darstellt, wurde der zweite Grund für das Item *prestige2* in den digitalen Raum übertragen und zudem negativ formuliert, wofür die Codierung entsprechend umgepolt wurde. Je höher der Mittelwert ausfällt, desto mehr suchen die Teilnehmenden nach persönlicher Anerkennung.

Mit der nächsten Itembatterie zum Karrierestreben wird der Bezug zum Studium hergestellt (Tab. 8):

Tab. 8: Operationalisierung des Karrierestrebens (Quelle: Eigene Darstellung, basierend auf Bankier & Smith, 2010, S. 253; Borgman, 2007, S. 238; Caprio, 2014, S. 149; Exline, 2016, S. 26; Rozum et al., 2015, S. 810)

Konstrukt	Korrespondierende Items		Code
Karrierestreiben (<i>career</i>)	career1	Gute Leistungen im Studium erhöhen meine Chancen auf dem Arbeitsmarkt.	
		• stimme überhaupt nicht zu	1
		• stimme weniger zu	2
		• stimme weitgehend zu	3
		• stimme voll und ganz zu	4
	career2	Meine Noten haben keinen Einfluss auf meine späteren Berufsaussichten.	
		• stimme überhaupt nicht zu	4
		• stimme weniger zu	3
		• stimme weitgehend zu	2
• stimme voll und ganz zu		1	

Beide Items zum Karrierestreiben wurden auf Basis der Literatur zu studentischen Inhalten in IR neu formuliert, in der erhöhte Wettbewerbschancen auf dem Arbeitsmarkt oder bei Graduiertenschulen für Studierende mit Publikationserfahrung vermutet werden (s. Kap. 2.2.2). Erneut wird das zweite Item negativ formuliert und entsprechend umcodiert. Je höher der Mittelwert ausfällt, desto mehr Wert legen die Teilnehmenden darauf, bereits im Studium die Grundlage für ihren künftigen beruflichen Erfolg zu legen.

Die letzte Itembatterie zum schulischen Fleiß fokussiert sich auf die Studienleistungen der Teilnehmenden (Tab. 9):

Tab. 9: Operationalisierung des schulischen Fleißes (Quelle: Eigene Darstellung, basierend auf Pickton & McKnight, 2006, S. 213)

Konstrukt	Korrespondierende Items		Code
Schulischer Fleiß (<i>diligence</i>)	diligence1	Ich möchte die Qualität meiner Studienleistungen ständig verbessern.	
		• stimme überhaupt nicht zu	1
		• stimme weniger zu	2
		• stimme weitgehend zu	3
		• stimme voll und ganz zu	4
	diligence2	Ich lege keinen Wert auf konstruktives Feedback.	
		• stimme überhaupt nicht zu	4
		• stimme weniger zu	3
		• stimme weitgehend zu	2
		• stimme voll und ganz zu	1

Das Item *diligence1* geht auf die Ergebnisse von Pickton und McKnight zurück, die den Wunsch nach Feedback und Anmerkungen als einen der wichtigsten Publikationsmotivatoren der Promotionsstudierenden identifizierten (2006, S. 213; s. Kap. 3). Das zweite Item stellt eine negative Umformulierung des ersten Items dar und wird entsprechend umgepolt. Je höher der Mittelwert ausfällt, desto mehr Wert legen die Studierenden auf ihre schulische Leistung.

7.2 Frageblock 2: Hochschule

Der zweite Frageblock untersucht die Beziehung der Teilnehmenden zu ihrer Hochschule. Die erste Itematterie fragt nach allgemeinen Identifikationsgefühlen (Tab. 10):

Tab. 10: Operationalisierung der Identifikation (Quelle: Eigene Darstellung, basierend auf Kim, 2011, Anhang A)

Konstrukt	Korrespondierende Items		Code
Identifikation (<i>identification</i>)	identification1	Ich fühle mich meiner Hochschule zugehörig.	
		• stimme überhaupt nicht zu	1
		• stimme weniger zu	2
		• stimme weitgehend zu	3
		• stimme voll und ganz zu	4
	identification2	Ich bin stolz darauf, an meiner Hochschule zu studieren.	
		• stimme überhaupt nicht zu	1
		• stimme weniger zu	2
		• stimme weitgehend zu	3
		• stimme voll und ganz zu	4

Beide Items wurden nahezu unverändert von Kim (2011, Anhang A) übernommen; lediglich beim Item *identification2* wurde die Rolle von Professor:in zu Student:in geändert. Je höher der Mittelwert ist, desto mehr identifizieren sich die Studierenden mit ihrer Hochschule. Die zweite Itembatterie untersucht die Variable der Vertrauenswürdigkeit (Tab. 11):

Tab. 11: Operationalisierung der Vertrauenswürdigkeit (Quelle: Eigene Darstellung, basierend auf Kim, 2011, Anhang A)

Konstrukt	Korrespondierende Items		Code
Vertrauenswürdigkeit (<i>trustworthiness</i>)	trustworthiness1	Inhalte, die meine Hochschule veröffentlicht, sind von guter Qualität.	
		• stimme überhaupt nicht zu	1
		• stimme weniger zu	2
		• stimme weitgehend zu	3
		• stimme voll und ganz zu	4
	trustworthiness2	Inhalte, die von meinen Lehrenden begutachtet wurden, sind von guter Qualität.	
		• stimme überhaupt nicht zu	1
		• stimme weniger zu	2
		• stimme weitgehend zu	3
		• stimme voll und ganz zu	4

Erneut basieren beide Items auf der Operationalisierung von Kim (2011, Anhang A), wobei ihre Formulierungen vereinfacht und auf die studentische Perspektive angepasst wurden. Während das Item *trustworthiness1* allgemein nach dem Vertrauen der Teilnehmenden in ihre Hochschule als herausgebende Institution fragt, stellt das Item *trustworthiness2* das Vertrauen in den Begutachtungsprozess in den Vordergrund. Gemäß ihrer Zielgruppe nimmt Kim dabei den Peer-Review in den Fokus, der bei studentischen Inhalten, wie in Kapitel 5.3.1 dargestellt, häufig durch qualitätssichernde Maßnahmen, wie z. B. die Begutachtung und Empfehlung durch eine Lehrperson, ersetzt wird. Analog hierzu wird auch das Item umformuliert, indem nach dem Vertrauen der Teilnehmenden in durch Lehrpersonen begutachtete Inhalte gefragt wird. Je höher der Mittelwert ist, desto mehr Vertrauen haben die Studierenden in die qualitätssichernden Prozesse ihrer Hochschule.

7.3 Frageblock 3: Upload-Prozess

Der dritte Frageblock befasst sich mit dem potenziellen Kostenfaktor der Zeit- und Aufwandsinvestition. Um eine gemeinsame Ausgangsbasis zu schaffen, wird den Teilnehmenden ein konkretes Szenario vor Augen geführt:

Stelle dir vor, dass deine **Hochschule** eine eigene **Website** eingerichtet hat, auf der du deine im Studium entstandenen Arbeiten veröffentlichen kannst. Voraussetzung ist, dass sie von einer **Lehrperson** begutachtet und zur Veröffentlichung **empfohlen** wurden. Deine Arbeiten können **frei** und **kostenlos** eingesehen werden. Beim Upload musst du ein **Formular** ausfüllen, in dem du deine Arbeit beschreibst (Anhang A).

Wieder werden Schlüsselbegriffe fett gedruckt, damit sie auch beim Überfliegen des Textes erfasst werden können. Auch auf Fachjargon, wie z. B. ‚institutionelles Repositorium‘ oder ‚Open Access‘, wird erneut verzichtet (s. Kap. 7.1). Weil in der Literatur zu studentischen IR-Inhalten die Bedeutung qualitätssichernder Maßnahmen hervorgehoben wird (s. Kap. 2.2.2), wird in dem Szenario die Begutachtung durch eine Lehrperson vorausgesetzt. Das Upload-Formular wird durch eine Mockup-Grafik visualisiert (Abb. 5; s. Kap. 8.1):

Titel:	<input type="text"/>
Sprache:	<input type="text"/>
Autorenschaft:	<input type="text"/>
Schlagwörter:	<input type="text"/>
Zusammenfassung:	<input type="text"/>
Einrichtung:	<input type="text"/>
Dokumenttyp:	<input type="text"/>
Hauptgutachter*in:	<input type="text"/>

Abb. 5: Mockup des Upload-Formulars (Quelle: Eigene Darstellung, basierend auf HAW Hamburg, o. J.b)

Die in der Mockup-Grafik aufgeführten Metadatenfelder basieren auf den Metadatenfeldern von REPOSIT, dem IR der HAW Hamburg (o. J.b). Nahezu alle Felder wurden unverändert von REPOSIT übernommen; lediglich das Feld der Einrichtung wurde für eine verbesserte Verständlichkeit in ‚Department‘ umbenannt, weil den Teilnehmenden die Möglichkeit fehlt, die auf REPOSIT beim Einrichtungsfeld hinterlegte und über seine Funktion aufschlussgebende *Gemeinsame Normdatei* (GND) zurate zu ziehen. Das Feld ‚Hauptgutachter*in‘ entspricht in seiner Funktion einem „peer reviewed‘ indicator“ bzw. einer Kennzeichnung qualitätsgeprüfter Inhalte, wie sie Pickton und McKnight in ihren Handlungsempfehlungen nahelegen (2006, S. 216).

Auf die Mockup-Grafik folgt eine Wiederholung der Ausfüllinstruktion aus dem ersten Frageblock (s. Kap. 7.1), wodurch den Teilnehmenden die vier Stufen der Likert-Skala in Erinnerung gerufen werden. Anschließend folgt die Itematterie der Zeit- und Aufwandsinvestition (Tab. 12):

Tab. 12: Operationalisierung der Zeit- und Aufwandsinvestition (Quelle: Eigene Darstellung, basierend auf Kim, 2011, Anhang A)

Konstrukt	Korrespondierende Items		Code
Zeit- und Aufwandsinvestition (<i>time_effort</i>)	time_effort1	Es ist leicht, das von meiner Hochschule geforderte Upload-Formular auszufüllen.	
		• stimme überhaupt nicht zu	4
		• stimme weniger zu	3
		• stimme weitgehend zu	2
		• stimme voll und ganz zu	1
	time_effort2	Meine Arbeitsergebnisse auf der Website meiner Hochschule zu veröffentlichen, kostet mich Zeit, die ich besser in mein Studium investieren könnte.	
		• stimme überhaupt nicht zu	1
		• stimme weniger zu	2
		• stimme weitgehend zu	3
		• stimme voll und ganz zu	4
	time_effort3	Es ist aufwendig, meine Arbeitsergebnisse auf der Website meiner Hochschule zu veröffentlichen.	
		• stimme überhaupt nicht zu	1
		• stimme weniger zu	2
• stimme weitgehend zu		3	
• stimme voll und ganz zu		4	

Alle drei Items spiegeln die Operationalisierung von Kim (2011, Anhang A) wider. Je höher der Mittelwert ist, desto höher schätzen die Studierenden den Zeit- und Arbeitsaufwand des Selbstarchivierens ein. Das erste Item nimmt direkten Bezug auf die Mockup-Grafik und fragt nach dem wahrgenommenen Aufwand der Dateneingabe, wobei eine positive Umformulierung und entsprechende Umcodierung vorgenommen wird. Für das zweite Item wurde Kims Formulierung dahingehend überarbeitet, dass die Arbeit des wissenschaftlichen Forschens durch das Studium ersetzt wird. Das letzte Item fragt nach dem wahrgenommenen Aufwand einer Selbstarchivierung und leitet so zum nächsten Frageblock über.

7.4 Frageblock 4: Selbstarchivierungsbereitschaft

Im vierten Frageblock steht die konkrete Selbstarchivierungsbereitschaft der Studierenden im Fokus. Der AV werden zwei Konstrukte vorangestellt, die in der bestehenden Forschung als besonders einflussreich identifiziert wurden (s. Kap. 3;5.2.3), weshalb die Teilnehmenden bei der Beantwortung des AV-Items unter dem Eindruck dieser Konstrukte stehen sollten. Dies ist zum einen die Variable der Plagiatsfurcht (Tab. 13):

Tab. 13: Operationalisierung der Plagiatsfurcht (Quelle: Eigene Darstellung, basierend auf Kim, 2011, Anhang A; Pickton & McKnight, 2006, S. 212)

Konstrukt	Korrespondierende Items		Code
Plagiatsfurcht (<i>plagiarism</i>)	plagiarism1	Ich habe Angst, dass andere von mir abschreiben, wenn ich meine Arbeitsergebnisse auf der Website meiner Hochschule veröffentliche.	
		• stimme überhaupt nicht zu	1
		• stimme weniger zu	2
		• stimme weitgehend zu	3
		• stimme voll und ganz zu	4
	plagiarism2	Andere könnten meine Arbeitsergebnisse als ihre eigenen ausgeben, wenn ich sie auf der Website meiner Hochschule veröffentliche.	
		• stimme überhaupt nicht zu	1
		• stimme weniger zu	2
		• stimme weitgehend zu	3
		• stimme voll und ganz zu	4

Beide Items wurden auf Basis von Kims Operationalisierung ihrer Vertrauensvariable formuliert, welche ein Item zur Plagiatsfurcht enthält (2011, Anhang A). Zusätzlich wird die Formulierung der Items von den von der Selbstarchivierung abschreckenden Faktoren beeinflusst, die Pickton und McKnight identifizieren (2006, S. 212; s. Kap. 3). Je höher der Mittelwert ist, desto mehr Furcht vor Plagiaten verspüren die Studierenden.

Auf den potenziellen Kostenfaktor der Plagiatsfurcht folgt der potenzielle Gewinnfaktor des Altruismus (Tab. 14):

Tab. 14: Operationalisierung des Altruismus (Quelle: Eigene Darstellung, basierend auf Kim, 2011, Anhang A)

Konstrukt	Korrespondierende Items		Code
Altruismus (<i>altruism</i>)	altruism1	Ich teile mein Wissen mit meinen Kommiliton:innen, wenn sie Hilfe brauchen.	
		• stimme überhaupt nicht zu	1
		• stimme weniger zu	2
		• stimme weitgehend zu	3
		• stimme voll und ganz zu	4
	altruism2	Wenn mir geholfen wurde, möchte ich etwas zurückgeben und anderen helfen.	
		• stimme überhaupt nicht zu	1
		• stimme weniger zu	2
		• stimme weitgehend zu	3
		• stimme voll und ganz zu	4
	altruism3	Die Ergebnisse wissenschaftlicher Forschung sollten für alle frei und kostenlos zugänglich sein.	
		• stimme überhaupt nicht zu	1
		• stimme weniger zu	2
• stimme weitgehend zu		3	
• stimme voll und ganz zu		4	

Wieder gehen die Items auf Kims Operationalisierung zurück (2011, Anhang A), wobei erneut Umformulierungen aufgrund der studentischen Zielgruppe vorgenommen werden. Die Items *altruism1* und *altruism2* reflektieren die Facette der Hilfsbereitschaft, die in Kapitel 5.3.1 als eine der zwei Facetten des Altruismus-Konstrukts identifiziert wurde. Während *altruism1* für einen reinen Altruismus steht, drückt *altruism2* die schwer von diesem zu trennende ‚goldene Regel‘ aus. Das Item *altruism3* fragt ab, inwieweit die Teilnehmenden eine Philosophie der Offenheit im Wissenschaftskontext unterstützen, was die zweite Facette des Altruismus-Konstrukts darstellt. Im Gegensatz zu Kim wird hierbei auf die Verwendung des OA-Fachbegriffes verzichtet. Je höher der Mittelwert ausfällt, desto altruistischer sind die Studierenden veranlagt.

Nachdem die Teilnehmenden über ihren eigenen Standpunkt zu OA reflektiert haben, folgt die AV der Selbstarchivierungsbereitschaft (Tab. 15):

Tab. 15: Operationalisierung der Selbstarchivierungsbereitschaft (Quelle: Eigene Darstellung)

Konstrukt	Korrespondierende Items		Code
Selbstarchivierungsbereitschaft (<i>selfarchive</i>)	selfarchive	Ich kann mir vorstellen, meine Arbeitsergebnisse auf der Website meiner Hochschule zu veröffentlichen.	
		• stimme überhaupt nicht zu	1
		• stimme weniger zu	2
		• stimme weitgehend zu	3
		• stimme voll und ganz zu	4

Wie in Kapitel 5.3.5 begründet, liefert Kims Operationalisierung an dieser Stelle keine sinnvolle Grundlage für das Messen der studentischen Selbstarchivierungsbereitschaft, weshalb das Item *selfarchive* neu formuliert wurde. Je höher der Wert ist, desto mehr sind die Studierenden zur Selbstarchivierung bereit.

Auf die AV folgt die Moderatorvariable der wissenschaftlichen Karriere (Tab. 16):

Tab. 16: Operationalisierung der wissenschaftlichen Karriere (Quelle: Eigene Darstellung)

Konstrukt	Korrespondierende Items		Code
Wissenschaftliche Karriere (<i>acad_career</i>)	acad_career	Ich kann mir vorstellen, eine wissenschaftliche Karriere einzuschlagen und selbst an einer Hochschule zu lehren und zu forschen.	
		• stimme überhaupt nicht zu	1
		• stimme weniger zu	2
		• stimme weitgehend zu	3
		• stimme voll und ganz zu	4

Das Item *acad_career* wurde ebenfalls neu formuliert.

7.5 Frageblock 5: Demografische Daten

Im abschließenden fünften Frageblock werden die demografischen Daten der Teilnehmenden zum Zweck der Stichprobenkontrolle (s. Kap. 9.2.1) erhoben (Tab. 17):

Tab. 17: Operationalisierung der demografischen Daten (Quelle: Eigene Darstellung)

Konstrukt	Korrespondierende Items		Code
Alter (age)	age	Wie alt bist du?	
		<ul style="list-style-type: none"> • <i>Kurzantwort-Textfeld</i> 	
Geschlecht (gender)	gender	Welchem Geschlecht fühlst du dich zugehörig?	
		<ul style="list-style-type: none"> • weiblich 	0
		<ul style="list-style-type: none"> • männlich 	1
		<ul style="list-style-type: none"> • divers 	2
Studienfach (discipline)	discipline	Welches Fach studierst du?	
		<ul style="list-style-type: none"> • <i>Kurzantwort-Textfeld</i> 	0-n
Abschluss (degree)	degree	Welchen Abschluss strebst du gerade an?	
		<ul style="list-style-type: none"> • Bachelor 	0
		<ul style="list-style-type: none"> • Master 	1
		<ul style="list-style-type: none"> • Sonstiges (<i>Kurzantwort-Textfeld</i>) 	2-n

Neben Alter und Geschlecht werden auch das Studienfach sowie der angestrebte Abschluss abgefragt. Um keine Informationen zu verlieren, wird – im Gegensatz zu Kim (s. Kap. 5.2.2) – davon abgesehen, die Studienfächer im Vorweg in übergeordnete Kategorien, wie z. B. Geistes- oder Naturwissenschaften, einzuordnen. Am Ende des Fragebogens erhalten die Studierenden die Möglichkeit, ihre E-Mail-Adresse für die Teilnahme am incentivierenden Gewinnspiel zu hinterlegen. Nach einer Danksagung werden abschließend die Kontaktadresse sowie der Hinweis wiederholt, dass diese anonyme Umfrage rein wissenschaftlichen Zwecken dient.

Auf die Operationalisierung der Konstrukte folgt im nächsten Kapitel die Beschreibung der empirischen Studie.

8 Empirische Studie

Die in Kapitel 6 aufgestellten Hypothesen wurden im Rahmen einer empirischen Studie überprüft, die im Zeitraum vom 09.12.2024 bis zum 23.01.2025 durchgeführt wurde. Insgesamt wurden 97 Studierende im Alter zwischen 17 und 42 Jah-

ren befragt. In diesem Kapitel wird zunächst die technische Konstruktion des in Kapitel 7 erarbeiteten Fragebogens dargestellt, wobei auch die im Rahmen des qualitativen Pretests vorgenommenen Änderungen begründet werden. Anschließend wird der Prozess der Teilnehmendenakquise erläutert.

8.1 Konstruktion des Fragebogens und qualitativer Pretest

Die Datenerhebung fand in Form einer Online-Befragung statt, für die das Umfrage-Tool *Google Formulare* gewählt wurde. Aufgrund des responsiven Designs konnte die Umfrage sowohl an stationären als auch an mobilen Endgeräten durchgeführt werden. Um den wissenschaftlichen Zweck der Umfrage hervorzuheben, wurde das Layout an das Corporate Design der HAW Hamburg angepasst, indem das Hochschullogo als Header und der hierin enthaltene Blauton als Akzentfarbe verwendet wurde. Ein Fortschrittsbalken unterstützte die Teilnehmenden dabei, den verbleibenden zeitlichen Aufwand abzuschätzen (Döring, 2023, S. 410). Der vollständige Fragebogen ist Anhang A zu entnehmen.

Eine erste Version des Online-Fragebogens wurde im Rahmen eines qualitativen Pretests mit einer Testperson durchgegangen. Der Pretest bestand aus zwei Stufen, wobei die erste Stufe die Verständlichkeit der Items thematisierte und die zweite Stufe nach Hinweisen zur Optimierung der anstehenden Teilnehmendenakquise suchte. In Stufe eins fand die *Methode des lauten Denkens* Anwendung, bei der die Testperson ihre Gedanken verbalisierte, die bei der Rezeption des Fragebogens auftraten (Döring, 2023, S. 366). Ursprünglich wurden die Endpunkte der Skalen mit „überhaupt nicht“ bzw. „voll und ganz“ benannt. Dies führte zu Missverständnissen bei negativ umformulierten Items, wie z. B. „Ich mag es nicht, wenn mein Name im Internet auftaucht“ (*prestige2*), wo „überhaupt nicht“ sowohl im Sinne von „Ich mag dies überhaupt nicht“ als auch im Sinne von „Ich stimme überhaupt nicht zu“ interpretiert werden konnte. Um diesen Missverständnissen vorzubeugen, wurden die Skalenendpunkte, wie bereits in Kapitel 7.1 erwähnt, mit „stimme überhaupt nicht zu“ bzw. „stimme voll und ganz zu“ benannt.

Bei der Itematterie zur Zeit- und Aufwandsinvestition wurde das Einschätzen der Ausfüllschwierigkeit eines hypothetischen Formulars von der Testperson als zu abstrakt empfunden, woraufhin eine ergänzende Mockup-Grafik erstellt wurde. In der ersten Fragebogenversion folgten zudem die Items zur Vertrauenswürdigkeit unmittelbar auf die Items der Zeit- und Aufwandsinvestition, was bei der Testper-

son die Unsicherheit auslöste, ob diese Fragen ebenfalls in Bezug auf das hypothetische Formular zu beantworten seien. Um diese Unsicherheit auszuräumen, wurden alle den Upload-Prozess betreffenden Items in einem eigenen Fragebogenabschnitt gruppiert, der mit der Beschreibung des Upload-Szenarios inklusive Mockup-Grafik eingeleitet wird.

In Stufe zwei wurde im Rahmen einer formlosen Befragung untersucht, welche Formulierungen im Ansprachetext die Rücklaufquote erhöhen könnten. Die zeitliche Umfangsangabe von sieben Minuten wurde von der Testperson als wenig aussagekräftig eingeschätzt, da jede Person in ihrem eigenen Tempo arbeitet und dementsprechend mehr oder weniger Zeit für das Ausfüllen des Fragebogens benötigt. Aus diesem Grund wurde die zeitliche Angabe durch eine quantitative Angabe der zu beantwortenden Fragen ergänzt. Die Testperson gab an, bei mehr als 20 Fragen die Motivation zu verlieren, weshalb ein Item entfernt wurde, um insgesamt auf eine runde Anzahl von 20 Items zu kommen. Die Items der demografischen Angaben wurden dabei nicht mitgezählt.

Ein weiterer Hinweis der Testperson betraf die Fragetechnik, die in einer ersten Formulierung des Ansprachetextes nicht näher beschrieben wurde. Schon der Verdacht darauf, dass der Fragebogen offene Fragen enthalten könnte, senkte die Teilnahmemotivation der Testperson, weshalb dem Ansprachetext der Hinweis hinzugefügt wurde, dass es sich um 20 geschlossene Fragen handelt. Weil die Begrifflichkeit der geschlossenen Frage ihrerseits Verwirrung auslösen und die Teilnahmemotivation senken könnte, wurde dabei der umgangssprachliche Begriff der Ankreuzfrage verwendet.

8.2 Teilnehmendenakquise

Die Teilnehmenden wurden sowohl digital als auch analog rekrutiert. Für die digitale Rekrutierung wurde der Link zum Fragebogen zunächst am 09. Dezember 2024 über den Messengerdienst *WhatsApp* sowie per E-Mail an insgesamt 27 Studierende im persönlichen Bekanntenkreis verschickt. Hierbei handelte es sich ausschließlich um Studierende der HAW Hamburg sowie der Universität Hamburg, doch weil der Fragebogen von diesen wiederum an ihre Bekannten weitergeleitet wurde, können auch Studierende anderer Hochschulen in der Stichprobe vertreten sein. Der Ansprachetext wurde abhängig vom Vertrautheitsgrad in der Tonalität angepasst. Konstante Bestandteile des Textes waren die Benennung des zeitlichen

und quantitativen Bearbeitungsaufwands (sieben Minuten bzw. 20 Ankreuzfragen) sowie der Hinweis auf das incentivierende Gewinnspiel, bei dem drei Amazon-Gutscheine im Wert von je 15 Euro verlost wurden. Am 14. Januar 2025 wurde die Umfrage in einer zweiten digitalen Befragungsrunde per E-Mail an insgesamt 733 Studierende des Departments für Information und Medienkommunikation der HAW Hamburg verschickt, wobei folgender Ansprachetext verwendet wurde:

Hallo liebe Kommiliton*innen,

sicher kennt ihr es: Man steckt viel Arbeit in eine Studienleistung, die dann nur eine Lehrperson je zu Gesicht bekommt.

Was haltet ihr von der Idee, eure **studentischen Arbeiten** online zu **veröffentlichen**? Mit dieser Frage beschäftige ich mich in meiner Masterarbeit, für die ich eine kleine Umfrage mit **20 Ankreuzfragen** (Dauer ca. **7 Min.**) erstellt habe: [Link]

Unter allen Teilnehmenden verlose ich **3x15€** als **Amazon-Gutscheine**.

Vielen Dank für eure Unterstützung!

Eure Silja

silja.wiegmann@haw-hamburg.de

MA Digitale Transformation

Der Text der Rundmail stellt eine kondensierte Fassung des Einleitungstextes des Fragebogens dar, dessen Aufbau und Tonalität in Kapitel 7.1 begründet wurde.

Für die analoge Rekrutierung wurden am 10. Dezember 2024 Flyer in der *Fachbibliothek Design, Medien und Information* (DMI) der HAW Hamburg sowie in der *Fachbereichsbibliothek Kulturwissenschaften* der Universität Hamburg ausgehängt, was zuvor mit dem dortigen Bibliothekspersonal abgestimmt wurde. Abbildung 6 zeigt den in der Fachbibliothek DMI ausgehängten Flyer:

Deine Meinung ist gefragt!

Was hältst du von der Idee, deine studentischen Arbeiten online zu veröffentlichen? z. B.:

- | | |
|---|--|
|  Hausarbeiten |  Folien |
|  Projektberichte |  Code |



- 20 Ankreuzfragen
- Dauer: ca. 7 Min.
- Verlosung: 3x 15€ als Amazon-Gutscheine



Abb. 6: Rekrutierungsflyer für die Fachbibliothek DMI (Quelle: Eigene Darstellung)

Auf dem Flyer werden die wichtigsten Eckpunkte des Ansprachetextes grafisch aufbereitet. Die Befragung wird als ein an der HAW Hamburg stattfindendes Masterprojekt ausgewiesen, wobei auch Kontaktdaten genannt werden. Ein QR-Code leitet die Studierenden zum Fragebogen weiter. Die aufgelisteten möglichen Dokumentarten werden anhand von Piktogrammen veranschaulicht, wobei Anpassungen an die in den jeweiligen Bibliotheken vertretenen Fachkulturen vorgenommen wurden. Abbildung 7 zeigt den in der Fachbereichsbibliothek Kulturwissenschaften ausgehängten Flyer:

Deine Meinung ist gefragt!

Was hältst du von der Idee, deine studentischen Arbeiten online zu veröffentlichen? z. B.:

- | | |
|---|--|
|  Hausarbeiten |  Folien |
|  Projektberichte |  Essays |



- 20 Ankreuzfragen
- Dauer: ca. 7 Min.
- Verlosung: 3x15€ als Amazon-Gutscheine



Abb. 7: Rekrutierungsflyer für die Fachbereichsbibliothek Kulturwissenschaften (Quelle: Eigene Darstellung)

Während der Flyer der Fachbibliothek DMI Code als mögliche Dokumentart auflistet, wird dieser auf dem Flyer der Fachbereichsbibliothek Kulturwissenschaften durch Essays ersetzt.

Die Umfrage wurde am 23. Januar 2025 geschlossen. Am selben Tag wurden die Amazon-Gutscheine mithilfe eines Zufallsgenerators verlost und an die Gewinner:innen verschickt. Die Bruttoreклауquote beläuft sich auf 12,8%, wobei unbekannt bleibt, wie viele Bibliotheksnutzende die ausgehängten Flyer zur Kenntnis genommen haben und an wie viele zusätzliche Personen der Fragebogen durch Studierende im persönlichen Bekanntenkreis weitergeleitet wurde.

9 Datenanalyse

Die Analyse der erhobenen Daten wurde mithilfe der freien Programmiersprache R durchgeführt. Sowohl die statistischen Berechnungen als auch die Diagramme zur Datenvisualisierung wurden dabei in Form eines R-Notebooks festgehalten (s. Anhang B). Vor dem Datenimport wurden die Rohdaten einer ersten Plausibilitätskontrolle unterzogen, in deren Rahmen der angestrebte Abschluss einer einzelnen Person von „Master“ zu „Bachelor“ geändert wurde.² Die Studienfächer der Teilnehmenden wurden manuell vereinheitlicht und von Tippfehlern bereinigt. So wurde z. B. die Abkürzung „BIM“ konsequent ausgeschrieben als „Bibliotheks- und Informationsmanagement“. Wenn zwei Fächer angegeben wurden, ließ sich davon ausgehen, dass es sich bei dem zweitgenannten Fach um das Nebenfach handelte, welches dementsprechend entfernt wurde. Die Angabe „geheim“ wurde zu „NA“ geändert.

Analog zu Kims methodischem Vorgehen (s. Kap. 5.2.2) wird in Kapitel 9.1 zunächst eine Reliabilitätsanalyse vorgenommen, ehe die Items durch die Bildung der Mittelwerte zu gemeinsamen Skalen zusammengefasst werden. Anschließend folgt in Kapitel 9.2 die Beschreibung der Stichprobe, wobei sowohl auf demografische Merkmale als auch auf ausgewählte statistische Maßzahlen der Verteilung eingegangen wird. In Kapitel 9.3 werden zunächst die Einzelhypothesen (s. Kap. 6) anhand einfacher linearer Regressionen überprüft, woraufhin dann die Überprüfung des Gesamtmodells anhand einer multiplen linearen Regression folgt.

9.1 Reliabilitätsanalyse: Spearman-Brown-Koeffizient und Cronbachs Alpha

Wie bereits in Kapitel 5.2.2 erläutert, ist es vor der Bildung gemeinsamer Skalen erforderlich, die interne Konsistenz der Konstrukte zu überprüfen. Das von Kim herangezogene Reliabilitätsmaß *Cronbachs Alpha* wird auch bei der vorliegenden Analyse für alle Skalen mit mehr als zwei Items verwendet. Weil bei Skalen mit genau zwei Items die Gefahr besteht, dass Cronbachs Alpha die Reliabilität unter-

² Vor dem Hintergrund des geringen Alters der Person (17 Jahre) sowie dem angegebenen Studienfach (Bibliotheks- und Informationsmanagement) wurde klar ersichtlich, dass sie sich aktuell nicht in einem postgradualen Studiengang befinden kann, gleichwohl sie dies für die Zukunft anstreben mag.

schätzen könnte (Eisinga, 2013, S. 641), findet bei diesen der *Spearman-Brown-Koeffizient* Anwendung. Tabelle 18 zeigt die Ergebnisse der Reliabilitätsanalyse mittels Cronbachs Alpha:

Tab. 18: Ergebnisse der Reliabilitätsanalyse mittels Cronbachs Alpha (Quelle: Eigene Darstellung)

Konstrukt	Cronbachs Alpha	Reliabilität durch Item-Ausschluss	
time_effort	0,577	time_effort1	0,654
		time_effort2	0,377
		time_effort3	0,288
altruism	0,571	altruism1	0,330
		altruism2	0,282
		altruism3	0,738

Beim Konstrukt *time_effort* liegt Cronbachs Alpha mit einem Wert von 0,577 unter dem akzeptablen Mindestwert von 0,70 (s. Kap. 5.2.2), sodass hier keine hinreichende interne Konsistenz besteht. Durch den Ausschluss des Items *time_effort1* lässt sich ein verbesserter Wert von 0,654 erzielen, weshalb dieses Item vor der Mittelwertbildung entfernt wird. Auch beim Konstrukt *altruism* liegt Cronbachs Alpha mit einem Wert von 0,571 unter dem akzeptablen Mindestwert. Durch den Ausschluss des Items *altruism3*, das den ‚OA-Spirit‘ (Kim, 2010, S. 1916; s. Kap. 5.3.1) bzw. den Grad der Zustimmung zu einer Philosophie der Offenheit im Wissenschaftskontext misst, lässt sich ein sehr guter Wert von 0,738 erreichen, weshalb dieses Item ebenfalls entfernt wird.

Tabelle 19 zeigt die Ergebnisse der Reliabilitätsanalyse mittels des Spearman-Brown-Koeffizienten:

Tab. 19: Ergebnisse der Reliabilitätsanalyse mittels des Spearman-Brown-Koeffizienten (Quelle: Eigene Darstellung)

Konstrukt	Spearman-Brown-Koeffizient
prestige	0,598
career	0,508
diligence	0,102
identification	0,825
trustworthiness	0,710
plagiarism	0,795

Anders als bei Cronbachs Alpha gibt es für den Spearman-Brown-Koeffizienten keinen Grenzwert, der sich weitläufig in der Literatur durchgesetzt hat. Vereinzelt wird ein Wert $>0,5$ als akzeptabel und ein Wert $>0,7$ als gut erachtet (Öztürk & Kудay, 2024, S. 9; Vet et al., 2017, S. 48), was an dieser Stelle als Richtlinie übernommen wird. Bei den Konstrukten *identification*, *trustworthiness* und *plagiarism* liegt der Spearman-Brown-Koeffizient über dem Wert von $0,7$, sodass hier eine gute interne Konsistenz vorliegt. Die Werte der Konstrukte *prestige* und *career* liegen zwischen $0,5$ und $0,7$, weshalb auch hier eine hinreichende interne Konsistenz gegeben ist. Beim Konstrukt *diligence* ist der Spearman-Brown-Koeffizient mit einem Wert von $0,102$ auffällig schlecht, sodass diese Items nicht sinnvoll zu einer gemeinsamen Skala kombiniert werden können. Die Kontrollfrage *diligence2*, bei der womöglich auch soziale Erwünschtheit eine Rolle gespielt haben könnte (s. Kap. 10), wird entfernt.

9.2 Deskriptive Statistik

Von insgesamt 97 Teilnehmenden haben 94 Personen den Fragebogen vollständig ausgefüllt. Unvollständige Antworten wurden von der Datenanalyse ausgeschlossen, sodass $n=94$ beträgt. Allein bei den demografischen Angaben wurden Lücken akzeptiert und als „NA“ erfasst, weshalb die Größe der Stichprobe hier nicht einheitlich ist. Im Folgenden wird zunächst die Stichprobe anhand ihrer demografischen Merkmale beschrieben. Anschließend werden anhand von ausgewählten statistischen Maßzahlen der Verteilung die Einstellungen der Teilnehmenden dargestellt.

9.2.1 Stichprobenkontrolle

Soziodemografische Merkmale lassen Rückschlüsse über die Ausgewogenheit der Stichprobe zu. Abbildung 8 zeigt die relative Häufigkeit des Auftretens der Geschlechter in Form eines Kreisdiagramms:

Anteile der Geschlechter in % (n=94)

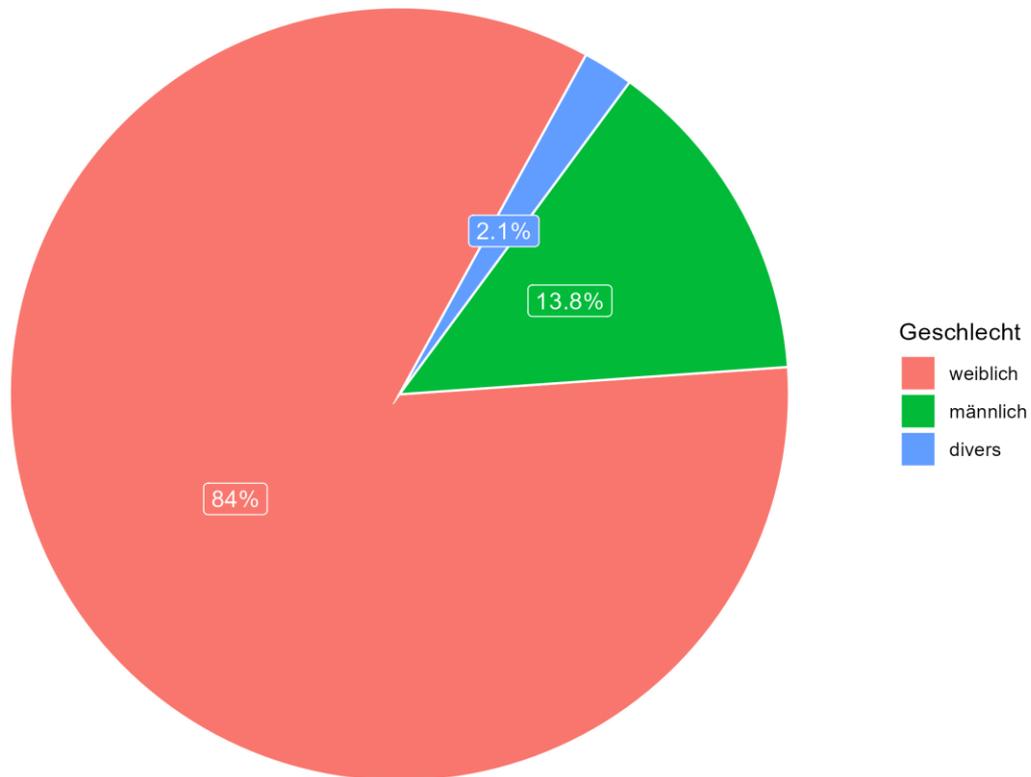


Abb. 8: Anteile der Geschlechter in % (n=94) (Quelle: Eigene Darstellung. Wegen Rundungsdifferenzen ergibt die Summe der Beschriftungen nicht 100%.)

84,0% der Teilnehmenden sind weiblich, sodass die Stichprobe in Hinsicht auf das Merkmal Geschlecht nicht repräsentativ ist.

Abbildung 9 zeigt die Altersverteilung in Form eines Histogramms:

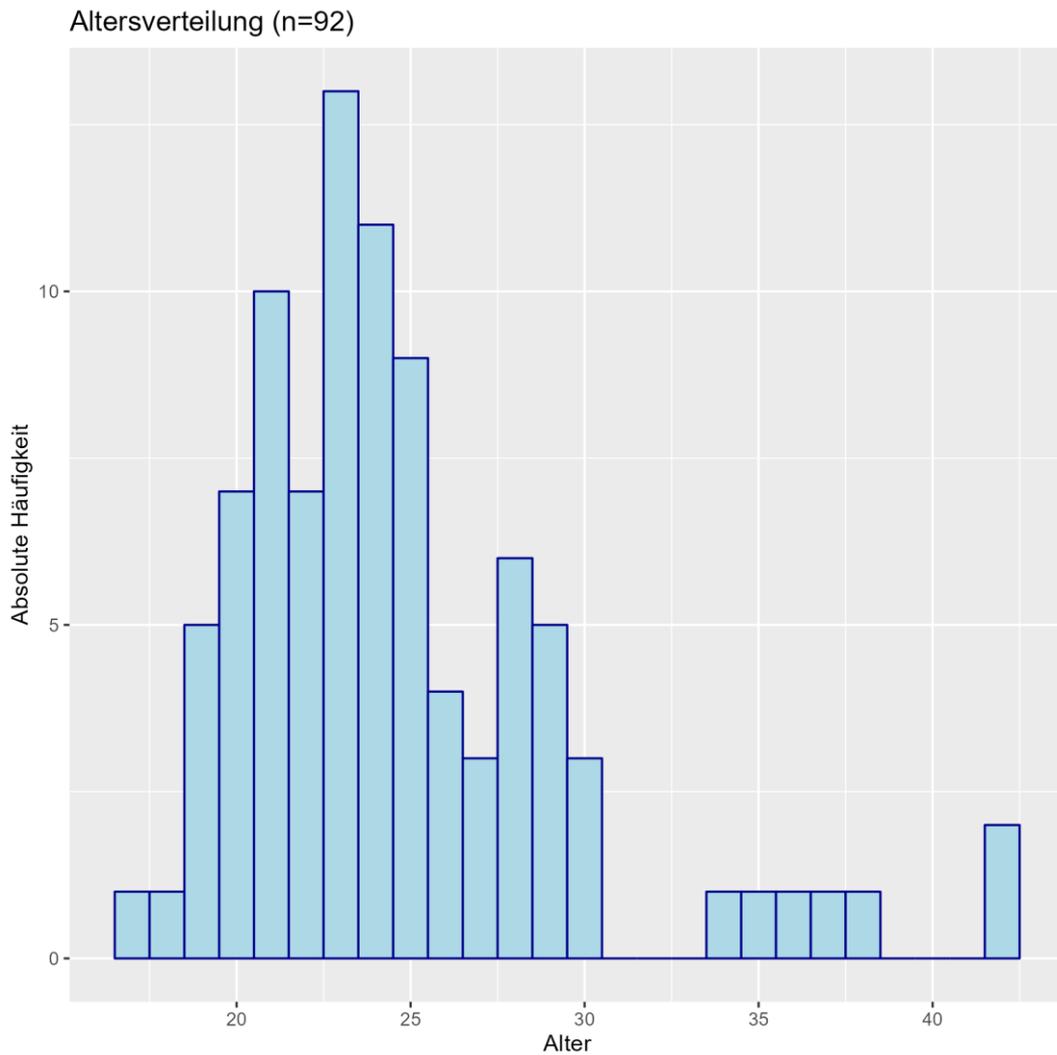


Abb. 9: Altersverteilung (n=92) (Quelle: Eigene Darstellung)

In Relation zur Gesamtbevölkerung ist die Stichprobe in Hinsicht auf das Merkmal Alter nicht repräsentativ, was jedoch auch nicht zu erwarten war, da gemäß des Forschungsinteresses bewusst nur Studierende befragt wurden. In der Gruppe der 17- bis 30-jährigen (92,4% der Teilnehmenden) ähnelt die Verteilung einer Glockenform, sodass die studentische Zielgruppe insgesamt gut abgedeckt wurde. Die sieben älteren Ausreißer (7,6% der Teilnehmenden) führen jedoch zu einer linkssteilen Verteilung.

Um eine diverse Bandbreite an Selbstarchivierungskulturen abzubilden, ist es erstrebenswert, dass möglichst viele verschiedene Fachbereiche in der Stichprobe enthalten sind (Pickton & McKnight, 2006, S. 206). Tabelle 20 zeigt die vertretenen Studienfächer, die in absteigender Reihenfolge nach der absoluten Häufigkeit der Studierenden sortiert sind:

Tab. 20: Vertretene Studienfächer (n=92) (Quelle: Eigene Darstellung)

Studienfach	Absolute Häufigkeit
Bibliotheks- und Informationsmanagement	36
Medien und Kommunikation	23
Digitale Transformation der Informations- und Medienwirtschaft	10
Klassische Archäologie	5
Digitale Kommunikation	4
Ethnologie	3
Germanistik	3
Geschichte	3
Kunstgeschichte	3
Empirische Kulturwissenschaft	1
Physik	1

Die Mehrheit der Teilnehmenden (79,3%) studiert ein informations- bzw. kommunikationswissenschaftliches Fach. Die zweitgrößte Gruppe (13,0%) ist den Kulturwissenschaften zuzuordnen. Sechs Teilnehmende (6,5%) gehören den Geisteswissenschaften an und lediglich eine Person studiert ein naturwissenschaftliches Fach. Vor dem Hintergrund der Teilnehmendenrekrutierung, die vorrangig unter Studierenden der Informations- und Kulturwissenschaften stattfand (s. Kap. 8.2), war diese Verteilung zu erwarten. Insgesamt ist die Stichprobe in Hinsicht auf das Merkmal der Studiendisziplin nicht repräsentativ.

Abbildung 10 zeigt die relative Häufigkeit der angestrebten Studienabschlüsse in Form eines Kreisdiagramms:

Anteile der angestrebten Abschlüsse in % (n=94)

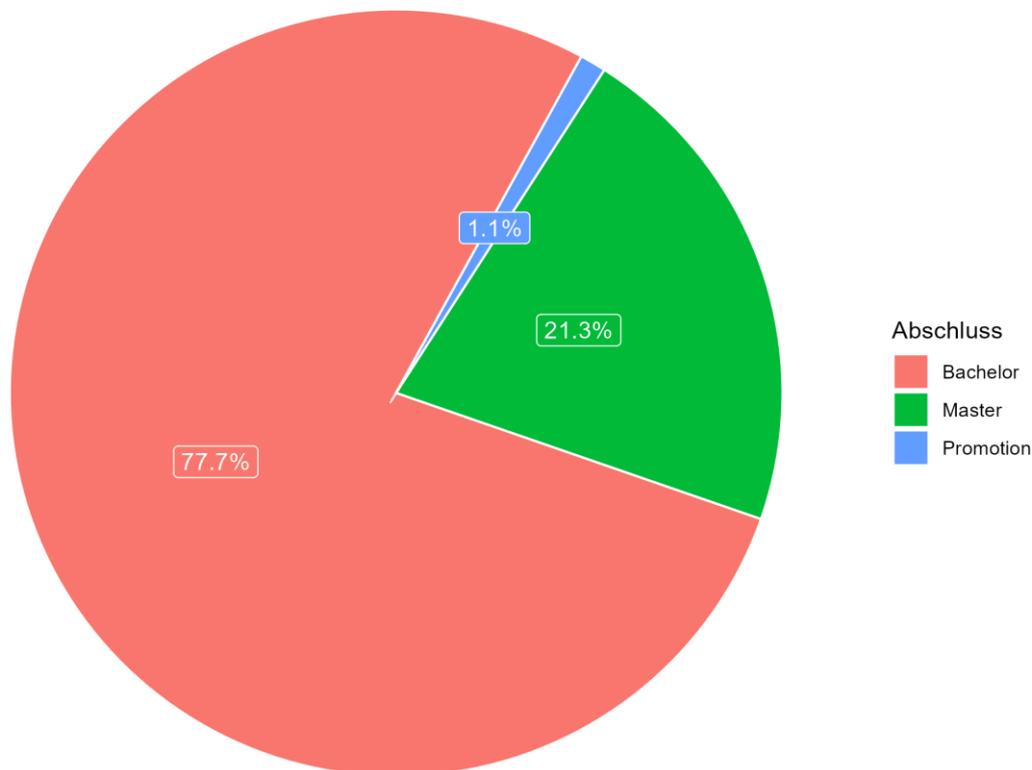


Abb. 10: Anteile der angestrebten Studienabschlüsse in % (n=94) (Quelle: Eigene Darstellung. Wegen Rundungsdifferenzen ergibt die Summe der Beschriftungen nicht 100%.)

77,7% Prozent der Teilnehmenden befinden sich aktuell in einem Bachelorstudienengang. 21,3% sind Masterstudierende und lediglich eine Person (1,1%) befindet sich im Promotionsstudium. Um einschätzen zu können, wie repräsentativ die Stichprobe in Hinsicht auf das Merkmal des angestrebten Studienabschlusses ist, wird die Anzahl der Bachelor- und Masterabschlüsse sowie der Promotionen an Hochschulen in Deutschland im Prüfungsjahr 2023 herangezogen: Während 243.633 Personen (58,47%) einen Bachelorabschluss absolvierten, vollendeten 146.484 Personen (35,15%) erfolgreich einen Masterstudiengang. 26.570 Personen (6,38%) promovierten (Statistisches Bundesamt, 2024).³ Im Vergleich wird deutlich, dass Bachelorstudierende in der Stichprobe überrepräsentiert sind, während Master- und Promotionsstudierende unterrepräsentiert sind.

³ Nicht berücksichtigt werden an dieser Stelle Staatsexamen sowie künstlerische und sonstige Abschlüsse, die sich von Bachelor- und Masterabschlüssen unterscheiden. Auch die Abbrecherquote wird nicht berücksichtigt.

9.2.2 Statistische Maßzahlen der Verteilung

Im Folgenden werden die in der Stichprobe vorherrschenden Einstellungen anhand des Medians, der im Gegensatz zum arithmetischen Mittelwert unempfindlich gegenüber Ausreißern ist (Kronthaler, 2021, S. 30), und der Schiefe beschrieben. Tabelle 21 fasst diese Werte für die jeweiligen Variablen zusammen:

Tab. 21: Mediane und Schiefen der Variablen (Quelle: Eigene Darstellung)

Variable	Median	Schiefe
(H1) time_effort	2,0	0,72
(H2) plagiarism	1,5	0,68
(H3) altruism	4,0	-0,78
(H4) diligence	3,0	-0,71
(H5) prestige	2,5	0,03
(H6) trustworthiness	3,0	-0,19
(H7a) career	3,0	-0,44
(H7b) acad_career	2,0	0,77
(H8) identification	2,5	-0,20
(AV) selfarchive	3,0	-0,84

Ein Median von 3,0 bei den linksschief verteilten Variablen *career* und *diligence* zeigt, dass die Studierenden in der Stichprobe eher mehr Wert auf die eigene schulische Leistung legen und überwiegend auch im Studium bereits die Grundlage für ihren künftigen beruflichen Erfolg legen möchten. Den Wunsch nach einer wissenschaftlichen Karriere verspüren sie dabei jedoch eher weniger. Bei der Variable *prestige* liegt eine leicht rechtsschiefe Verteilung vor, die sich jedoch auffällig nahe an einer symmetrischen Normalverteilung befindet (Abb. 11), sodass bei diesem Faktor keine starke Verzerrung in der Stichprobe vorliegt.

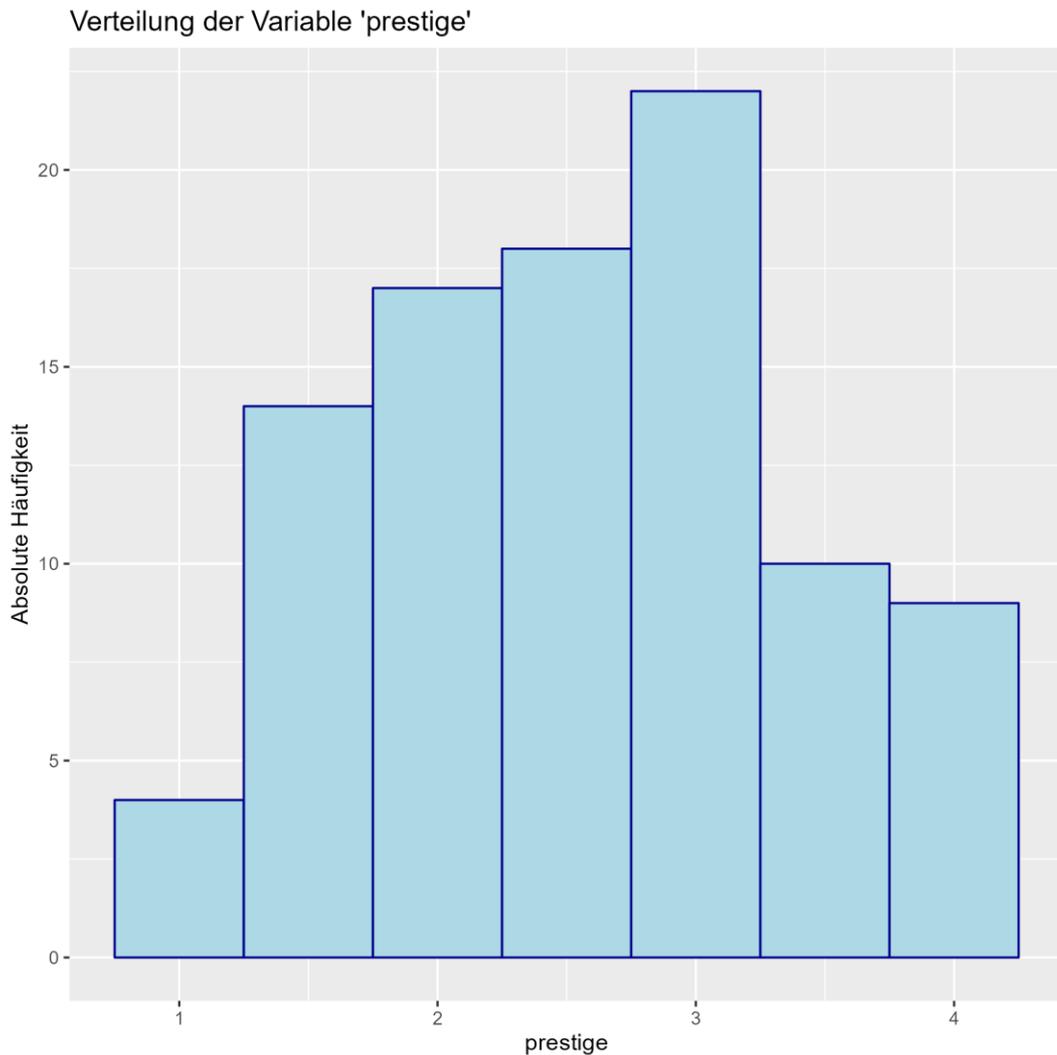


Abb. 11: Verteilung der Variable *prestige* ($n=94$) (Quelle: Eigene Darstellung)

Mediane von 2,5 bzw. 3,0 bei den linksschief verteilten Variablen *identification* und *trustworthiness* legen nahe, dass die Studierenden sich eher mehr mit ihrer Hochschule identifizieren und überwiegend auch Vertrauen in die qualitätssichernden Prozesse ihrer Institution haben. Den Zeit- und Arbeitsaufwand des Selbstarchivierens schätzen die Studierenden als eher gering ein. Die Variable *plagiarism* weist den niedrigsten Median (1,5) sowie eine rechtsschiefe Verteilung auf, was zeigt, dass die Furcht vor Plagiaten in der Stichprobe äußerst gering ausgeprägt ist. Stark ausgeprägt ist hingegen die altruistische Veranlagung der Studierenden, wie der hohe Median (4,0) der linksschief verteilten Variable *altruism* zeigt. Bei der abhängigen Variable *selfarchive* wird anhand des Medians von 3,0 und der linksschiefen Verteilung deutlich, dass die Studierenden in der Stichprobe eher mehr zur Selbstarchivierung bereit sind.

9.3 Regressionsanalyse

Um die in Kapitel 6 aufgestellten Hypothesen zu überprüfen, wird eine Regressionsanalyse nach der Methode der kleinsten Quadrate durchgeführt (s. Kap. 5.2.2). Neben einer Moderationsanalyse werden hierbei sowohl einfache lineare Regressionen für die Einzelhypothesen als auch eine multiple lineare Regression für das Gesamtmodell vorgenommen.

9.3.1 Einfache lineare Regression: Überprüfung der Einzelhypothesen

Im Zuge der einfachen linearen Regressionen wurden die linearen Modelle zunächst in Hinblick auf Heteroskedastizität überprüft. Die Streuung der Residuen war bei keinem der Modelle linear, was die Berechnung robuster Standardfehler erforderlich machte, um einer Verzerrung der p -Werte der Regressionskoeffizienten vorzubeugen (Hayes, 2007, S. 710-711; Kronthaler, 2021, S. 272). In der in Kapitel 9.1 durchgeführten Reliabilitätsanalyse wurde deutlich, dass mit dem Item *altruism3* („OA-Spirit“) vermutlich ein anderes Konstrukt als beabsichtigt gemessen wurde, was in Kapitel 10 näher ausgeführt wird. Auf Grundlage dieses Ergebnisses wurde bei der Regressionsanalyse zusätzlich ein möglicher gerichteter Zusammenhang zwischen dem Prädiktor *altruism3* („OA-Spirit“) und der AV untersucht, was im Folgenden als Zusatzüberprüfung (ZÜ) bezeichnet wird. Abbildungen 12 bis 20 zeigen die Streudiagramme der einfachen linearen Regressionen, wobei die Anzahl gleichartiger Datenpunkte anhand des Durchmessers der Punkte ausgedrückt wird:

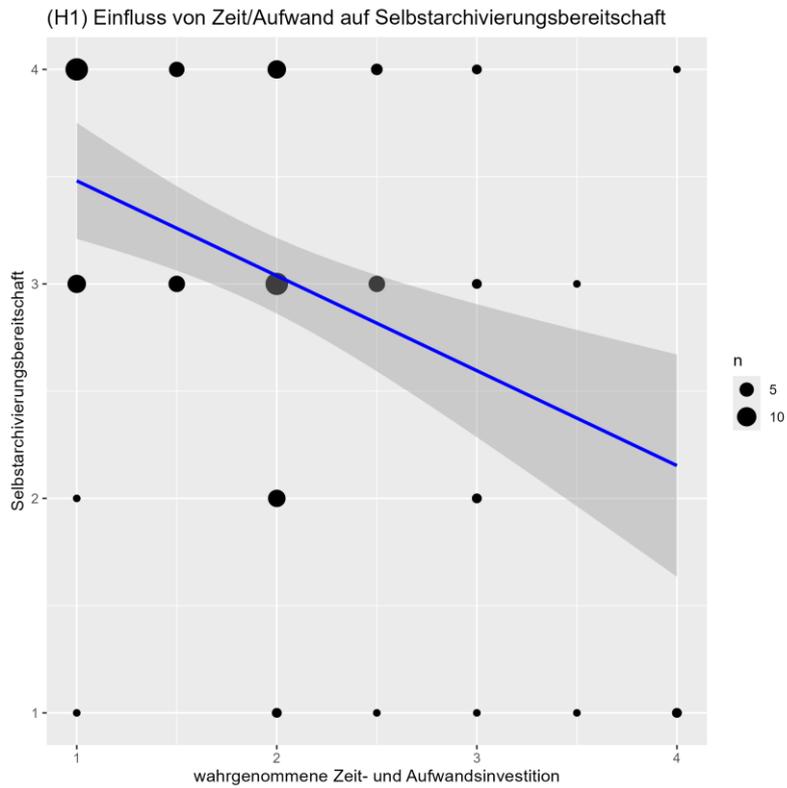


Abb. 12: Einfache lineare Regression der wahrgenommenen Zeit- und Aufwandsinvestition (Quelle: Eigene Darstellung)

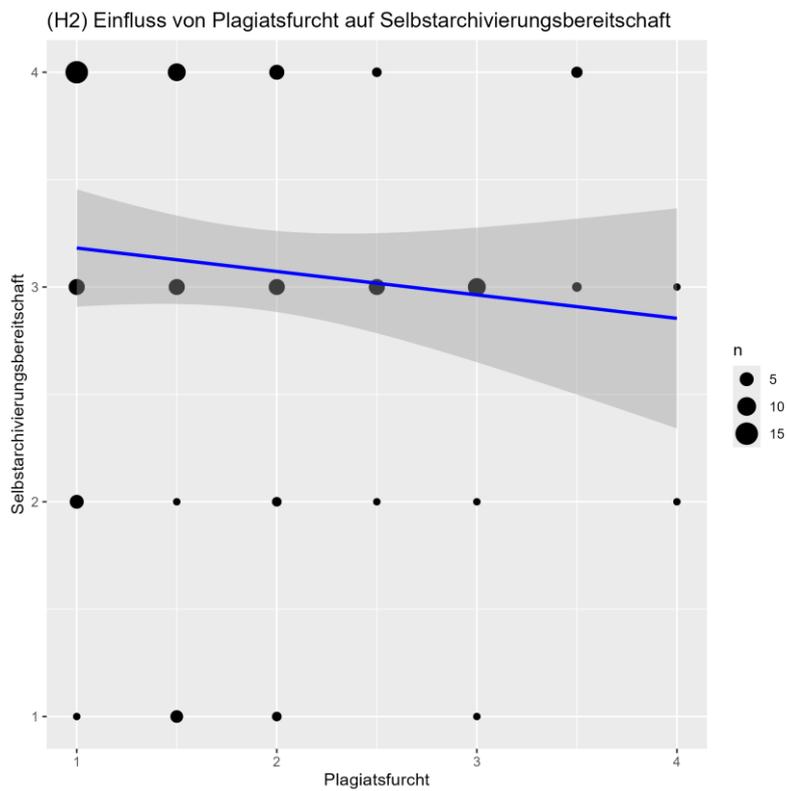


Abb. 13: Einfache lineare Regression der Plagiatsfurcht (Quelle: Eigene Darstellung)

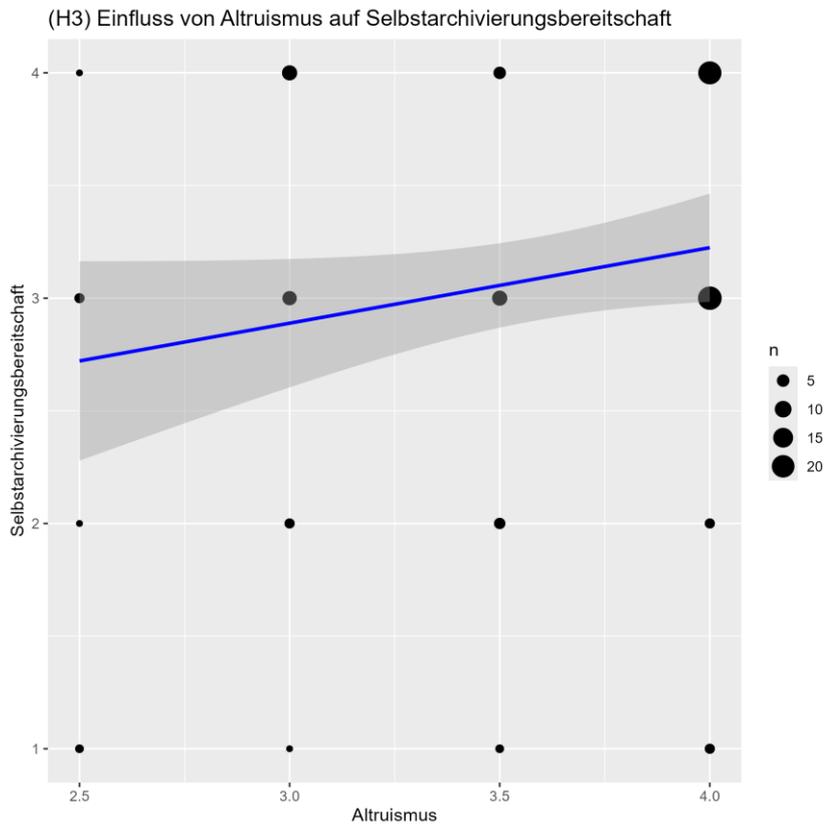


Abb. 14: Einfache lineare Regression des Altruismus
(Quelle: Eigene Darstellung)

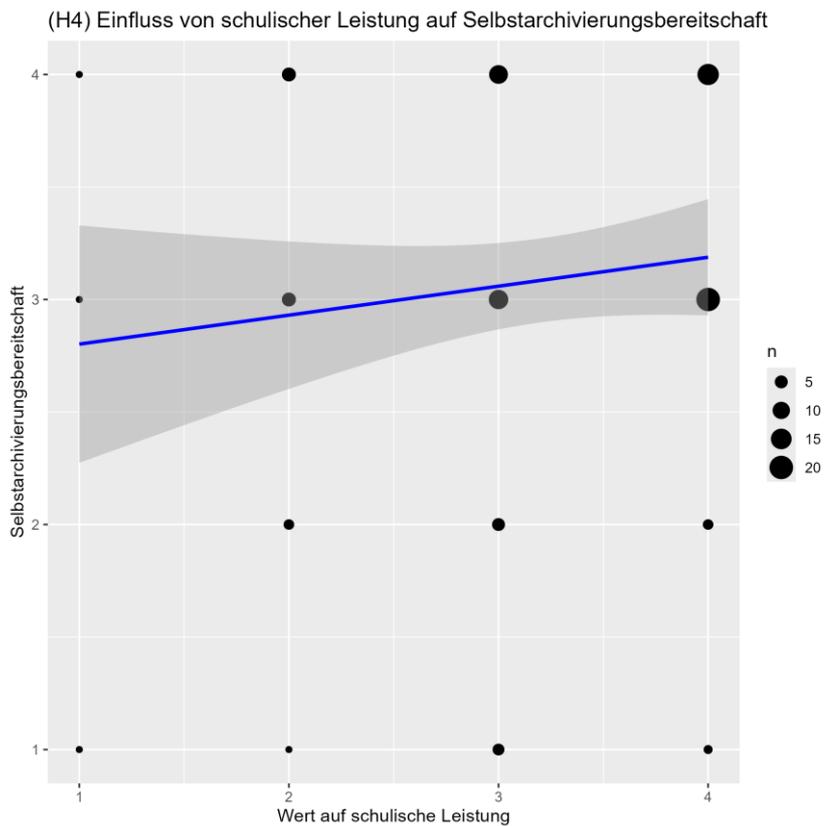


Abb. 15: Einfache lineare Regression des Wertes auf schulische Leistung
(Quelle: Eigene Darstellung)

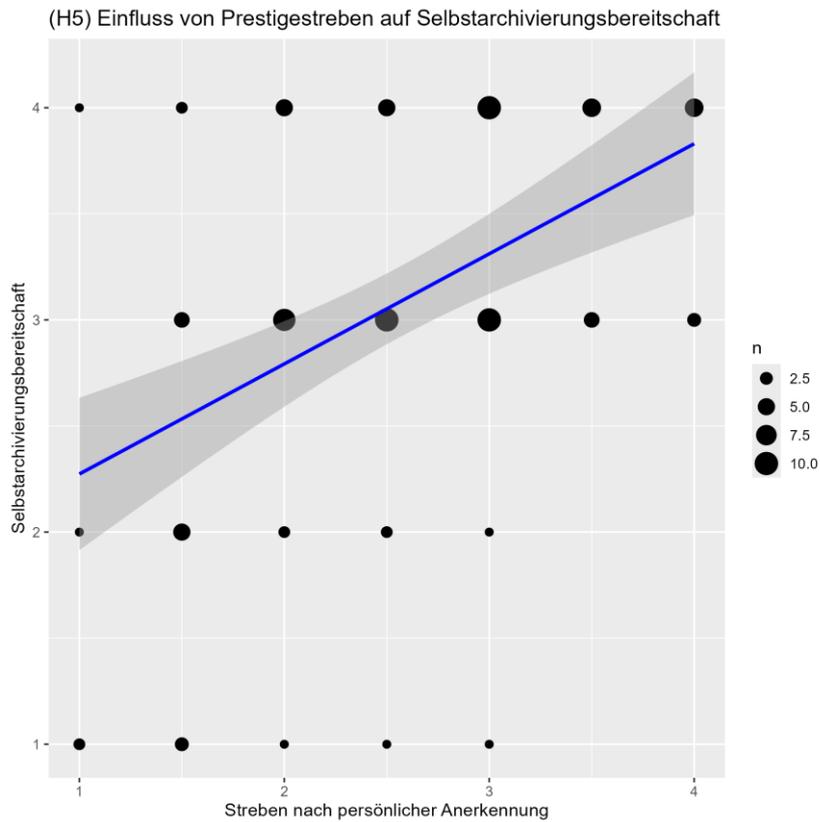


Abb. 16: Einfache lineare Regression des Strebens nach persönlicher Anerkennung (Quelle: Eigene Darstellung)

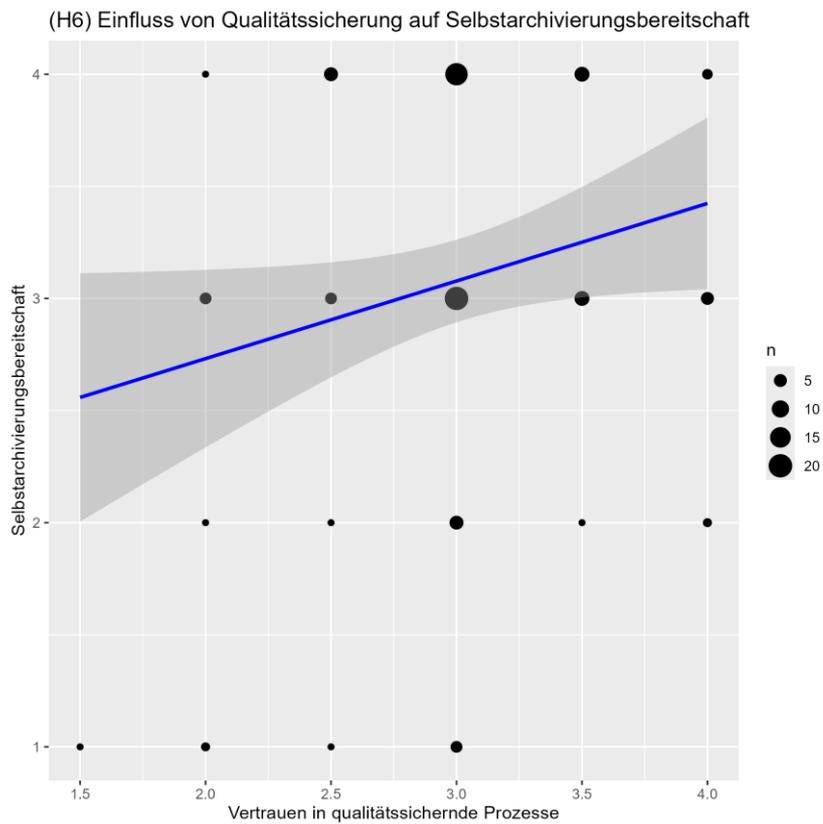


Abb. 17: Einfache lineare Regression des Vertrauens in qualitätssichernde Prozesse (Quelle: Eigene Darstellung)

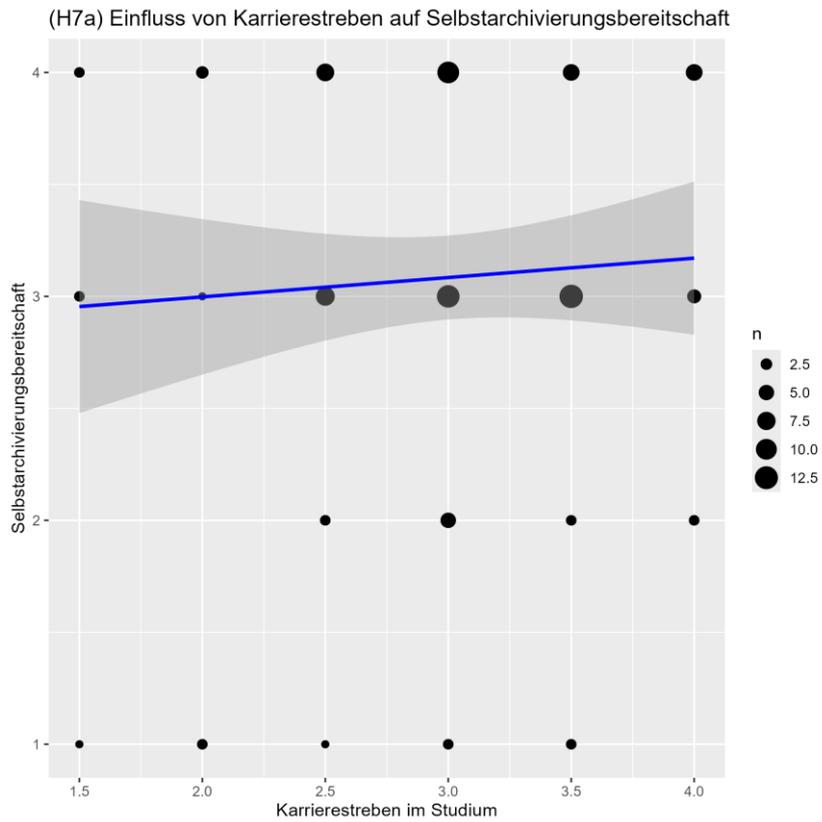


Abb. 18: Einfache lineare Regression des Karrierestrebens im Studium (Quelle: Eigene Darstellung)

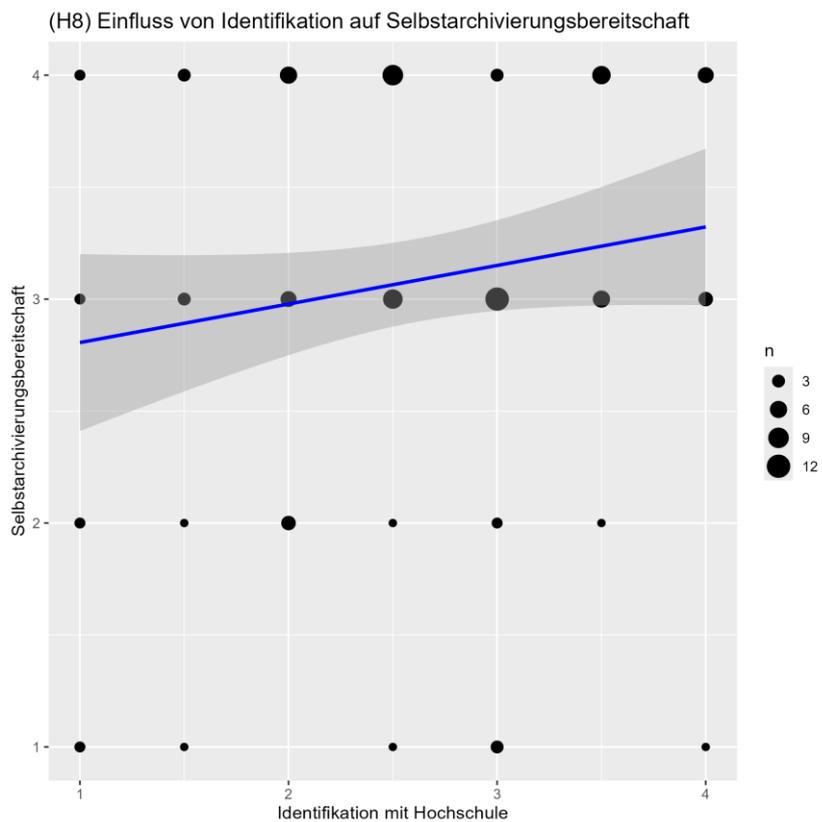


Abb. 19: Einfache lineare Regression der Identifikation mit der Hochschule (Quelle: Eigene Darstellung)

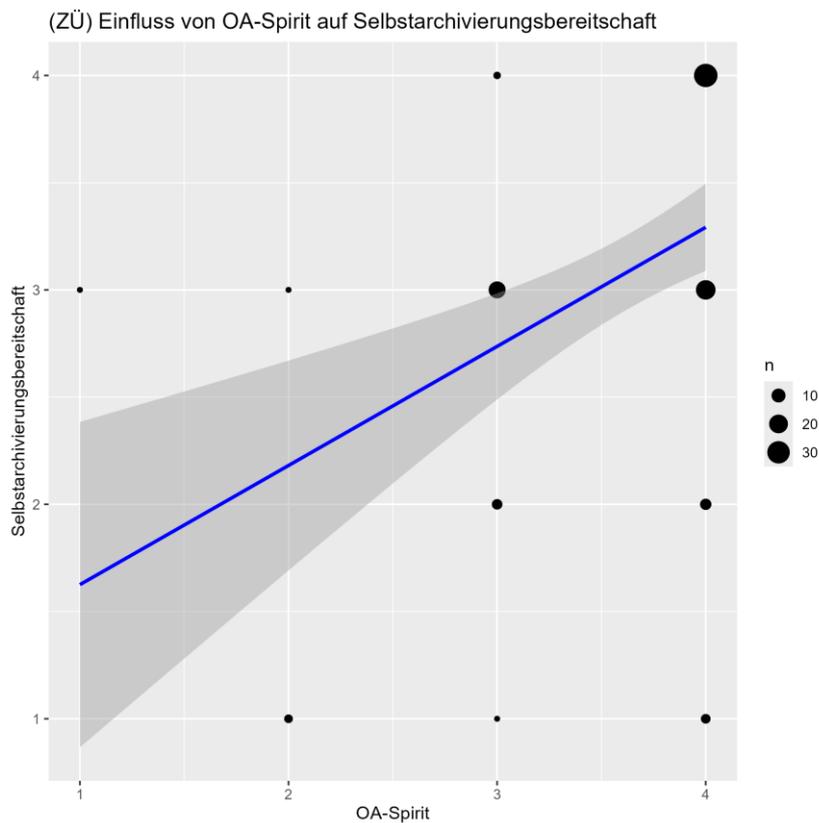


Abb. 20: Einfache lineare Regression des ‚OA-Spirits‘ (Quelle: Eigene Darstellung)

Anhand der Streudiagramme (Abb. 12-20) lässt sich beobachten, dass die Regressionslinien der in Kapitel 6 postulierten Kostenfaktoren in der Tendenz sinken, während die Regressionslinien der vermuteten Gewinn- sowie Kontextfaktoren tendenziell steigen. Die Konfidenzbänder zeigen auf, welcher Bereich den wahren Wert der Grundgesamtheit mit einer Wahrscheinlichkeit von 95% abdeckt (Kronthaler, 2021, S. 115-116). Tabelle 22 fasst die robusten p -Werte, Regressionskoeffizienten sowie Bestimmtheitsmaße (R^2) der einfachen linearen Regressionen zusammen:

Tab. 22: Ergebnisse der einfachen linearen Regressionen (Quelle: Eigene Darstellung)

Prädiktor	p -Wert (robust)	Koeffizient	R^2
time_effort	0,004	-0,443	0,135
plagiarism	0,290	-0,109	0,010
altruism	0,105	0,335	0,034
diligence	0,304	0,129	0,014
prestige	8,162e-06	0,519	0,218
trustworthiness	0,079	0,346	0,042

Tab. 22: Ergebnisse der einfachen linearen Regressionen (Fortsetzung)

career	0,603	0,086	0,004
identification	0,156	0,172	0,027
altruism3	0,004	0,555	0,144

Allein bei den Prädiktoren *time_effort*, *prestige* und *altruism3* liegen robuste *p*-Werte <0,05 vor. Die Bestimmtheitsmaße dieser signifikanten Zusammenhänge fallen zwar größer aus als bei den nicht signifikanten Zusammenhängen, bleiben dabei jedoch insgesamt niedrig.

9.3.2 Moderationsanalyse zu H7b

Im Rahmen einer Moderationsanalyse wurde überprüft, ob ein stärkerer Wunsch nach einer wissenschaftlichen Karriere (*acad_career*) den vermuteten positiven Einfluss, den das Karrierestreben (*career*) auf die Bereitschaft zur Selbstarchivierung in IR (*selfarchive*) ausübt, verstärkt (H7b). Der *F*-Statistik ist zu entnehmen, dass der *p*-Wert des Gesamtmodells 0,400 beträgt und damit deutlich über dem Signifikanzniveau von 0,05 liegt. Das Bestimmtheitsmaß fällt mit einem Wert von 0,032 äußerst niedrig aus. Tabelle 23 zeigt die robusten *p*-Werte und Koeffizienten des Modells:

Tab. 23: Ergebnisse der Moderationsanalyse zu H7b (Quelle: Eigene Darstellung)

Prädiktor	<i>p</i> -Wert (robust)	Koeffizient
career	0,851	-0,076
acad_career	0,895	-0,068
career x acad_career	0,656	0,073

Sowohl der *p*-Wert des Interaktionseffektes als auch die *p*-Werte der Haupteffekte liegen deutlich über dem Signifikanzniveau von 0,05.

9.3.3 Multiple lineare Regression: Überprüfung des Gesamtmodells

Auf die Untersuchung der Einzelhypothesen und die Moderationsanalyse folgt die Überprüfung des Gesamtmodells im Rahmen einer multiplen linearen Regression.

Hierbei wurde, wie in Kapitel 9.3.1 bereits begründet und in Kapitel 10 näher ausgeführt, auch das im Zuge der Reliabilitätsanalyse aus der Altruismus-Skala entfernte Item *altruism3* („OA-Spirit“) als eigenständiger Prädiktor miteinbezogen.

Weil starke Korrelationen zwischen den Prädiktoren die Aussagekraft der Koeffizienten verringern (Eckstein, 2016, S. 216), wurde zunächst eine Multikollinearitätsdiagnose durchgeführt. Um aufzuzeigen, inwieweit die Prädiktoren miteinander in Beziehung stehen, wurden *generalisierte Varianzinflationsfaktoren* (GVIF) ermittelt (Tab. 24), welche die Varianzen der Koeffizienten ausdrücken. Je stärker die UVs miteinander korrelieren, desto höher fallen die Varianzen aus. Im Gegensatz zum VIF vermag es der GVIF, auch Prädiktoren mit mehreren Freiheitsgraden (Df bzw. *degrees of freedom*) in die Diagnose miteinzubeziehen, was aufgrund des angenommenen Interaktionseffektes im Modell erforderlich wird. Um Werte mit einer unterschiedlichen Anzahl an Dimensionen miteinander vergleichen zu können, empfehlen Fox und Weisberg (2019, S. 432), den GVIF mit $1/(2 \cdot Df)$ zu potenzieren, was der Quadratwurzel des VIF gleichkommt. Bei eindimensionalen Werten entspricht der GVIF dem VIF (Fox & Monette, 1992, S. 180; Fox & Weisberg, 2019, S. 432; Marcoulides & Raykov, 2019, S. 875-876).

Tab. 24: Ergebnisse der Multikollinearitätsdiagnose (Quelle: Eigene Darstellung)

Prädiktor	GVIF	Df	GVIF $1/(2 \cdot Df)$	Interaktion
time_effort	1,233259	1	1,110522	/
plagiarism	1,230841	1	1,109433	/
altruism	1,288062	1	1,134928	/
diligence	1,495101	1	1,222743	/
prestige	1,208580	1	1,099354	/
trustworthiness	1,603029	1	1,266108	/
career	1,696595	3	1,092102	acad_career
acad_career	1,696595	3	1,092102	career
identification	1,647494	1	1,283547	/
altruism3	1,239625	1	1,113385	/

In der Literatur ist umstritten, welcher Wertebereich für den VIF als akzeptabel gilt; die diskutierten Werte reichen von ≤ 4 über ≤ 5 bis hin zu ≤ 10 (Eckstein, 2016, S. 216; Kronthaler & Zöllner, 2021, S. 105; Marcoulides, 2019, S. 876; Wooldridge, 2020, S. 92). Für diese Arbeit wird angenommen, dass ein $VIF \leq 4$ bzw.

$\text{GVIF}^{1/(2 \cdot \text{DF})} \leq \sqrt{4}$ Multikollinearität ausschließt. Weil bei allen Prädiktoren $\text{GVIF}^{1/(2 \cdot \text{DF})} < 2$ ist, liegt keine für die multiple Regression problematische Multikollinearität vor.

Die F -Statistik zeigt, dass der p -Wert des Gesamtmodells $1,303\text{e-}06$ beträgt und damit weit unter dem Signifikanzniveau von $0,05$ liegt. Das Bestimmtheitsmaß R^2 beträgt $0,428$. Vor der Betrachtung der Koeffizienten wird das Modell in Hinsicht auf Heteroskedastizität überprüft (Abb. 13):⁴

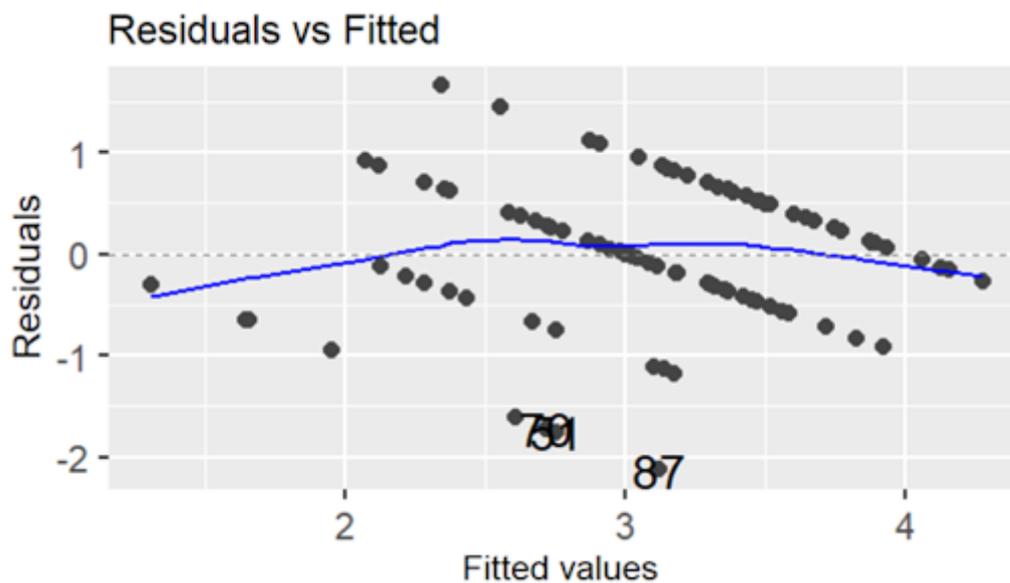


Abb. 21: Heteroskedastizität des Gesamtmodells (Quelle: Eigene Darstellung)

Erneut macht die nicht-lineare Streuung der Residuen, welche anhand der nicht-horizontal verlaufenden blauen Linie grafisch ersichtlich wird, die Berechnung robuster Standardfehler erforderlich, um einer Verzerrung der p -Werte der Regressionskoeffizienten vorzubeugen (s. Kap. 9.3.1). Tabelle 25 zeigt die robusten p -Werte und Koeffizienten der multiplen linearen Regression:

Tab. 25: Ergebnisse der multiplen linearen Regression (Quelle: Eigene Darstellung)

Prädiktor	p -Wert (robust)	Koeffizient
time_effort	0,016	-0,371
plagiarism	0,208	-0,110
altruism	0,722	0,073

⁴ Bei den numerisch beschrifteten Datenpunkten handelt es sich um auffällige Werte, die an dieser Stelle für die Interpretation jedoch keine hervorgehobene Rolle spielen.

Tab. 25: Ergebnisse der multiplen linearen Regression (Fortsetzung)

diligence	0,848	0,025
prestige	7,116e-05	0,406
trustworthiness	0,969	-0,009
career	0,885	0,051
acad_career	0,346	0,401
identification	0,545	0,071
altruism3	0,016	0,385
career x acad_career	0,510	-0,094

Die Ergebnisse der multiplen linearen Regression decken sich mit jenen der einfachen linearen Regressionen (s. Kap. 9.3.1): Erneut liegen für die drei Prädiktoren *time_effort*, *prestige* und *altruism3* robuste *p*-Werte $<0,05$ vor, während die übrigen Zusammenhänge deutlich über dem Signifikanzniveau bleiben.

Im folgenden Kapitel werden die Studienergebnisse zusammenfassend dargestellt.

9.4 Ergebnisse

Die einfachen linearen Regressionen haben ergeben, dass die robusten *p*-Werte der in den Hypothesen 2-4 sowie 6, 7a und 8 angenommenen Zusammenhänge deutlich über dem Signifikanzniveau von 0,05 liegen (s. Tab. 22). Auch der in H7b vermutete Moderationseffekt ist nicht signifikant (s. Tab. 23), sodass die Irrtumswahrscheinlichkeit hier ebenfalls hoch ist. Die Nullhypothese lässt sich nicht verwerfen, weshalb die Interpretation nicht sinnvoll fortgesetzt werden kann. Bei den Prädiktoren *time_effort* (H1), *prestige* (H5) und *altruism3* (ZÜ) hingegen liegen robuste *p*-Werte $<0,05$ vor, sodass die Irrtumswahrscheinlichkeit gering ist. Die Nullhypothese kann verworfen und stattdessen die Alternativhypothese angenommen werden. Die Bestimmtheitsmaße der drei signifikanten Prädiktoren fallen zwar größer aus als bei den nicht signifikanten, bleiben dabei jedoch insgesamt niedrig. Dieses Phänomen tritt in den Sozialwissenschaften, in denen oft mit schwer messbaren und daher fehleranfälligen Konzepten gearbeitet wird, allerdings häufig auf, weshalb auch sehr kleine R^2 als akzeptabel gelten (Kronthaler, 2021, S. 251-252;

Wooldridge, 2020, S. 35). Dementsprechend üben die niedrigen Bestimmtheitsmaße an dieser Stelle keinen weiteren Einfluss auf die Interpretation aus.

Die Betrachtung der Koeffizienten der einfachen linearen Regressionen zeigt, welchen Einfluss die signifikanten Prädiktoren isoliert auf die AV ausüben. Beim Prädiktor *time_effort* (H1) beträgt der Koeffizient -0,443, sodass ein negativer Zusammenhang zwischen den Variablen besteht: Wenn die wahrgenommene Zeit- und Aufwandsinvestition um eine Einheit steigt, dann sinkt die Bereitschaft zur Selbstarchivierung in IR um 0,443. Beim Prädiktor *prestige* (H5) beträgt der Koeffizient 0,519, sodass ein positiver Zusammenhang besteht: Wenn das Streben nach persönlicher Anerkennung um eine Einheit steigt, dann steigt die Selbstarchivierungsbereitschaft um 0,519. Beim Prädiktor *altruism3* (ZÜ) beträgt der Koeffizient 0,555, sodass erneut ein positiver Zusammenhang besteht: Wenn der ‚OA-Spirit‘ um eine Einheit steigt, dann steigt die Bereitschaft zur Selbstarchivierung um 0,555.

Im Rahmen der multiplen linearen Regression wurde untersucht, welchen Einfluss die Prädiktoren im Zusammenspiel auf die AV ausüben. Der hochsignifikante *p*-Wert in der *F*-Statistik zeigt, dass die Nullhypothese verworfen und stattdessen die Alternativhypothese angenommen werden kann, nach der die Prädiktoren linear auf die AV abbilden. Das Modell vermag die Varianz in der AV zu 43% zu erklären. Die in den einfachen Regressionen signifikanten Prädiktoren *time_effort* (H1), *prestige* (H5) und *altruism3* (ZÜ) behalten auch in der multiplen Regression ihre Signifikanz bei. Die übrigen Prädiktoren bleiben weiterhin insignifikant und leisten keinen Erklärungsbeitrag (s. Tab. 25). Im Zusammenspiel behalten die Koeffizienten der signifikanten Prädiktoren ihre Vorzeichen bei, verringern sich dabei jedoch betragsmäßig: Beim Prädiktor *prestige* (H5) sinkt der Koeffizient auf 0,406, womit die Variable den größten Einfluss⁵ auf die AV ausübt. Beim Prädiktor *altruism3* (ZÜ) sinkt der Koeffizient auf 0,385, womit die Variable den zweitgrößten Einfluss ausübt. Beim Prädiktor *time_effort* (H1), der den drittgrößten Einfluss auf die AV ausübt, sinkt der Koeffizient auf -0,371.

Im folgenden Kapitel werden die Ergebnisse der Regressionsanalyse diskutiert, wobei auch auf die statistischen und methodischen Limitationen der Studie eingegangen wird.

⁵ Weil die Wertebereiche der Variablen einheitlich zwischen 1 und 4 liegen, ist keine Z-Standardisierung erforderlich, um die Koeffizienten miteinander vergleichbar zu machen (Eckstein, 2016, S. 110).

10 Diskussion und Limitationen

Der in dieser Arbeit durchgeführten empirischen Studie liegt die Hypothese zugrunde, dass sich anhand von Gewinn-, Kosten- und Kontextfaktoren die Selbstarchivierungsbereitschaft von Studierenden erklären lässt. Die Ergebnisse der Studie liefern Anhaltspunkte, die diese Hypothese im Fall der Gewinn- und Kostenfaktoren unterstützen. Der signifikante p -Wert des Gesamtmodells (s. Kap. 9.4) zeigt, dass die Prädiktoren wie vermutet linear auf die AV abbilden, sodass die gewählte Methode einer linearen Regressionsanalyse geeignet für das Untersuchen der gestellten Forschungsfrage ist. Dieses Erkenntnis muss mit der Einschränkung betrachtet werden, dass möglicherweise nicht alle für die Selbstarchivierungsbereitschaft von Studierenden in IR relevanten Faktoren im empirischen Modell berücksichtigt wurden. Weil weitere, bisher unbekannte Faktoren eine Rolle spielen und das reale Verhalten der Studierenden beeinflussen können, lassen sich aufgrund der Ergebnisse keine gesicherten Prognosen aufstellen.

Auch sind die Ergebnisse der Regressionsanalyse vor dem Hintergrund der statistischen Limitationen der empirischen Studie zu betrachten. Zunächst ist kritisch anzumerken, dass der Fragebogen nicht die für Likert-Skalen übliche Skalenkonstruktion durchlaufen hat, bei der die Items vor der Selektion des finalen Item-Pools im Rahmen eines quantitativen Pretests überprüft werden (Döring, 2023, S. 271-272). Darüber hinaus ist auf die Einseitigkeit der Stichprobe hinzuweisen, die dazu führt, dass die sozialwissenschaftlichen Standards für Repräsentativität nicht erfüllt sind. Wie im Zuge der Stichprobenkontrolle bereits dargestellt (s. Kap. 9.2.1), sind Frauen sowie Bachelorstudierende in der Stichprobe überrepräsentiert. Weil sich die Ansichten zur Selbstarchivierungsbereitschaft im Laufe eines Studiums möglicherweise verändern könnten, könnte insbesondere Letzteres die Stichprobe verzerren.

Besonders problematisch in Bezug auf den Untersuchungsgegenstand ist, dass nur Studierende einseitig ausgewählter Fachbereiche befragt wurden. So entstammen die Teilnehmenden vorrangig den Informations- und Kommunikations- sowie Kulturwissenschaften, während Geistes- und Naturwissenschaftler:innen stark unterrepräsentiert sind. Andere Fachdisziplinen, wie z. B. Wirtschafts- und Ingenieurwissenschaften oder Psychologie, sind überhaupt nicht vertreten. Diese Homogenität der Stichprobe ist kritisch zu betrachten, weil unterschiedliche Fachdisziplinen über unterschiedlich ausgeprägte Selbstarchivierungskulturen verfügen (s. Kap. 5.2.1), weshalb davon auszugehen ist, dass eine Unausgewogenheit bei

den studierten Fächern die Stichprobe hinsichtlich der Selbstarchivierungsbereitschaft verzerrt. So ist auch die Feststellung, dass die befragten Studierenden tendenziell eher mehr zur Selbstarchivierung bereit sind (s. Kap. 9.2.2), vor dem Hintergrund zu betrachten, dass Studierende des Bibliotheks- und Informationsmanagements an der HAW Hamburg, die bereits im Studium für OA-Themen sensibilisiert werden, in der Stichprobe stark überrepräsentiert sind (s. Kap. 9.2.1). Dass die Teilnehmenden vorrangig Hamburger Hochschulen besuchen (s. Kap. 8.2), könnte möglicherweise auch Auswirkungen auf die Faktoren der Identifikation sowie der Vertrauenswürdigkeit haben, die in der Stichprobe tendenziell eher hoch bewertet werden (s. Kap. 9.2.2). Insgesamt ist die gefundene Zielgruppe relevant für die gestellte Forschungsfrage, dabei jedoch überrepräsentiert, sodass sich auf Grundlage der Ergebnisse keine generalisierten Aussagen über die Selbstarchivierungsbereitschaft von Studierenden im Allgemeinen treffen lassen.

Im Rahmen der Datenanalyse wurden zwei signifikante Gewinnfaktoren sowie ein signifikanter Kostenfaktor identifiziert. Den größten Einfluss auf die AV übt der Gewinnfaktor der persönlichen Anerkennung (*prestige*) aus, der als einziger Faktor nicht auf dem Kim-Modell basiert, sondern allein aus der bestehenden Literatur zu studentischen Inhalten in IR abgeleitet wurde. Weil die verwendeten Items potenzielle Publikationsmotivatoren widerspiegeln, die in der Literatur von Studierenden selbst genannt wurden (s. Kap. 5.3.1), haben sie einen direkten Realitätsbezug, der sich möglicherweise positiv auf die Konstruktvalidität ausgewirkt haben könnte. Bei der Reliabilitätsanalyse wurde eine akzeptable interne Konsistenz der Items untereinander festgestellt (s. Kap. 9.1), sodass das Konstrukt hinreichend abgebildet wurde. Trotzdem ist bezüglich der Operationalisierung kritisch anzumerken, dass die Skala lediglich zwei Items umfasst. Dies verringert u. U. die Konstruktvalidität, weil die Verwendung mehrerer heterogener Items die Wahrscheinlichkeit erhöht, ein gewünschtes Konstrukt adäquat abzubilden (Eisinga et al., 2013, S. 637).

Abbildung 14 zeigt die relative Selbstarchivierungsbereitschaft nach Prestigestreben. Hierfür wurden die Teilnehmenden in zwei Gruppen eingeteilt, wobei eine Gruppe eher weniger nach persönlicher Anerkennung strebt ($prestige \leq 2$, insgesamt 35 Teilnehmende) und eine Gruppe hiernach eher mehr strebt ($prestige > 2$, insgesamt 59 Teilnehmende).

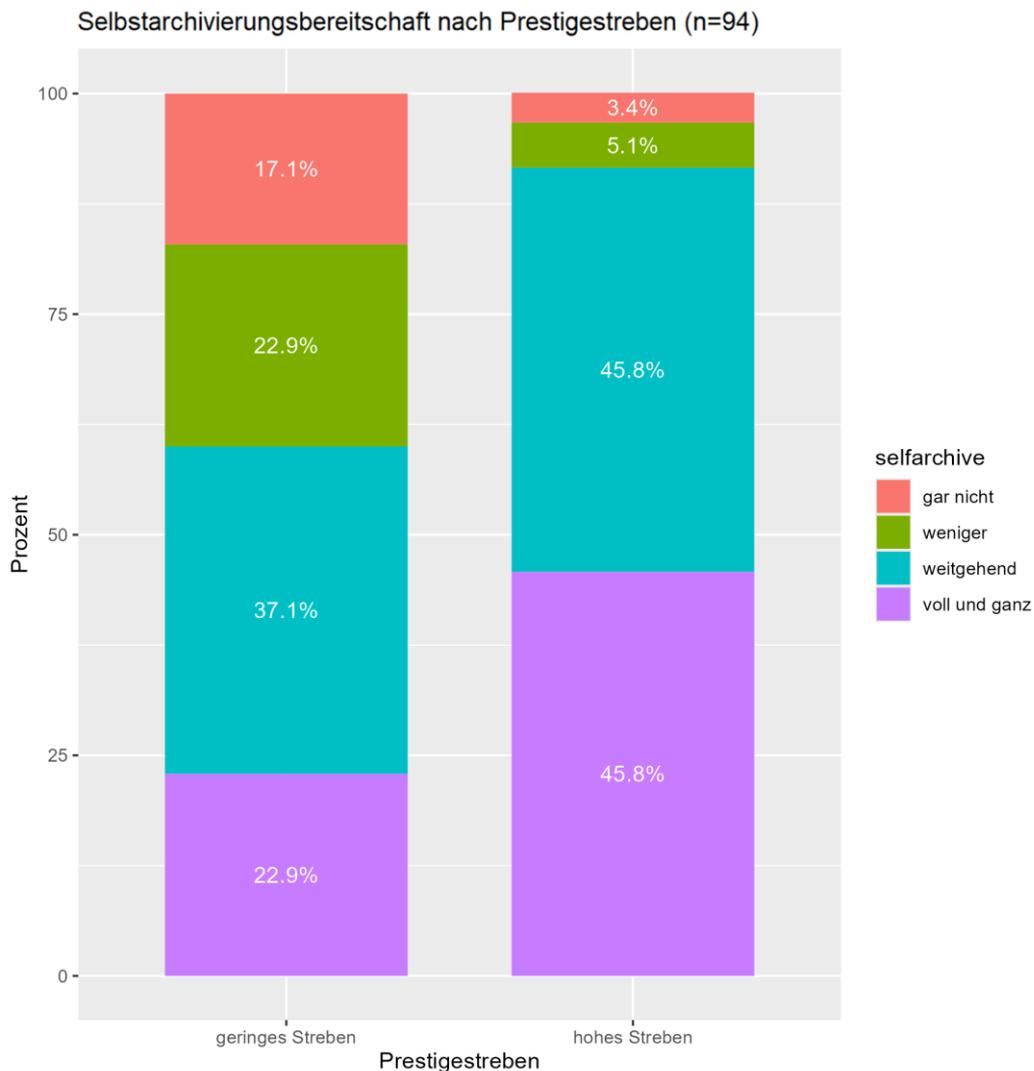


Abb. 22: Selbstarchivierungsbereitschaft nach Prestigestreben (Quelle: Eigene Darstellung. Wegen Rundungsdifferenzen ergibt die Summe der Beschriftungen nicht 100%.)

Die Abbildung führt vor Augen, dass Studierende, die eher mehr nach persönlicher Anerkennung streben, weitaus mehr zur Selbstarchivierung bereit sind als jene mit eher geringem Prestigestreben. Während 91,6% derjenigen mit hohem Prestigestreben weitgehend oder voll und ganz zur Selbstarchivierung bereit sind, trifft dies auf lediglich 60,0% derjenigen mit geringem Prestigestreben zu. Umgekehrt sind 17,1% letzterer gar nicht zur Selbstarchivierung bereit, was nur von 3,4% derjenigen mit hohem Prestigestreben angegeben wird. Die starken Unterschiede in den beiden Gruppen visualisieren die hohe Relevanz des Prädiktors. Das Ergebnis liefert einen deutlichen Hinweis darauf, dass tatsächlich – wie in H5 vermutet – ein positiver Zusammenhang zwischen dem Streben nach persönlicher Anerkennung und der Bereitschaft zur Selbstarchivierung in IR bestehen könnte, weshalb es sich lohnt, diese Hypothese in Nachfolgestudien näher zu untersuchen.

Den zweitgrößten Einfluss auf die AV übt der Gewinnfaktor des ‚OA-Spirits‘ aus, der den Grad der Zustimmung bezeichnet, den die Teilnehmenden einer Philosophie der Offenheit im Wissenschaftskontext beimessen. Das Item war ursprünglich Teil der Altruismus-Skala, wurde jedoch im Zuge der Reliabilitätsanalyse aus dieser entfernt (s. Kap. 9.1) und anschließend im Rahmen einer Zusatzüberprüfung als eigenständiger Prädiktor untersucht (s. Kap. 9.3.1). Wie in Kapitel 7.4 erläutert, geht das Item auf Kims Operationalisierung ihrer eigenen Altruismus-Variable zurück, bei der jedes Item konkreten Bezug auf das wissenschaftliche Publizieren nimmt. Es ist zu vermuten, dass die Übertragung auf den studentischen Kontext bei diesem Faktor nicht gelungen ist. In Kapitel 5.3.1 wurden in Kims Forschung zwei Facetten des Altruismus-Konstruktes identifiziert, wobei die erste Facette den ‚OA-Spirit‘ umfasst und die zweite Facette eine allgemeine Hilfsbereitschaft ausdrückt. Letztere Facette kam jedoch erst in den Interviews zutage, die auf Kims empirische Studie folgten (2010, S. 1916), und war somit kein Bestandteil der Operationalisierung ihres Fragebogens. Die Ergebnisse dieser Arbeit weisen darauf hin, dass es sich bei ‚OA-Spirit‘ und Hilfsbereitschaft nicht – wie zunächst angenommen – um zwei Facetten desselben Konstruktes, sondern vielmehr um zwei unterschiedliche Konstrukte handelt, wodurch die für diese Arbeit konzipierte Altruismus-Skala unbeabsichtigt mehrdimensional wurde.

Abbildung 15 zeigt die relative Selbstarchivierungsbereitschaft nach ‚OA-Spirit‘. Hierfür wurden die Teilnehmenden erneut in zwei Gruppen eingeteilt, wobei eine Gruppe eine Philosophie der Offenheit im Wissenschaftskontext eher weniger teilt ($altruism3 \leq 2$) und eine Gruppe diese eher mehr teilt ($altruism3 > 2$):

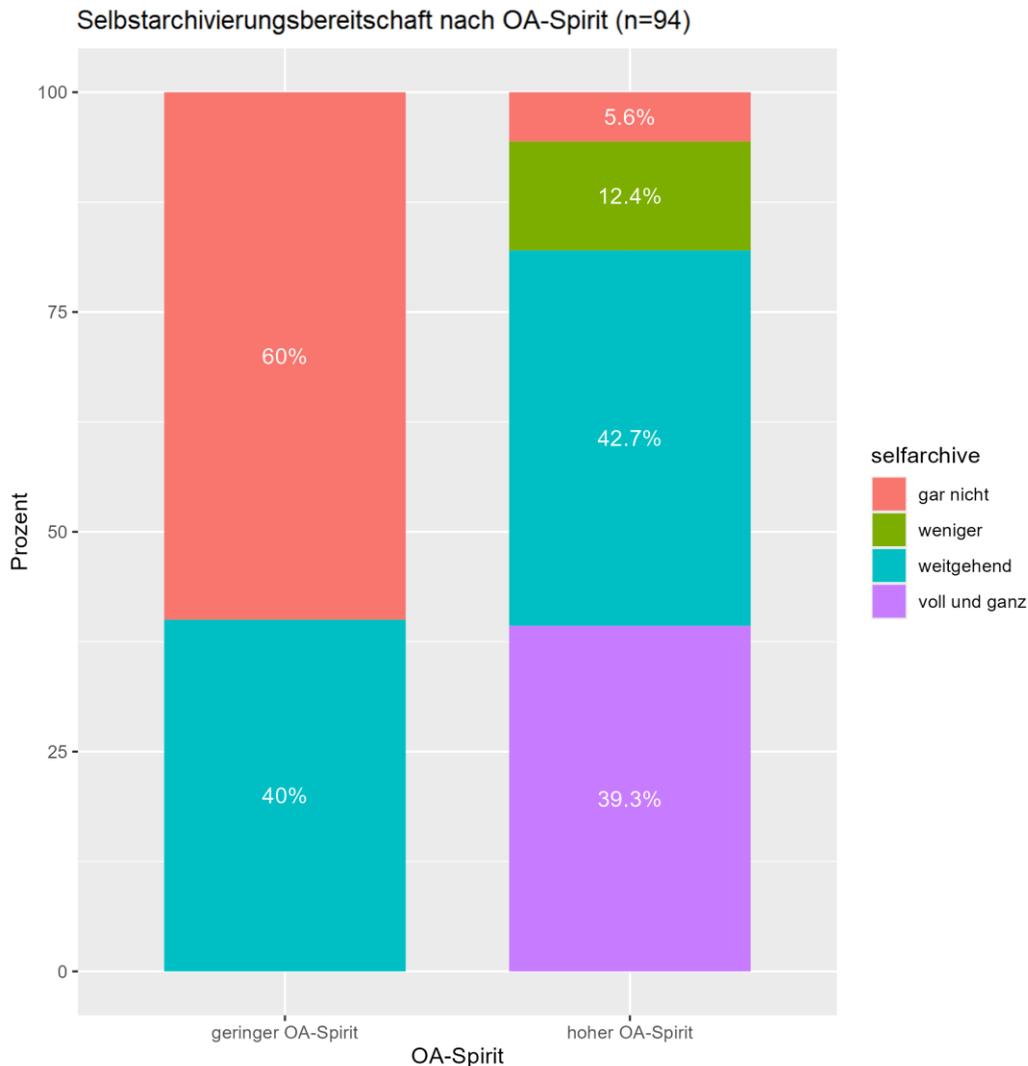


Abb. 23: Selbstarchivierungsbereitschaft nach ‚OA-Spirit‘ (Quelle: Eigene Darstellung)

Die Unterschiede zwischen den beiden Gruppen sind plakativ: Während 39,3% derjenigen mit hohem ‚OA-Spirit‘ voll und ganz zur Selbstarchivierung in IR bereit sind, trifft dies auf keine einzige Person mit geringem ‚OA-Spirit‘ zu. Gleichzeitig schließt mit 60% über die Hälfte derjenigen mit geringem ‚OA-Spirit‘ eine Selbstarchivierung kategorisch aus, was auf lediglich 5,6% der Gruppe mit hohem ‚OA-Spirit‘ zutrifft. Auffällig ist zudem, dass in der Gruppe mit niedrigem ‚OA-Spirit‘ niemand weniger, dafür jedoch 40% weitgehend zur Selbstarchivierung bereit sind, sodass das Selbstarchivieren entweder auf starke Ablehnung oder zögerliche Zustimmung trifft. So gibt es also eine Gruppe von Studierenden, die zwar persönlich weitgehend zur Selbstarchivierung bereit sind, dabei jedoch die Idee von Forschungsergebnissen als öffentliches Gut ablehnen. Die Vermutung, dass sich die Publikationsmotivation dieser Gruppe eventuell durch ein hohes Prestigestreben erklären ließe, wird durch die vorliegenden Daten nicht unterstützt, weil die Gruppe zu glei-

chen Anteilen aus Studierenden mit hohem bzw. geringem Prestigestreben besteht.

Die Interpretation des ‚OA-Spirits‘ muss an dieser Stelle durch den Hinweis eingeschränkt werden, dass diese Eigenschaft in der Stichprobe sehr hoch ausgeprägt ist: In absoluten Zahlen ausgedrückt besteht die Gruppe jener mit hohem ‚OA-Spirit‘ aus 89 Teilnehmenden, während die Gruppe jener mit geringem ‚OA-Spirit‘ lediglich fünf Studierende umfasst, sodass Nachfolgeuntersuchungen anhand einer fachdisziplinär repräsentativen Stichprobe vonnöten sind. Trotzdem kann festgehalten werden, dass sich die in dieser Arbeit festgestellte Signifikanz des ‚OA-Spirits‘ mit Kims Ergebnissen deckt, deren Altruismus-Konstrukt – wie oben erläutert – inhaltlich mit ‚OA-Spirit‘ gleichzusetzen ist und in ihren Studien konstant signifikant war. Somit ist das Ergebnis ein Hinweis darauf, dass auch bei Studierenden ein positiver Zusammenhang zwischen dem ‚OA-Spirit‘ und der Selbstarchivierungsbereitschaft in IR besteht, den es sich in Nachfolgestudien zu untersuchen lohnt. Um die Konstruktvalidität zu erhöhen, sollte die Operationalisierung des ‚OA-Spirits‘ dabei durch weitere Items ergänzt werden.

Nicht signifikant hingegen ist die Altruismus-Variable dieser Arbeit, die mit Hilfsbereitschaft gleichzusetzen ist (s. oben). Es lässt sich vermuten, dass bei dem in der Stichprobe stark ausgeprägten Altruismus (s. Kap. 9.2.2) auch bewusstes Antwortverhalten in Form von sozialer Erwünschtheit eine Rolle gespielt und sich auf den p -Wert niedergeschlagen haben könnte. So ist davon auszugehen, dass die Teilnehmenden beim Ausfüllen des Fragebogens den sozialen Druck verspürten, sich als hilfsbereit darzustellen, doch ob sie anderen im Ernstfall wirklich zu Hilfe kommen, bleibt ungewiss. Insgesamt liefert das Ergebnis keinen Hinweis, der die in H3 formulierte Vermutung unterstützt, dass ein positiver Zusammenhang zwischen einer altruistischen Veranlagung (im Zuge der Reliabilitätsanalyse und Diskussion präziser definiert als Hilfsbereitschaft) und der Bereitschaft zur Selbstarchivierung in IR besteht. Möglicherweise würde sich der vermutete Effekt durch eine überarbeitete Operationalisierung, die den Einfluss der sozialen Erwünschtheit bestmöglich minimiert, wahrnehmen lassen.

Den drittgrößten Einfluss auf die AV übt der Kostenfaktor des zusätzlichen Zeit- und Arbeitsaufwands (*time_effort*) aus. Die Items *time_effort1* („Es ist leicht, das von meiner Hochschule geforderte Upload-Formular auszufüllen“) und *time_effort3* („Es ist aufwendig, meine Arbeitsergebnisse auf der Website meiner Hochschule zu veröffentlichen“) sollten unterschiedliche Formulierungen derselben Frage darstel-

len. Tatsächlich wurde dies von den Teilnehmenden jedoch nicht so aufgefasst, wie die stark rechtsschiefe Verteilung bei *time_effort1* zeigt (Skewness von 1,27 und Median von 1,0), die sich bei *time_effort3* (Skewness von 0,77 und Median von 2,0) so nicht widerspiegelt. Möglicherweise bestanden Unklarheiten bei der Mockup-Grafik, die im qualitativen Pretest nicht zutage kamen. Auch besteht die Möglichkeit, dass es neben dem Ausfüllen des Upload-Formulars noch weitere Aspekte des Selbstarchivierungsprozesses gibt, die die Studierenden als aufwendig empfinden. Diese wahrgenommen, aber nicht explizit abgefragten Hürden könnten das Antwortverhalten unterschwellig beeinflusst haben. Auch die negative Polung des Items *time_effort1* könnte problematisch gewesen sein, weil die Vermischung positiver und negativer Items die Gefahr einer ungewollten Mehrdimensionalität der Skala erhöht, bei der Korrelationen nur zwischen den positiven bzw. negativen Items untereinander bestehen (Döring, 2023, S. 272). Während die Existenz eines negativ gepolten Items bei der *prestige*-Variable die interne Konsistenz nicht auffällig geschmälert hat, führt das Entfernen des negativen Items beim Faktor des Zeit- und Arbeitsaufwands zu einer deutlichen Verbesserung der internen Konsistenz, weshalb es im Zuge der Reliabilitätsanalyse entfernt wurde (s. Kap. 9.1).

Abbildung 16 zeigt die relative Selbstarchivierungsbereitschaft nach dem wahrgenommenen Zeit- und Arbeitsaufwand. Auch dieses Mal wurden die Teilnehmenden in zwei Gruppen eingeteilt, wobei eine Gruppe den Aufwand einer Selbstarchivierung in IR als eher gering einschätzt (*time_effort* ≤ 2) und eine Gruppe diesen als eher hoch einschätzt (*time_effort* > 2):

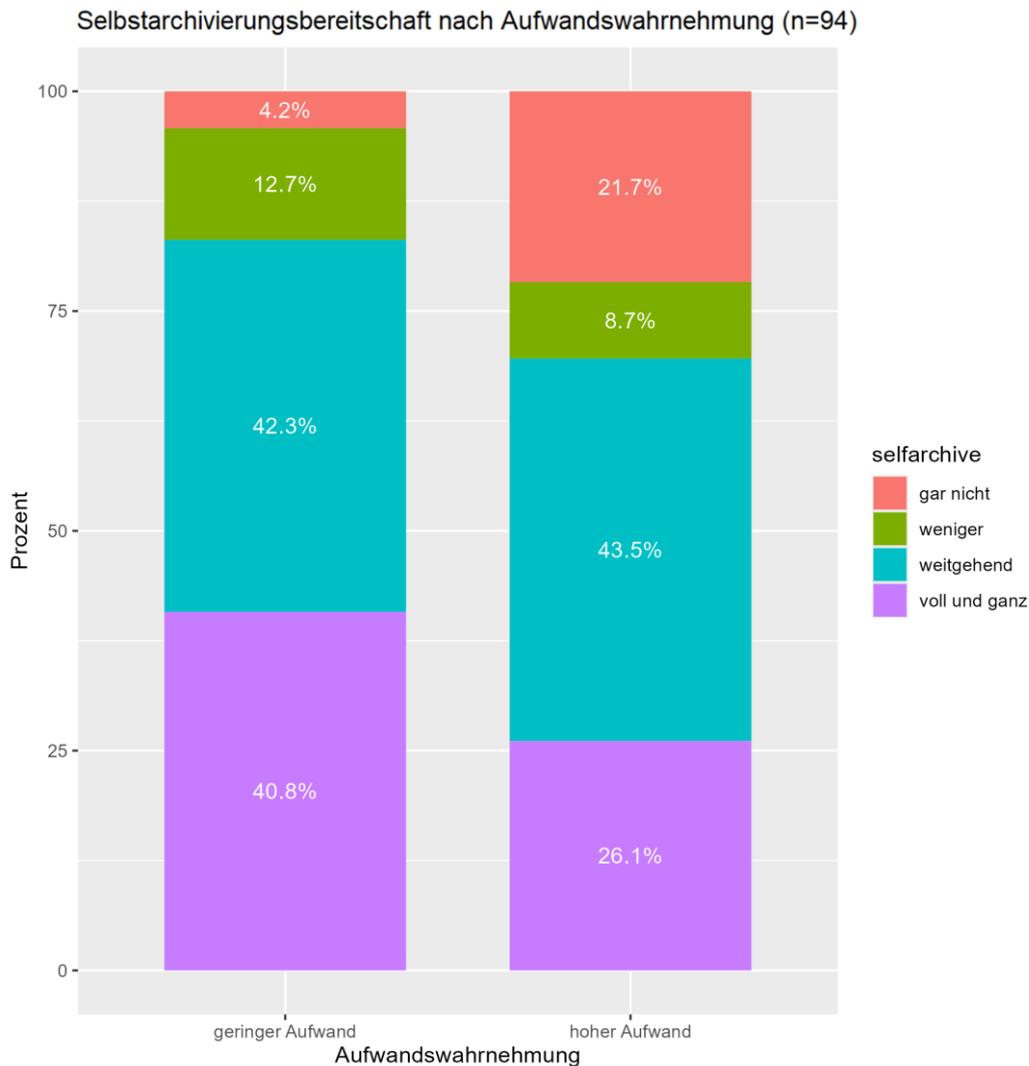


Abb. 24: Selbstarchivierungsbereitschaft nach Aufwandswahrnehmung (Quelle: Eigene Darstellung)

Das Mittelfeld jener Studierenden, die weniger oder weitgehend zur Selbstarchivierung bereit sind, ist in den beiden Gruppen ähnlich verteilt, doch an den Rändern lassen sich erneut deutliche Unterschiede feststellen: Während 40,8% jener Studierenden mit einer geringen Aufwandswahrnehmung voll und ganz zur Selbstarchivierung bereit sind, trifft dies auf lediglich 26,1% derjenigen mit einer hohen Aufwandswahrnehmung zu. Noch deutlicher fällt die Diskrepanz bei jenen aus, die überhaupt nicht zum Publizieren bereit sind: Während 21,7% der Teilnehmenden mit einer hohen Aufwandswahrnehmung das Selbstarchivieren komplett ablehnen, trifft dies nur auf 4,2% derjenigen mit einer geringen Aufwandswahrnehmung zu.

Die Interpretation der Zeit- und Aufwandswahrnehmung ist vor dem Hintergrund zu betrachten, dass die Studierenden in der Stichprobe den Aufwand einer

Selbstarchivierung überwiegend als gering einschätzen. So umfasst die Gruppe derjenigen mit einer geringen Aufwandswahrnehmung 71 Teilnehmende, während die Gruppe jener Studierenden, die den Aufwand als eher hoch einschätzen, lediglich 23 Teilnehmende umfasst. Gleichwohl die Verzerrung weniger stark ausfällt als beim Faktor des ‚OA-Spirits‘, ist sie doch auffälliger als bei der *prestige*-Variable. Wie der hoch ausgeprägte ‚OA-Spirit‘ könnte auch die gering ausgeprägte Aufwandswahrnehmung damit zusammenhängen, dass für OA sensibilisierte Studierende des Bibliotheks- und Informationsmanagements in der Stichprobe deutlich überrepräsentiert sind. Darüber hinaus könnte der Umstand, dass Studierende weitgehend internetaffin und geübt im Umgang mit Datenmasken sind, zur Verzerrung beitragen, sodass diese auch charakteristisch für die studentische Zielgruppe sein könnte. Insgesamt deckt sich die festgestellte Signifikanz des Kostenfaktors mit Kims Ergebnis ihrer Nachfolgestudie, in der der wahrgenommene Zeit- und Arbeitsaufwand ebenfalls ein signifikanter Prädiktor war (s. Kap. 5.2.2). Auch Pickton und McKnight vermuten implizit, dass dieser Faktor eine Rolle bei der Selbstarchivierungsbereitschaft von Forschungsstudierenden spielen könnte (s. Kap. 5.3.2). Somit ist das Ergebnis dieser Studie ein weiterer Hinweis darauf, dass – wie in H1 vermutet – auch bei Studierenden ein negativer Zusammenhang zwischen dem wahrgenommenen Zeit- und Arbeitsaufwand des Selbstarchivierens und der Selbstarchivierungsbereitschaft in IR bestehen könnte, weshalb es sich lohnt, dieser Hypothese in weiteren Studien nachzugehen.

Die übrigen Prädiktoren, für die gerichtete Zusammenhänge mit der AV angenommen wurden, sind – wie auch der bereits diskutierte Altruismus-Faktor – nicht signifikant. Wie bei der Variable *time_effort* wurde auch bei der Operationalisierung des schulischen Fleißes (*diligence*) ein negativ gepoltes Item eingesetzt, das möglicherweise eine ungewollte Mehrdimensionalität der Skala hervorgerufen haben könnte, die sich im auffällig niedrigen Spearman-Brown-Koeffizienten niederschlägt. Wie in Kapitel 9.1 angedeutet, kommt hinzu, dass beim negativ gepolten Item *diligence2* („Ich lege keinen Wert auf konstruktives Feedback“) eventuell bewusstes Antwortverhalten eine Rolle gespielt haben könnte: Um den Eindruck von Arroganz zu vermeiden, könnten die Teilnehmenden den sozialen Druck verspürt haben, ihre eigene Kritikfähigkeit zu überhöhen. Aufgrund dieses Verdachts wurde das Item aus der intern inkonsistenten Skala entfernt, was jedoch zur Folge hat, dass das Konstrukt nur noch von einem einzigen Item abgebildet wird. Da bereits Skalen mit zwei Items Gefahr laufen, die Konstruktvalidität zu verringern (s. o.), ist bei Verwendung eines einzelnen Items nicht davon auszugehen, dass das Konstrukt adäquat erfasst wurde, was sich wiederum auf den *p*-Wert niedergeschlagen

haben könnte. Aufgrund der in Kims Nachfolgestudie festgestellten Signifikanz des Faktors der akademischen Belohnung, der eine der Bezugsgrundlagen für den Faktor des schulischen Fleißes darstellt (s. Kap. 5.2.2;5.3.1), lohnt es sich, die Operationalisierung zu überarbeiten und im Rahmen einer Wiederholungsstudie erneut zu überprüfen, ob der vermutete positive Einflusseffekt beobachtet werden kann.

Wie bei der Variable *time_effort* sollten auch beim Faktor des Karrierestrebens zwei Items unterschiedliche Formulierungen derselben Frage darstellen. Dass bei beiden Items (*career1* und *career2*) ein Median von 3,0 vorliegt, ist ein erster Hinweis darauf, dass die Operationalisierung dieses Konstruktes insgesamt erfolgreich verlaufen ist. Der zweite Hinweis ist, dass das negativ gepolte Item *career2* die interne Konsistenz nicht auffällig schmälert. Trotzdem lässt sich der in H7a angenommene Effekt in der Realität nicht feststellen. Auch der in H7b vermutete Moderationseffekt hat sich nicht bestätigt. Dieses Ergebnis deckt sich mit den Befunden von Kim (2010;2011), deren Variable der professionellen Anerkennung, die als Referenzgrundlage des *career*-Konstruktes dient, weder in ihrer Nachfolge noch in ihrer Abschlussstudie signifikant ist (s. Kap. 5.2;5.3.1). Die Ergebnisse dieser Studie legen nahe, dass auch bei Studierenden kein positiver Zusammenhang zwischen dem Karrierestreben und der Selbstarchivierungsbereitschaft in IR besteht.

Die Faktoren der Identifikation und der Vertrauenswürdigkeit sind trotz hoher interner Konsistenz nicht signifikant, sodass sich die in H6 und H8 angenommenen Zusammenhänge in der Realität nicht feststellen lassen. Auch in Kims Studien bleiben die Faktoren insignifikant. Das Ergebnis dieser Untersuchung unterstützt die Vermutung, dass auch bei Studierenden kein positiver Zusammenhang zwischen den Prädiktoren der Identifikation und Vertrauenswürdigkeit und der AV der Selbstarchivierungsbereitschaft in IR besteht. Für den Faktor der Plagiatsfurcht, der ebenfalls über eine hohe interne Konsistenz verfügt, ließ sich in dieser Studie auch keine Signifikanz feststellen, die H2 unterstützen würde. Bei Kim floss die Furcht vor Plagiaten als ein Item mit in ihre Vertrauensvariable (*trust*) ein, die in ihrer Abschlussstudie signifikant war, sodass der Faktor für Wissenschaftler:innen eine Rolle spielt (s. Kap. 5.2.3). Es lohnt sich, die Variable der Plagiatsfurcht im Rahmen einer Nachfolgestudie anhand einer fachdisziplinär repräsentativen Stichprobe erneut zu untersuchen, um nachzuprüfen, ob der Faktor auch für die Zielgruppe der Studierenden relevant sein könnte.

Abschließend für die Diskussion lässt sich festhalten, dass die Operationalisierung der latenten Konstrukte insgesamt noch nicht hinreichend war und vor der Durchführung von Nachfolgestudien verbessert werden sollte. Weil die Anzahl der Items pro Skala mit meistens zwei und höchstens drei Items zu niedrig ausfällt, bleibt auch bei jenen Faktoren mit einer hohen internen Konsistenz fraglich, inwieweit die Konstruktvalidität gegeben ist. Um die Reliabilität der Skalen zu erhöhen und vermutete Zusammenhänge besser aufdecken zu können, sollten alle Konstrukte durch weitere Items ergänzt werden (Eisinga et al., 2013, S. 641). Orientiert an einem bei Likert-Skalen häufig gewählten Vorgehen könnte zudem vollständig auf negativ gepolte Items verzichtet werden, um ungewollte Mehrdimensionalität zu vermeiden und eine höhere Korrelation der Items untereinander zu erreichen (Döring, 2023, S. 272). Die Repräsentativität der Stichprobe muss insbesondere in Hinblick auf die fachdisziplinären Hintergründe der Studierenden verbessert werden. Erstrebenswert ist eine umfangreiche Abbildung existierender Selbstarchivierungskulturen. Auch sollte deutschlandweit befragt werden, um regionale Verzerrungen, wie z. B. die OA-Schwerpunktsetzung der HAW Hamburg, zu reduzieren.

11 Handlungsempfehlungen

Auf Grundlage der in dieser Studie identifizierten Gewinn- und Kostenfaktoren werden in diesem Kapitel Handlungsempfehlungen für Hochschulbibliotheken zur Akquise studentischer Autor:innen formuliert. Das Etablieren und aktive Bewerben bibliothekarischer Dienstleistungen leistet einen wichtigen Beitrag zur Incentivierung der Stakeholder, deren Partizipation den Erfolg von IR bedingt (Scherer, 2016, S. 171; s. Kap. 2.1). Die Signifikanz des Faktors der persönlichen Anerkennung zeigt, dass Studierende mit einem hohen Prestigestreben eher mehr dazu bereit sind, als studentische Autor:innen zu IR beizutragen. Um diese Gruppe gezielt anzusprechen, könnte von den Bibliotheken eine turnusmäßige Preisverleihung veranstaltet werden, die exzellente Arbeiten hervorhebt und neu erreichte Meilensteine des IR publikmacht. Parallel zur Präsenzveranstaltung sollte die Preisverleihung auch über die sozialen Medien der Bibliothek bzw. Hochschule in Szene gesetzt werden, um den Bekanntheitsgrad des IR und seiner Inhalte weiter zu steigern. Empfehlenswert ist auch das Einrichten einer eigenständigen Website für das IR, auf der neben Publikationsrichtlinien, Workflows und Kontaktinformationen die ausgezeichneten Arbeiten präsentiert werden können. Eine Verlinkung der IR-

Website auf den Websites der Departments lenkt zusätzliche Aufmerksamkeit auf die IR-Inhalte (Rozum & Thoms, 2016, S. 316; Scherer, 2016, S. 170-171).

Ein weiteres incentivierendes Instrument ist die Integration von Nutzungsstatistiken in die technische Infrastruktur von IR. Nutzungsstatistiken können einerseits als Benchmark der IR-betreibenden Institution fungieren, doch von größerer Relevanz aus Sicht der Autor:innen ist die soziale Wirkkraft ihrer Publikationen, die anhand von Zugriffs-, Ansichts- und Downloadzahlen messbar gemacht wird (Bashir et al., 2022, S. 525). Der Wunsch von Autor:innen nach einer Quantifizierung des Impacts ihrer Arbeit trat sowohl in Kims Pilot- und Abschlussstudie (s. Kap. 5.2.1;5.2.3) als auch in Pickton und McKnights strukturierten Interviews (2006, S. 215-216) zutage. Eine Quantifizierung durch Nutzungsstatistiken und die hierdurch ermöglichte Vergleichbarkeit der persönlichen Wirkkraft könnte auf Studierende mit einem hohen Prestigestreben einen ähnlich bestätigenden Effekt ausüben wie das Sammeln von Likes und Shares in den sozialen Medien und sie dementsprechend zur Produktion von IR-Inhalten motivieren.

Das Streben nach persönlicher Anerkennung ist eine intrinsische Veranlagung, die sich bei Studierenden mit einem gering ausgeprägten Prestigestreben nicht künstlich hervorrufen lässt. Eine Möglichkeit, diese Gruppe trotzdem zu erreichen, besteht darin, an den ebenfalls signifikanten Faktor des ‚OA-Spirits‘ zu appellieren und das IR in eine breit angelegte Strategie zur Förderung einer Philosophie der Offenheit an der Hochschule einzubetten. Eine Maßnahme zur Förderung des OA-Bewusstseins ist die Ernennung von „liaison librarians“ (Scherer, 2016, S. 161) bzw. Verbindungsbibliothekar:innen, die in engem Kontakt zu den jeweiligen Departments stehen. Das ausgewählte Bibliothekspersonal sollte dabei selbst über einen hohen ‚OA-Spirit‘ verfügen, um diesen authentisch nach außen hin kommunizieren zu können. Weil unterschiedliche Selbstarchivierungskulturen eine fachlich zugeschnittene Herangehensweise erfordern, müssen die Verbindungsbibliothekar:innen darüber hinaus ein präzises Verständnis für die Bedürfnisse, Erwartungen und Sorgen der von ihnen betreuten Fachkulturen mitbringen, um erfolgreich zwischen den Departments und der Bibliothek vermitteln zu können. Die primäre Aufgabe der Verbindungsbibliothekar:innen besteht darin, die Gewinnfaktoren einer Selbstarchivierung zu kommunizieren und bei der Überwindung der Kostenfaktoren zu unterstützen (Pickton & McKnight, 2006, S. 217; Rozum & Thoms, 2016, S. 318; Scherer, 2016, S. 161; Shearer, 2003, S. 257).

Pickton und McKnight stellen fest, dass Studierende dem Rat zur Selbstarchivierung eher folgen, wenn dieser anstatt des Bibliothekspersonals von den Lehrenden erteilt wird. Daher ist zu empfehlen, dass die Verbindungsbibliothekar:innen zunächst Kontakt zu den Lehrenden aufnehmen, welche dann wiederum gezielt Studierende aus ihren Veranstaltungen oder Betreuungsprojekten rekrutieren können (2006, S. 214;217). Lehrende mit einem hohen ‚OA-Spirit‘, die selbst über Erfahrung im Umgang mit IR verfügen, können eine Vorbildfunktion einnehmen und stellen damit ein wichtiges Bindeglied zwischen potenziellen studentischen Autor:innen und den Hochschulbibliotheken dar (Davis-Kahl, 2012, S. 215; 2021, S. 93). Darüber hinaus können die Lehrenden dazu beitragen, der Gefahr einer Kompromittierung der Online-Identität der Studierenden entgegenzuwirken, indem sie sich an der Formulierung der Auswahlkriterien beteiligen, die studentische Arbeiten erfüllen müssen, um im IR publiziert zu werden (Davis-Kahl, 2021, S. 88). In Zusammenhang mit der Qualitätssicherung ist ebenfalls zu empfehlen, bei allen Publikationen standardmäßig die begutachtende Person in den Metadaten zu erfassen und ein nachträgliches Löschen studentischer Beiträge zu ermöglichen (s. Kap. 2.2.2).

Um den signifikanten Kostenfaktor des zusätzlichen Zeit- und Arbeitsaufwands abzuschwächen, sollten die Bibliotheken einen Upload-Service anbieten, der den Studierenden das Hochladen ihrer Arbeiten inklusive des Ausfüllens der Metadaten sowie der Verschlagwortung abnimmt (Kim, 2010, S. 1909; Pickton & McKnight, 2006, S. 216; Rozum & Thoms, 2016, S. 317; Scherer, 2016, S. 167; s. Kap. 5.3.2). Gleichzeitig sollten die Kompetenzen der Studierenden dahingehend gestärkt werden, dass sie auch selbstständig zum Publizieren ihrer Arbeiten fähig sind. Hierfür sollten die Bibliotheken Veranstaltungsformate konzipieren, die neben einer technischen Schulung im Umgang mit dem IR auch weitergefasste Aspekte des freien Publizierens behandeln, wie z. B. Zitation und Quellenfindung oder Copyright und Lizenzen. Die Schulungsmaterialien sollten frei auf der Website des Repositoriums zur Verfügung stehen (Passehl-Stoddart & Monge, 2014, S. 9; Pickton & McKnight, 2006, S. 217).

Um die Teilnahmequote vonseiten der Studierenden zu erhöhen und den Mehraufwand für die Lehrenden, die im Austausch mit den Verbindungsbibliothekar:innen stehen, zu reduzieren, ist eine Integration derartiger Schulungsformate in das feste Curriculum zu begrüßen. Eine Best Practice stellt die HAW Hamburg dar, die für Bachelorstudierende des Bibliotheks- und Informationsmanagements im zweiten Semester die Profilverkstatt Metadaten anbietet, in deren Rahmen die

Studierenden Repositorien kennenlernen und erproben. Auch die Drittsemester des Masterstudiengangs Digitale Transformation der Informations- und Medienwirtschaft praktizieren im Rahmen des Gestaltungsmoduls die Anwendung offener Publikationsstandards (HAW Hamburg, 2021, S. 40-41; 2024, S. 11). Wenn die Überrepräsentation von HAW-Studierenden des Bibliotheks- und Informationsmanagements wirklich, wie in Kapitel 9.4 vermutet, die Stichprobe hinsichtlich ihrer Selbstarchivierungsbereitschaft verzerrt haben sollte, wäre dies ein Hinweis darauf, dass die Integration von OA-Themen ins Curriculum auch für andere Studiengänge und Hochschulen eine sinnvolle Maßnahme zur Förderung der Selbstarchivierungsbereitschaft von Studierenden in IR sein kann. Bei Studiengängen ohne informationswissenschaftlichen Bezug könnten die Verbindungsbibliothekar:innen den Studierenden das IR und seine Möglichkeiten im Rahmen von Veranstaltungen zum wissenschaftlichen Schreiben nahebringen.

12 Desiderata

Die in Kapitel 6 aufgestellten Hypothesen lassen sich anhand der in dieser Arbeit durchgeführten empirischen Studie nicht abschließend belegen oder verwerfen. Wie in Kapitel 10 erläutert, stellen die Ergebnisse lediglich erste Hinweise auf die Plausibilität der Hypothesen dar, sodass Nachfolgestudien mit einer verbesserten Operationalisierung, welche die Hypothesen anhand einer repräsentativen Stichprobe genauer untersuchen, erforderlich sind. Darüber hinaus kamen bei der Diskussion der Ergebnisse neue Fragen auf, denen ebenfalls in weiteren Studien nachgegangen werden sollte. So wurde eine mögliche Gruppe von Studierenden identifiziert, die zwar persönlich weitgehend zur Selbstarchivierung bereit sind, dabei jedoch die Idee von Forschungsergebnissen als öffentliches Gut ablehnen. Die geringe Größe der Gruppe ($n=2$) lässt keine sinnvolle Interpretation dieser Beobachtung zu, doch aus den identifizierten signifikanten Gewinnfaktoren lässt sich die Vermutung ableiten, dass sich die Selbstarchivierungsbereitschaft der Gruppe – sofern es sich bestätigt, dass sie überhaupt existiert – möglicherweise durch ein hohes Prestigestreben erklären ließe. Es lohnt sich, dieser neuen Hypothese im Rahmen einer repräsentativen Untersuchung nachzugehen.

Eine weitere Frage, die in der Diskussion offenblieb, betrifft mögliche von den Teilnehmenden wahrgenommene, aber nicht explizit abgefragte Hürden im Selbstarchivierungsprozess, die zur Signifikanz des wahrgenommenen Zeit- und Arbeitsaufwands beigetragen haben könnten. Eine geeignete Methode zur nähe-

ren Erforschung dieser möglichen Publikationshürden stellt die Durchführung von Usability-Studien anhand eines programmierten Website-Prototypen dar, in denen konkrete Verständnisschwierigkeiten aufgedeckt werden können. Ergänzend hierzu bietet sich das Durchführen qualitativer Interviews an, in denen die Studierenden zu möglichen Bedenken befragt werden, die sie über den konkreten Upload-Prozess hinaus von einer Selbstarchivierung abhalten könnten.

Neben den in dieser Studie untersuchten Hypothesen lassen sich weitere potenzielle Untersuchungsgegenstände identifizieren, die zu einem besseren Verständnis der Selbstarchivierungsbereitschaft von Studierenden in IR beitragen. Hierzu gehört einerseits die Rolle von Lizenzen, die in dieser Arbeit durch die Beschränkung auf das Gratis-Gold-Modell weitgehend ausgeklammert wurde (s. Kap. 4). Es könnte näher untersucht werden, inwieweit sich die Selbstarchivierungsbereitschaft von Studierenden in Abhängigkeit von verschiedenen Lizenzmodellen, wie z. B. dem Gratis- oder Libre-Publizieren (s. Kap. 1), verändert. Eine weitere Thematik, die es sich zu erforschen lohnt, sind fachdisziplinäre Unterschiede zwischen den Selbstarchivierungskulturen, wie Kim sie für die Zielgruppe des wissenschaftlichen Personals festgestellt hat (2010, S. 1916-1917; s. Kap. 5.2.2). Zu erforschen wäre, inwieweit sich die für eine jeweilige Fachdisziplin charakteristische Einstellung zum freien Publizieren bereits bei den Studierenden niederschlägt. Orientiert an Kim ließe sich anhand einer repräsentativen Stichprobe eine einfaktorielle Varianzanalyse (ANOVA) durchführen, um mögliche fachdisziplinäre Unterschiede aufzudecken.

Auch Methoden der ethnografischen Feldforschung bergen die Chance, nähere Informationen über verschiedene Fachkulturen aus studentischer Perspektive zu gewinnen, wobei insbesondere die Beschaffenheit der konkreten Arbeitsergebnisse von Interesse ist, die bei Fächern wie Textildesign, Maschinenbau oder Philosophie sehr unterschiedlich ausfallen können (s. Kap. 2.2.1). Um hierüber konkrete Erkenntnisse zu gewinnen, bietet sich die Durchführung von *Photo Elicitation Interviews* an. Hierbei werden die Studierenden in die Rolle von Expert:innen versetzt und aktiv am Forschungsprozess beteiligt, indem sie zunächst eine Fotostrecke anhand vorgegebener Themen bzw. Prompts erstellen, welche anschließend im Rahmen eines qualitativen Interviews gemeinsam betrachtet und besprochen wird (Briden, 2007, S. 40-42). Auf Grundlage des Datenmaterials lassen sich Handlungsempfehlungen zur sinnvollen Abbildung vielfältiger Studienleistungen in IR formulieren sowie möglicherweise auch bisher unbekannte fachspezifische Gewinn- oder

Kostenfaktoren ableiten, die eventuell in Abhängigkeit des jeweiligen Dokumenttyps auftreten könnten.

13 Fazit

Mit dem Ziel, Handlungsempfehlungen für Hochschulbibliotheken zur Akquise studentischer Autor:innen aufzustellen, wurde in dieser Arbeit der Frage nachgegangen, welche Faktoren die Bereitschaft von Studierenden zur Selbstarchivierung in IR beeinflussen. Der hierfür durchgeführten empirischen Studie liegt die Hypothese zugrunde, dass sich anhand von Gewinn-, Kosten- und Kontextfaktoren die Selbstarchivierungsbereitschaft von Studierenden erklären lässt. Die Ergebnisse liefern Anhaltspunkte, die diese Hypothese im Fall der Gewinn- und Kostenfaktoren unterstützen: Im Rahmen einer linearen Regressionsanalyse wurden mit dem Prestigestreben und dem ‚OA-Spirit‘ zwei signifikante Gewinnfaktoren sowie mit dem wahrgenommenen Zeit- und Arbeitsaufwand ein signifikanter Kostenfaktor identifiziert, die Einfluss auf die Publikationsbereitschaft nehmen. Vor 15 Jahren prognostizierte die Informationswissenschaftlerin Kim, deren Forschung die Grundlage des in dieser Arbeit aufgestellten empirischen Modells bildet, dass das Selbstarchivieren im Zuge des voranschreitenden Generationenwechsels nach und nach zu einer gängigen Praxis werden würde (2010, S. 1917;1920). Unter Vorbehalt aller diskutierten Limitationen der Stichprobe ist die festgestellte hohe Selbstarchivierungsbereitschaft unter den Teilnehmenden ein erster Hinweis darauf, dass sich Kims Prognose bewahrheiten und die nachrückende Generation eine neue Aufgeschlossenheit gegenüber freiem Publizieren an die Hochschulen tragen könnte.

Gleichzeitig darf nicht auf die Eigeninitiative motivierter Individuen gebaut werden, um IR nachhaltig in der intellektuellen Infrastruktur der Hochschulen zu verankern. Individuelles Verhalten wird sich erst dann ändern, wenn von den Institutionen attraktive Rahmenbedingungen geschaffen werden, die eine Entscheidung zur Selbstarchivierung begünstigen. Dies beinhaltet sowohl Incentivierungen, die das Selbstarchivieren belohnen, als auch Unterstützung beim Überwinden von Publikationshürden (Borgman, 2007, S. 225). Das Veranlassen von Preisverleihungen und Schulungsformaten, die Ernennung von Verbindungsbibliothekar:innen sowie das Angebot eines Upload-Services stellen erste Handlungsempfehlungen für Hochschulbibliotheken zur Erhöhung der Partizipationsrate von Studierenden dar, die nach der Implementierung von IR lange Zeit als potenzielle Inhaltskonsu-

ment:innen und -produzent:innen vernachlässigt wurden. Das gezielte Werben um studentische Autor:innen ist ein wichtiger Bestandteil des strukturellen Wandels, den Hochschulbibliotheken im Rahmen des digitalen Transformationsprozesses wissenschaftlicher Kommunikations- und Forschungspraktiken vollziehen.

14 Literaturverzeichnis

- Abrizah, A., Hilmi, M., & Kassim, N. A. (2015). Resource-sharing through an inter-institutional repository: Motivations and resistance of library and information science scholars. *The Electronic Library*, 33(4), 730-748. <https://doi.org/10.1108/EL-02-2014-0040>
- Alavi, M., & Leidner, D. E. (2001). Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS Quarterly*, 25(1), 107-136. <https://doi.org/10.2307/3250961>
- Alby, T. (2022). *Data Science in der Praxis: Eine verständliche Einführung in alle wichtigen Verfahren*. Rheinwerk Verlag.
- Ammarukleart, S., & Kim, J. (2017). Institutional repository research 2005-2015: A trend analysis using bibliometrics and text mining. *Digital Library Perspectives*, 33(3), 264-278. <https://doi.org/10.1108/DLP-07-2016-0027>
- Arzheimer, K. (2016). *Strukturgleichungsmodelle: Eine anwendungsorientierte Einführung*. Springer VS. <https://doi.org/10.1007/978-3-658-09609-0>
- Ball, D. (2016). Open access: Effects on publishing behaviour of scientists, peer review and interrelations with performance measures. In P. Weingart & N. Taubert (Hrsg.), *Wissenschaftliches Publizieren: Zwischen Digitalisierung, Leistungsmessung, Ökonomisierung und medialer Beobachtung* (S. 179-209). De Gruyter Akademie Forschung.
- Bankier, J.-G., & Smith, C. (2010). Repository collection policies: Is a liberal and inclusive policy helpful or harmful? *Australian Academic & Research Libraries*, 41(4), 245-259. <https://doi.org/10.1080/00048623.2010.10721479>
- Barandiaran, D., Rozum, B., & Thoms, B. (2014). Focusing on student research in the institutional repository: DigitalCommons@USU. *College & Research Libraries News*, 75(10), 546-549. <http://dx.doi.org/10.5860/crln.75.10.9209>

- Bashir, S., Gul, S., Bashir, S., Nisa, N. T., & Ganaie, S. A. (2022). Evolution of institutional repositories: Managing institutional research output to remove the gap of academic elitism. *Journal of Librarianship and Information Science*, 54(3), 518-531. <https://doi.org/10.1177/09610006211009592>
- Bergstrom, T. C. (2001). Free labor for costly journals? *The Journal of Economic Perspectives*, 15(3), 183-198. <http://econ.ucsb.edu/~tedb/archive/freelabor.pdf>
- Borgman, C. L. (2007). *Scholarship in the digital age: Information, infrastructure, and the internet*. MIT Press.
- Briden, J. (2007). Photo surveys: Eliciting more than you knew to ask for. In N. F. Foster & S. Gibbons (Hrsg.), *Studying students: The undergraduate research project at the University of Rochester* (S. 40-47). Association of College and Research Libraries. <http://hdl.handle.net/1802/7520>
- Brown, P. O., Cabell, D., Chakravarti, A., Cohen, B., Delamothe, T., Eisen, M., Grivell, L., Guédon, J.-C., Hawley, R. S., Johnson, R. K., Kirschner, M. W., Lipman, D., Lutzker, A. P., Marincola, E., Roberts, R. J., Rubin, G. M., Schloegl, R., Siegel, V., So, A. D., Suber, P., Varmus, H. E., Velterop, J., Walport, M. J., & Watson, L. (2003, 20. Juni). *Bethesda statement on open access publishing*. <http://nrs.harvard.edu/urn-3:HUL.InstRepos:4725199>
- Buckland, M. K. (1991). Information as thing. *Journal of the American Society for Information Science* 42(5), 351-360.
- Budapest Open Access Initiative (BOAI). (2002, 14. Februar). *Budapest Open Access Initiative*. <https://www.budapestopenaccessinitiative.org/read/>
- Callicott, B. B., Scherer, D., & Wesolek, A. (2016). Introduction. In B. B. Callicott, D. Scherer & A. Wesolek (Hrsg.), *Making institutional repositories work* (xv-xxvi). Purdue University Press. <https://www.jstor.org/stable/j.ctt1wf4drg.4>
- Caprio, M. J. (2014). Student publishing: Future scholars as change agents, *OCLC Systems & Services: International digital library perspectives*, 30(3), 144-157. <https://doi.org/10.1108/OCLC-01-2014-0003>

- Carroll, M. W. (2011). Why full open access matters. *PLoS Biology* 9(11), e1001210. <https://doi.org/10.1371/journal.pbio.1001210>
- Crow, R. (2002). The case for institutional repositories: A SPARC position paper. *Scholarly Publishing*, 223, 1-37.
- Davis-Kahl, S. (2012). Engaging undergraduates in scholarly communication: Outreach, education, and advocacy. *College & Research Libraries News*, 73(4), 212-222. <https://doi.org/10.5860/crln.73.4.8744>
- Davis-Kahl, S. (2021). Undergraduate student work in the institutional repository. In S. C. Finlay (Hrsg.), *The complete guide to institutional repositories* (S. 87-93). ALA Editions.
- Demetres, M. R., Delgado, D., Wright, D. N. (2020). The impact of institutional repositories: A systematic review. *Journal of the Medical Library Association*, 108(2), 177-184. <https://doi.org/10.5195/jmla.2020.856>
- Deppe, A., & Beucke, D. (2017). Ursprünge und Entwicklung von Open Access. In K. Söllner & B. Mittermaier (Hrsg.), *Praxishandbuch Open Access* (S. 12-20). De Gruyter Saur. <https://doi.org/10.1515/9783110494068-002>
- Deutsche Forschungsgemeinschaft (DFG). (2022, 22. Juni). *Was ist Open Access?* Deutsche Forschungsgemeinschaft. https://www.dfg.de/foerderung/programme/infrastruktur/lis/open_access/was_ist_open_access/index.html
- Dewenter, R., & Rösch, J. (2015). *Einführung in die neue Ökonomie der Medienmärkte: Eine wettbewerbsökonomische Betrachtung aus Sicht der Theorie der zweiseitigen Märkte*. Springer Gabler. <https://doi.org/10.1007/978-3-658-04736-8>
- Dijk, J. van. (2020). *The digital divide*. Polity.
- Döring, N. (2023). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (6. Aufl.). Springer. <https://doi.org/10.1007/978-3-662-64762-2>

- Eckstein, P. P. (2016). *Angewandte Statistik mit SPSS: Praktische Einführung für Wirtschaftswissenschaftler* (8. Aufl.). Gabler Verlag. <http://dx.doi.org/10.1007/978-3-658-10918-9>
- Eisinga, R., Grotenhuis, M. t., & Pelzer, B. (2013). The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown? *International Journal of Public Health*, 58, 637-642. <https://doi.org/10.1007/s00038-012-0416-3>
- Exline, E. (2016). Extending the institutional repository to include undergraduate research. *College & Undergraduate Libraries*, 23(1), 16-27. <https://doi.org/10.1080/10691316.2014.950782>
- Fox, J., & Monette, G. (1992). Generalized collinearity diagnostics. *Journal of the American Statistical Association*, 87(417), 178-183. <https://doi.org/10.2307/2290467>
- Fox, J., & Weisberg, S. (2019). *An R companion to applied regression* (3. Aufl.). Sage.
- Gantert, K. (2016). *Bibliothekarisches Grundwissen*. De Gruyter Saur. <https://doi.org/10.1515/9783110321500>
- Gilbert, S. F. (2004). Points of view: Should students be encouraged to publish their research in student-run publications? A case against undergraduate-only journal publications. *Cell Biology Education*, 3(1), 22-23. <https://doi.org/10.1187/cbe.04-01-0023>
- Guo, H., Wang, C., Su, Z., & Wang, D. (2020). Technology push or market pull? Strategic orientation in business model design and digital start-up performance. *Journal of Product Innovation Management*, 37(4), 352-372. <https://doi.org/10.1111/jpim.12526>
- Harris, E., & Issakson, P. (2019, 20. Oktober). *How to use open access repositories* [Infografik]. Zenodo. <https://doi.org/10.5281/zenodo.3666627>
- HAW Hamburg. (o. J.a). *SAIL HAW Hamburg: Studentisches Publikationsportal: Austausch – Information – Lernen*. Reposit. <https://reposit.haw-hamburg.de/sail>

- HAW Hamburg. (o. J.b). *REPOSIT - das institutionelle Repositorium der HAW Hamburg*.
Reposit. <https://reposit.haw-hamburg.de>
- HAW Hamburg. (2021). *Modulhandbuch des Bachelor-Studiengangs Bibliotheks- und Informationsmanagement*. HAW Hamburg. <https://www.haw-hamburg.de/studium/studienorganisation/ordnungen/modulhandbuecher/>
- HAW Hamburg. (2024). *Modulhandbuch Digitale Transformation der Informations- und Medienwirtschaft (DiTra)*. HAW Hamburg. <https://www.haw-hamburg.de/studium/studienorganisation/ordnungen/modulhandbuecher/>
- Hayes, A. F. (2022). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Press.
- Hayes, A. F., & Cai, L. (2007). Using heteroskedasticity-consistent standard error estimators in OLS regression: An introduction and software implementation. *Behavior Research Methods* 39(4), 709-722. <https://doi.org/10.3758/BF03192961>
- Heinrich, E. (2004). Electronic repositories of marked student work and their contributions to formative evaluation. *Journal of Educational Technology & Society*, 7(3), 82-96. <http://www.jstor.org/stable/jeductechsoci.7.3.82>
- Herb, U. (2016). Empfehlungen, Stellungnahmen, Deklarationen und Aktivitäten wissenschaftspolitischer Akteure zur Gestaltung des wissenschaftlichen Kommunikationssystems. In P. Weingart & N. Taubert (Hrsg.), *Wissenschaftliches Publizieren: Zwischen Digitalisierung, Leistungsmessung, Ökonomisierung und medialer Beobachtung* (S. 147-177). De Gruyter Akademie Forschung.
- Hertenstein, E. (2014). Student scholarship in institutional repositories. *Journal of Librarianship and Scholarly Communication*, 2(3), eP1135. <http://dx.doi.org/10.7710/2162-3309.1135>
- Hüther, M. (2017). Digitalisierung: Motor im Strukturwandel: Herausforderung für die Wirtschaftspolitik. *ORDO: Jahrbuch für die Ordnung von Wirtschaft und Gesellschaft*, 68(1), 179-216. <https://doi.org/10.1515/ordo-2018-0009>

- Kim, J. (2007). Motivating and impeding factors affecting faculty contribution to institutional repositories. *Journal of Digital Information*, 8(2).
<https://ils.unc.edu/tibbo/JCDL2006/Kim-JCDLWorkshop2006.pdf>
- Kim, J. (2010). Faculty self-archiving: Motivations and barriers. *Journal of the American Society for Information Science and Technology*, 61(9), 1909-1922.
<https://doi.org/10.1002/asi.21336>
- Kim, J. (2011). Motivations of faculty self-archiving in institutional repositories. *The Journal of Academic Librarianship*, 37(3), 246-254.
<https://doi.org/10.1016/j.acalib.2011.02.017>
- Kincaid, H. (2012). Mechanisms, causal modeling, and the limitations of traditional multiple regression. In H. Kincaid (Hrsg.), *The Oxford Handbook of Philosophy of Social Science* (S. 46-64). Oxford University Press.
- Kling, R., McKim, G., & King, A. (2003). A bit more to it: Scholarly communication forums as socio-technical interaction networks. *Journal of the American Society for Information Science and Technology*, 54(1), 47-67.
<https://doi.org/10.1002/asi.10154>
- Kronthaler, F. (2021). *Statistik angewandt mit Excel: Datenanalyse ist (k)eine Kunst* (2. Aufl.). Springer Spektrum. <https://doi.org/10.1007/978-3-662-62302-2>
- Kronthaler, F., & Zöllner, S. (2021). *Data analysis with RStudio: An easygoing introduction*. Springer Spektrum. <https://doi.org/10.1007/978-3-662-62518-7>
- Levin, J. A., Burbules, N. C., & Bruce, B. C. (2005). From student work to exemplary educational resources: The case of the CTER White Papers. *E-Learning and Digital Media*, 2(1), 39-49. <https://www.researchgate.net/publication/45532153>
- Lynch, C. A. (2003). Institutional repositories: Essential infrastructure for scholarship in the digital age. *Portal: Libraries and the Academy*, 3(2), 327-336.
<https://doi.org/10.1353/pla.2003.0039>

- Lynch, C. A. (2016). Foreword: A few reflections on the evolution of institutional repositories. In B. B. Callicott, D. Scherer & A. Wesolek (Hrsg.), *Making institutional repositories work* (xi–xiii). Purdue University Press.
<https://www.jstor.org/stable/j.ctt1wf4drg.3>
- Madalli, D. P. (2015). *Concepts of openness and open access*. United Nations Educational, Scientific and Cultural Organization (UNESCO).
<https://unesdoc.unesco.org/ark:/48223/pf0000232207>
- Marcoulides, K. M., & Raykov, T. (2019). Evaluation of variance inflation factors in regression models using latent variable modeling methods. *Educational and Psychological Measurement*, 79(5), 874-882.
<https://doi.org/10.1177/0013164418817803>
- Markey, K., St. Jean, B., Soo, Y. R., Yakel, E., & Kim, J. (2008). Institutional repositories: The experience of master's and baccalaureate institutions. *Portal: Libraries and the Academy* 8(2), 157-173. <https://doi.org/10.1353/pla.2008.0022>
- Max-Planck-Gesellschaft. (o. J.). *Berlin declaration on open access to knowledge in the sciences and humanities*. Open Access: Initiativen der Max-Planck-Gesellschaft.
<https://openaccess.mpg.de/Berliner-Erklaerung>
- Meyer, E. T., & Schroeder, R. (2015). *Knowledge machines: Digital transformations of the sciences and humanities*. The MIT Press.
<https://ieeexplore.ieee.org/book/7109386>
- Miedema, F. (2022). *Open science: The very idea*. Springer Nature.
<https://doi.org/10.1007/978-94-024-2115-6>
- O'Gara, G., & Drake Davis, L. (2017). Sharing the spotlight: Open access publishing and undergraduate research. In K. L. Smith & K. A. Dickson (Hrsg.), *Open access and the future of scholarly communication implementation* (S.143-168). Rowman & Littlefield.
- Öztürk, M. H., & Kудay, A. D. (2024). Psychometric evaluation of the Turkish version of the Work-Related Stress Scale: A study among search and rescue workers responding to the 2023 Kahramanmaraş earthquakes. *Journal of Health Psychology*, 2024, 1-12. <https://doi.org/10.1177/13591053241287675>

- Passehl-Stoddart, E., & Monge, R. (2014). From freshman to graduate: Making the case for student-centric institutional repositories. *Journal of Librarianship and Scholarly Communication*, 2(3), eP1130. <http://dx.doi.org/10.7710/2162-3309.1130>
- Pearl, J. (2000). *Causality: Models, reasoning, and inference*. Cambridge University Press.
- Pickton, M., & McKnight, C. (2006). Research students and the Loughborough institutional repository. *Journal of Librarianship and Information Science*, 38(4), 203-219. <https://doi.org/10.1177/0961000606066572>
- Picot, A., & Fischer, T. (2006). *Weblogs professionell: Grundlagen, Konzepte und Praxis im unternehmerischen Umfeld*. Dpunkt.Verlag.
- Plutchak, T. S., & Moore, K. B. (2017). Dialectic: The aims of institutional repositories. *The Serials Librarian*, 72(1-4), 27-35. <https://doi.org/10.1080/0361526X.2017.1320868>
- Porst, R. (2014). *Fragebogen: Ein Arbeitsbuch* (4. Aufl.). Springer VS. <https://doi.org/10.1007/978-3-658-02118-4>
- Rozum, B., & Thoms, B. (2016). Populating your institutional repository and promoting your students: IRs and undergraduate research. In B. B. Callicott, D. Scherer & A. Wesolek (Hrsg.), *Making institutional repositories work* (S. 311–318). Purdue University Press. <https://doi.org/10.2307/j.ctt1wf4drg.28>
- Rozum, B., Thoms, B., Bates, S., & Barandiaran, D. M. (2015, 25.-28. März). *We have only scratched the surface: The role of student research in institutional repositories* [Konferenzschrift]. ACRL 17th National Conference, Portland, OR, USA. <http://hdl.handle.net/11213/17921>
- Salganik, M. J. (2018). *Bit by bit: Social research in the digital age*. Princeton University Press.

- Sathe, N., Grady, J., & Giuse, N. (2002). Print versus electronic journals: A preliminary investigation into the effect of journal format on research processes. *Journal of the Medical Library Association*, 90(2), 235-243.
<https://www.researchgate.net/publication/11371986>
- Schallehn, V., & Schimmer, R. (2014). Open Access. In R. Griebel, H. Schäffler & K. Söllner (Hrsg.), *Praxishandbuch Bibliotheksmanagement* (S. 311-338). De Gruyter Saur. <https://doi.org/10.1515/9783110303261>
- Scherer, D. (2016). Incentivizing them to come: Strategies, tools, and opportunities for marketing an institutional repository. In B. B. Callicott, D. Scherer & A. We-solek (Hrsg.), *Making institutional repositories work* (S. 159-174). Purdue University Press. <https://www.jstor.org/stable/j.ctt1wf4drg.17>
- Schimmer, R., & Geschuhn, K. (2017). Open-Access-Transformation: Die Ablösung des Subskriptionswesens durch Open-Access-Geschäftsmodelle. In K. Söllner & B. Mittermaier (Hrsg.), *Praxishandbuch Open Access* (S. 173-180).
<https://doi.org/10.1515/9783110494068-020>
- Shearer, K. (2003). Institutional repositories: Towards the identification of critical success factors. *Proceedings of the Annual Conference of CAIS*, 250-263.
<https://doi.org/10.29173/cais532>
- Smith, E. T. (2003). Changes in faculty reading behaviors: The impact of electronic journals on the University of Georgia, *The Journal of Academic Librarianship*, 29(3), 162-168. [https://doi.org/10.1016/S0099-1333\(03\)00018-1](https://doi.org/10.1016/S0099-1333(03)00018-1).
- Söllner, K. (2017). Warum und für wen Open Access? In K. Söllner & B. Mittermaier (Hrsg.), *Praxishandbuch Open Access* (S. 3-11).
<https://doi.org/10.1515/9783110494068-001>
- Statistisches Bundesamt. (2024, 16. September). *Statistischer Bericht: Statistik der Prüfungen: Prüfungsjahr 2023* (Tab. 21321-01). Destatis.
<https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bildung-Forschung-Kultur/Hochschulen/Publikationen/Downloads-Hochschulen/statistischer-bericht-pruefungen-2110420237005.html>

- SUB Göttingen. (2022, 2. November). *Repositorien*. Open-Access-Network.
<https://open-access.network/informieren/publizieren/repositorien>
- Swan, A., Willmers, M., & King, T. (2014). Costs and benefits of open access: A guide for managers in Southern African higher education. *Scholarly Communication in Africa Programme*, 2. <http://hdl.handle.net/11427/2324>
- Taubert, N., & Weingart, P. (2016). Wandel des wissenschaftlichen Publizierens: Eine Heuristik zur Analyse rezenter Wandlungsprozesse. In P. Weingart & N. Taubert (Hrsg.), *Wissenschaftliches Publizieren: Zwischen Digitalisierung, Leistungsmessung, Ökonomisierung und medialer Beobachtung* (S. 3-38). De Gruyter Akademie Forschung.
- Ukwoma, S. C., & Dike, V. W. (2017). Academics' attitudes toward the utilization of institutional repositories in Nigerian universities. *Portal: Libraries and the Academy*, 17(1), 17-32. <https://doi.org/10.1353/pla.2017.0002>.
- Urban, D., & Mayerl, J. (2018). *Angewandte Regressionsanalyse: Theorie, Technik und Praxis* (5. Aufl.). Springer VS. <https://doi.org/10.1007/978-3-658-01915-0>
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *Management Information Systems (MIS) Quarterly*, 27(3), 425-478. <https://doi.org/10.2307/30036540>
- Vet, H. C. W. de, Mokkink, L. B., Mosmuller, D. G., & Terwee, C. B. (2017). Spearman-Brown prophecy formula and Cronbach's alpha: Different faces of reliability and opportunities for new applications. *Journal of Clinical Epidemiology*, 85, 45-49.
<http://dx.doi.org/10.1016/j.jclinepi.2017.01.013>
- Walkington, H. (2012). Developing dialogic learning space: The case of online undergraduate research journals. *Journal of Geography in Higher Education*, 36(4), 547-562. <https://doi.org/10.1080/03098265.2012.692072>
- Ware, M. (2004). Institutional repositories and scholarly publishing. *Learned Publishing*, 17(2), 115-124. <https://doi.org/10.1087/095315104322958490>
- Watkinson, C. (2014). The university as publisher revisited. *Insights*, 27(2), 181-185.
<http://dx.doi.org/10.1629/2048-7754.131>

Wooldridge, J. M. (2020). *Introductory econometrics: A modern approach* (7. Aufl.). Cengage.

Zeller, M., & Stenberg, E. S. (2017). Faculty require online distribution of student work: Enter the librarian. In K. L. Smith & K. A. Dickson (Hrsg.), *Open access and the future of scholarly communication implementation* (S. 31-52). Rowman & Littlefield.

Anhang A: Fragebogen



Hochschule für Angewandte
Wissenschaften Hamburg
Hamburg University of Applied Sciences

Umfrage zum Veröffentlichen von Studienarbeiten

In diesem Masterprojekt an der HAW Hamburg geht es um **eure Einstellung zum Veröffentlichen von Arbeiten**, die ihr **im Rahmen eures Studiums erstellt** habt (z. B. Abschluss-/Hausarbeiten, Folien, Projekt-/Laborberichte, Essays oder Code). Die Umfrage richtet sich an alle Studierenden und fragt nach euren ganz persönlichen Einschätzungen, sodass es keine richtigen oder falschen Antworten gibt. Für die Auswertung ist es wichtig, dass ihr bitte keine Frage auslasst.

Alle Angaben erfolgen **anonym** und dienen rein wissenschaftlichen Zwecken. Das Ausfüllen des Fragebogens mit **20 Ankreuzfragen** dauert **ca. 7 Minuten**. Unter allen Teilnehmenden verlose ich **3x15€** als **Amazon-Gutscheine**.

Vielen Dank für eure Unterstützung!

Anmerkungen und Fragen gerne an: Silja Wiegmann (silja.wiegmann@haw-hamburg.de),
MA Digitale Transformation

[In Google anmelden](#), um den Fortschritt zu speichern. [Weitere Informationen](#)

Bitte gib an, inwieweit du den folgenden Aussagen zustimmst:
Überhaupt nicht, weniger, weitgehend oder voll und ganz?

Ich möchte zeigen, dass ich erfolgreich bin.

1 2 3 4

stimme überhaupt nicht zu stimme voll und ganz zu

Ich mag es nicht, wenn mein Name im Internet auftaucht.

1 2 3 4

stimme überhaupt nicht zu stimme voll und ganz zu

Gute Leistungen im Studium erhöhen meine Chancen auf dem Arbeitsmarkt.

1 2 3 4

stimme überhaupt nicht zu stimme voll und ganz zu

Meine Noten haben keinen Einfluss auf meine späteren Berufsaussichten.

1 2 3 4

stimme überhaupt nicht zu stimme voll und ganz zu

Ich möchte die Qualität meiner Studienleistungen ständig verbessern.

1 2 3 4

stimme überhaupt nicht zu stimme voll und ganz zu

Ich lege keinen Wert auf konstruktives Feedback.

1 2 3 4

stimme überhaupt nicht zu stimme voll und ganz zu

Weiter



Seite 1 von 5

[Alle Eingaben löschen](#)

Ich fühle mich meiner Hochschule zugehörig.

1 2 3 4

stimme überhaupt nicht zu stimme voll und ganz zu

Ich bin stolz darauf, an meiner Hochschule zu studieren.

1 2 3 4

stimme überhaupt nicht zu stimme voll und ganz zu

Inhalte, die meine Hochschule veröffentlicht, sind von guter Qualität.

1 2 3 4

stimme überhaupt nicht zu stimme voll und ganz zu

Inhalte, die von meinen Lehrenden begutachtet wurden, sind von guter Qualität.

1 2 3 4

stimme überhaupt nicht zu stimme voll und ganz zu

Zurück

Weiter



Seite 2 von 5

Alle Eingaben

löschen

Stelle dir vor, dass deine **Hochschule** eine eigene **Website** eingerichtet hat, auf der du deine im Studium entstandenen Arbeiten veröffentlichen kannst. Voraussetzung ist, dass sie von einer **Lehrperson** begutachtet und zur Veröffentlichung **empfohlen** wurden. Deine Arbeiten können **frei** und **kostenlos** eingesehen werden. Beim Upload musst du ein **Formular** ausfüllen, in dem du deine Arbeit beschreibst.

Folgendes **Bild** zeigt, wie das Formular aussieht:

Titel:	<input type="text"/>
Sprache:	<input type="text"/>
Autorenschaft:	<input type="text"/>
Schlagwörter:	<input type="text"/>
Zusammenfassung:	<input type="text"/>
Department:	<input type="text"/>
Dokumenttyp:	<input type="text"/>
Gutachter*in:	<input type="text"/>

Bitte gib nun an, inwieweit du den folgenden Aussagen zustimmst:
Überhaupt nicht, weniger, weitgehend oder voll und ganz?

Es ist leicht, das von meiner Hochschule geforderte Upload-Formular auszufüllen.

1 2 3 4

stimme überhaupt nicht zu stimme voll und ganz zu

Meine Arbeitsergebnisse auf der Website meiner Hochschule zu veröffentlichen, kostet mich Zeit, die ich besser in mein Studium investieren könnte.

1 2 3 4

stimme überhaupt nicht zu stimme voll und ganz zu

Es ist aufwendig, meine Arbeitsergebnisse auf der Website meiner Hochschule zu veröffentlichen.

1 2 3 4

stimme überhaupt nicht zu stimme voll und ganz zu

Zurück

Weiter



Seite 3 von 5

Alle Eingaben

löschen

Ich habe Angst, dass andere von mir abschreiben, wenn ich meine Arbeitsergebnisse auf der Website meiner Hochschule veröffentliche.

1 2 3 4

stimme überhaupt nicht zu stimme voll und ganz zu

Andere könnten meine Arbeitsergebnisse als ihre eigenen ausgeben, wenn ich sie auf der Website meiner Hochschule veröffentliche.

1 2 3 4

stimme überhaupt nicht zu stimme voll und ganz zu

Ich teile mein Wissen mit meinen Kommiliton:innen, wenn sie Hilfe brauchen.

1 2 3 4

stimme überhaupt nicht zu stimme voll und ganz zu

Wenn mir geholfen wurde, möchte ich etwas zurückgeben und anderen helfen.

1 2 3 4

stimme überhaupt nicht zu stimme voll und ganz zu

Die Ergebnisse wissenschaftlicher Forschung sollten für alle frei und kostenlos zugänglich sein.

1 2 3 4

stimme überhaupt nicht zu stimme voll und ganz zu

Ich kann mir vorstellen, meine Arbeitsergebnisse auf der Website meiner Hochschule zu veröffentlichen.

1 2 3 4

stimme überhaupt nicht zu stimme voll und ganz zu

Ich kann mir vorstellen, eine wissenschaftliche Karriere einzuschlagen und selbst an einer Hochschule zu lehren und zu forschen.

1 2 3 4

stimme überhaupt nicht zu stimme voll und ganz zu

Zurück

Weiter



Seite 4 von 5

Alle Eingaben

löschen

Abschluss

Wie alt bist du?

Meine Antwort

Welchem Geschlecht fühlst du dich zugehörig?

- weiblich
- männlich
- divers

Welches Fach studierst du?

Meine Antwort _____

Welchen Abschluss strebst du gerade an?

Bachelor

Master

Sonstiges: _____

Wenn du am Gewinnspiel teilnehmen möchtest, gib bitte deine E-Mail-Adresse an:

Meine Antwort _____

Vielen Dank für deine Teilnahme!

Diese Umfrage ist anonym und dient rein wissenschaftlichen Zwecken.

Anmerkungen und Fragen gerne an: silja.wiegmann@haw-hamburg.de

Zurück

Senden

 Seite 5 von 5

Alle Eingaben

löschen

Anhang B: R-Notebook

Code ▾

1 Einleitung

Für die empirische Studie wurden im Zeitraum vom 09. Dezember 2024 bis zum 23. Januar 2025 insgesamt 97 Personen im Alter zwischen 17 und 42 Jahren befragt. Die Umfrage wurde mithilfe des Umfrage-Tools *Google Formulare* durchgeführt.

Erforderliche Bibliotheken importieren:

Hide

```
# Datenimport, -aufbereitung und -visualisierung
library(tidyverse)
library(ggfortify)
# Psychometrische Skalen, Reliabilität
library(psych)
library(splithalfr)
# Lineare Regression
library(car)
library(sandwich)
library(lmtest)
```

1.1 Hypothesen

H1: Je höher die Studierenden den Zeit- und Arbeitsaufwand des Selbstarchivierens einschätzen, desto weniger sind sie zur Selbstarchivierung in IR bereit.

H2: Je mehr Furcht vor Plagiaten die Studierenden verspüren, desto weniger sind sie zur Selbstarchivierung in IR bereit.

H3: Je altruistischer die Studierenden veranlagt sind, desto mehr sind sie zur Selbst-archivierung in IR bereit.

H4: Je mehr Wert die Studierenden auf die eigene schulische Leistung legen, desto mehr sind sie zur Selbstarchivierung in IR bereit.

H5: Je mehr die Studierenden persönliche Anerkennung suchen, desto mehr sind sie zur Selbstarchivierung in IR bereit.

H6: Je mehr Vertrauen die Studierenden in die qualitätssichernden Prozesse ihrer Institution haben, desto mehr sind sie zur Selbstarchivierung in IR bereit.

H7a: Je mehr die Studierenden bereits im Studium die Grundlage für ihren künftigen beruflichen Erfolg legen möchten, desto mehr sind sie zur Selbstarchivierung in IR bereit.

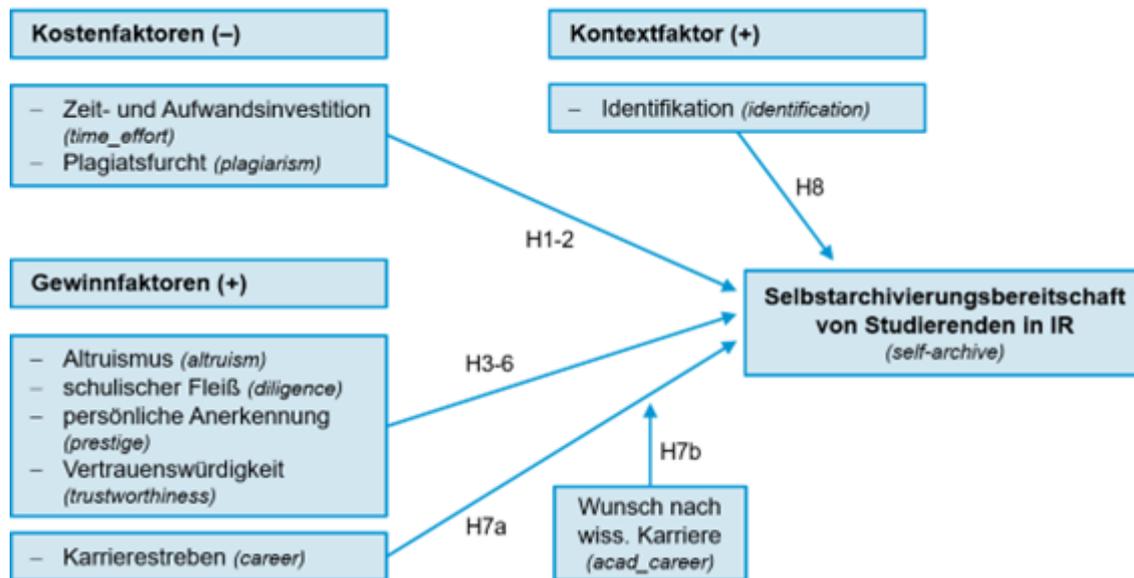
H7b: Ein stärkerer Wunsch nach einer wissenschaftlichen Karriere (*acad_career*) verstärkt den positiven Einfluss, den das Karrierestreben (*career*) auf die Bereitschaft zur Selbstarchivierung in IR (*self-archive*) ausübt.

H8: Je mehr sich die Studierenden mit ihrer Hochschule identifizieren, desto mehr sind sie zur Selbstarchivierung in IR bereit.

1.2 Empirisches Modell

Hide

```
knitr::include_graphics("model.png")
```



2 Datenaufbereitung

Aus *Google Formulare* exportierte Daten importieren:

Hide

```
df <- read_delim("selfarchive_responses.csv", ",")
```

Rows: 97 Columns: 26— Column specification

Delimiter: ","

chr (5): Zeitstempel, Welchem Geschlecht fühlst du dich zugehörig?, W...

dbl (21): Ich möchte zeigen, dass ich erfolgreich bin., Ich mag es nic...

i Use `spec()` to retrieve the full column specification for this data.

i Specify the column types or set `show_col_types = FALSE` to quiet this message.

2.1 Spalten auswählen und umbenennen

Spalte *Zeitstempel* sowie das Gewinnspiel entfernen:

Hide

```
df <- df %>%  
  select(-c("Zeitstempel", "Wenn du am Gewinnspiel teilnehmen möchtest, gib bitte deine  
  E-Mail-Adresse an:"))
```

Spalten umbenennen:

Hide

```

df <- df %>%
  rename(prestige1 = "Ich möchte zeigen, dass ich erfolgreich bin.",
         prestige2 = "Ich mag es nicht, wenn mein Name im Internet auftaucht.",
         career1 = "Gute Leistungen im Studium erhöhen meine Chancen auf dem Arbeitsmarkt.",
         career2 = "Meine Noten haben keinen Einfluss auf meine späteren Berufsaussichten.",
         diligence1 = "Ich möchte die Qualität meiner Studienleistungen ständig verbessern.",
         diligence2 = "Ich lege keinen Wert auf konstruktives Feedback.",
         identification1 = "Ich fühle mich meiner Hochschule zugehörig.",
         identification2 = "Ich bin stolz darauf, an meiner Hochschule zu studieren.",
         trustworthiness1 = "Inhalte, die meine Hochschule veröffentlicht, sind von guter Qualität.",
         trustworthiness2 = "Inhalte, die von meinen Lehrenden begutachtet wurden, sind von guter Qualität.",
         time_effort1 = "Es ist leicht, das von meiner Hochschule geforderte Upload-Formular auszufüllen.",
         time_effort2 = "Meine Arbeitsergebnisse auf der Website meiner Hochschule zu veröffentlichen, kostet mich Zeit, die ich besser in mein Studium investieren könnte.",
         time_effort3 = "Es ist aufwendig, meine Arbeitsergebnisse auf der Website meiner Hochschule zu veröffentlichen.",
         plagiarism1 = "Ich habe Angst, dass andere von mir abschreiben, wenn ich meine Arbeitsergebnisse auf der Website meiner Hochschule veröffentliche.",
         plagiarism2 = "Andere könnten meine Arbeitsergebnisse als ihre eigenen ausgeben, wenn ich sie auf der Website meiner Hochschule veröffentliche.",
         altruism1 = "Ich teile mein Wissen mit meinen Kommiliton:innen, wenn sie Hilfe brauchen.",
         altruism2 = "Wenn mir geholfen wurde, möchte ich etwas zurückgeben und anderen helfen.",
         altruism3 = "Die Ergebnisse wissenschaftlicher Forschung sollten für alle frei und kostenlos zugänglich sein.",
         selfarchive = "Ich kann mir vorstellen, meine Arbeitsergebnisse auf der Website meiner Hochschule zu veröffentlichen.",
         acad_career = "Ich kann mir vorstellen, eine wissenschaftliche Karriere einzuschlagen und selbst an einer Hochschule zu lehren und zu forschen.",
         age = "Wie alt bist du?",
         gender = "Welchem Geschlecht fühlst du dich zugehörig?",
         discipline = "Welches Fach studierst du?",
         degree = "Welchen Abschluss strebst du gerade an?"
  )

```

Whitespace entfernen:

Hide

```

df <- df %>%
  mutate(across(where(is.character), str_trim))

```

Unvollständige Antworten entfernen, unplausible Antworten korrigieren:

Hide

```
df <- df %>%
  drop_na(diligence1, time_effort1, time_effort2, time_effort3, acad_career)
df[37,24] <- "Bachelor"
```

2.2 Werte (um-)codieren

Negativ formulierte Items umcodieren:

Hide

```
df$prestige2_recode <- recode(df$prestige2, "1=4; 2=3; 3=2; 4=1")
df$career2_recode <- recode(df$career2, "1=4; 2=3; 3=2; 4=1")
df$diligence2_recode <- recode(df$diligence2, "1=4; 2=3; 3=2; 4=1")
df$time_effort1_recode <- recode(df$time_effort1, "1=4; 2=3; 3=2; 4=1")
df[,c("prestige2", "career2", "diligence2", "time_effort1")] <- list(NULL)
```

Item *gender* codieren, Datentyp *Factor*:

Hide

```
df <- df %>%
  mutate(gender = str_replace(gender, "weiblich", "0")) %>%
  mutate(gender = str_replace(gender, "männlich", "1")) %>%
  mutate(gender = str_replace(gender, "divers", "2")) %>%
  mutate(gender = as.factor(gender))
```

Item *degree* codieren, Datentyp *Factor*:

Hide

```
df <- df %>%
  mutate(degree = str_replace(degree, "Bachelor", "0")) %>%
  mutate(degree = str_replace(degree, "Master", "1")) %>%
  mutate(degree = str_replace(degree, "Promotion", "2")) %>%
  mutate(degree = as.factor(degree))
```

3 Items zu Skalen zusammenfassen

Die Items werden zu Skalen zusammengefasst.

3.1 Persönliche Anerkennung (*prestige*)

Überprüfung der internen Konsistenz anhand des Spearman-Brown-Koeffizienten:

Hide

```
spearman_brown(df$prestige1, df$prestige2_recode)
```

```
[1] 0.5976783
```

Der Spearman-Brown-Koeffizient beträgt 0.598: Die interne Konsistenz der Items ist akzeptabel, sodass die Items per Mittelwertbildung kombiniert werden:

Hide

```
df$prestige <- rowMeans(subset(df, select = c(prestige1, prestige2_recode)), na.rm = TRUE)
```

3.2 Karrierestreben (*career*)

Überprüfung der internen Konsistenz anhand des Spearman-Brown-Koeffizienten:

Hide

```
spearman_brown(df$career1, df$career2_recode)
```

```
[1] 0.5083561
```

Der Spearman-Brown-Koeffizient beträgt 0.508: Die interne Konsistenz der Items ist akzeptabel, sodass die Items per Mittelwertbildung kombiniert werden:

Hide

```
df$career <- rowMeans(subset(df, select = c(career1, career2_recode)), na.rm = TRUE)
```

3.3 Schulischer Fleiß (*diligence*)

Überprüfung der internen Konsistenz anhand des Spearman-Brown-Koeffizienten:

Hide

```
spearman_brown(df$diligence1, df$diligence2_recode)
```

```
[1] 0.1022523
```

Der Spearman-Brown-Koeffizient beträgt 0.102: Die interne Konsistenz der Items ist sehr schlecht, weshalb die Items nicht sinnvoll zu einer gemeinsamen Skala kombiniert werden können. Die Kontrollfrage *diligence2_recode*, die womöglich ein anderes Konstrukt misst als beabsichtigt oder bei der auch soziale Erwünschtheit eine Rolle gespielt haben könnte, wird entfernt:

Hide

```
df$diligence <- df$diligence1
```

3.4 Identifikation (*identification*)

Überprüfung der internen Konsistenz anhand des Spearman-Brown-Koeffizienten:

Hide

```
spearman_brown(df$identification1, df$identification2)
```

```
[1] 0.8249544
```

Der Spearman-Brown-Koeffizient beträgt 0.825: Die interne Konsistenz der Items ist gut, sodass die Items per Mittelwertbildung kombiniert werden:

Hide

```
df$identification <- rowMeans(subset(df, select = c(identification1, identification2)),
na.rm = TRUE)
```

3.5 Vertrauenswürdigkeit (*trustworthiness*)

Überprüfung der internen Konsistenz anhand des Spearman-Brown-Koeffizienten:

Hide

```
spearman_brown(df$trustworthiness1, df$trustworthiness2)
```

```
[1] 0.7098093
```

Der Spearman-Brown-Koeffizient beträgt 0.710: Die interne Konsistenz der Items ist gut, sodass die Items per Mittelwertbildung kombiniert werden:

Hide

```
df$trustworthiness <- rowMeans(subset(df, select = c(trustworthiness1, trustworthiness
2)), na.rm = TRUE)
```

3.6 Zeit- und Aufwandsinvestition (*time_effort*)

Überprüfung der internen Konsistenz mittels Cronbachs Alpha:

Hide

```
alpha(subset(df, select = c(time_effort1_recode, time_effort2, time_effort3)))
```

Reliability analysis

```
Call: alpha(x = subset(df, select = c(time_effort1_recode, time_effort2,
time_effort3)))
```

raw_alpha	std.alpha	G6(smc)	average_r	S/N	ase	mean	
<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
0.5771786	0.5645441	0.4939393	0.3017481	1.296444	0.07129736	1.751773	0.59

1 row

95% confidence boundaries

	lower	alpha	upper
	<dbl>	<dbl>	<dbl>
Feldt	0.40	0.58	0.71
Duhachek	0.44	0.58	0.72

2 rows

Reliability if an item is dropped:

	raw_alpha <dbl>	std.alpha <dbl>	G6(smc) <dbl>	average_r <dbl>	S/N <dbl>	alpha <dbl>
time_effort1_recode	0.6539916	0.6541276	0.4860249	0.4860249	1.8912394	0.07134
time_effort2	0.3772606	0.3922524	0.2439763	0.2439763	0.6454198	0.12205
time_effort3	0.2879707	0.2982244	0.1752431	0.1752431	0.4249569	0.14080

3 rows

Item statistics

	n <dbl>	raw.r <dbl>	std.r <dbl>	r.cor <dbl>	r.drop <dbl>	mean <dbl>	
time_effort1_recode	94	0.5701151	0.6470755	0.3133061	0.2436608	1.468085	0.650
time_effort2	94	0.7907614	0.7574345	0.5844486	0.4465732	1.904255	0.868
time_effort3	94	0.8220367	0.7887726	0.6437022	0.4952044	1.882979	0.890

3 rows

Non missing response frequency for each item

	1	2	3	4	miss
time_effort1_recode	0.61	0.33	0.05	0.01	0
time_effort2	0.35	0.47	0.11	0.07	0
time_effort3	0.39	0.39	0.15	0.06	0

Raw Alpha beträgt 0.577: Die interne Konsistenz der Items ist schlecht. Durch das Weglassen von *time_effort1_recode* lässt sich ein verbesserter Wert von 0.654 erreichen, weshalb dieses Item entfernt wird. Die verbleibenden Items werden per Mittelwertbildung kombiniert:

Hide

```
df$time_effort <- rowMeans(subset(df, select = c(time_effort2, time_effort3)), na.rm = TRUE)
```

3.7 Plagiatsfurcht (*plagiarism*)

Überprüfung der internen Konsistenz anhand des Spearman-Brown-Koeffizienten:

Hide

```
spearman_brown(df$plagiarism1, df$plagiarism2)
```

```
[1] 0.7951349
```

Der Spearman-Brown-Koeffizient beträgt 0.795: Die interne Konsistenz der Items ist gut, sodass die

Items per Mittelwertbildung kombiniert werden:

Hide

```
df$plagiarism <- rowMeans(subset(df, select = c(plagiarism1, plagiarism2)), na.rm = TRUE)
```

3.8 Altruismus (*altruism*)

Überprüfung der internen Konsistenz mittels Cronbachs Alpha:

Hide

```
alpha(subset(df, select = c(altruism1, altruism2, altruism3)))
```

Reliability analysis

Call: alpha(x = subset(df, select = c(altruism1, altruism2, altruism3)))

raw_alpha	std.alpha	G6(smc)	average_r	S/N	ase	mean	
<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
0.5711607	0.5907458	0.5539498	0.3248519	1.443469	0.07792734	3.599291	0.42

1 row

95% confidence boundaries

	lower	alpha	upper
	<dbl>	<dbl>	<dbl>
Feldt	0.40	0.57	0.70
Duhachek	0.42	0.57	0.72

2 rows

Reliability if an item is dropped:

	raw_alpha	std.alpha	G6(smc)	average_r	S/N	alpha se	var.r
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
altruism1	0.3296225	0.3296791	0.1973747	0.1973747	0.4918228	0.13825872	NA
altruism2	0.2821656	0.2903184	0.1698084	0.1698084	0.4090825	0.14309998	NA
altruism3	0.7376825	0.7557333	0.6073725	0.6073725	3.0938866	0.05096977	NA

3 rows

Item statistics

	n <dbl>	raw.r <dbl>	std.r <dbl>	r.cor <dbl>	r.drop <dbl>	mean <dbl>	sd <dbl>
altruism1	94	0.7565829	0.7988551	0.6937433	0.5050684	3.648936	0.4798621
altruism2	94	0.8172959	0.8112464	0.7143482	0.4886942	3.521277	0.6346402
altruism3	94	0.6457081	0.6145583	0.2476739	0.2064398	3.627660	0.6218952

3 rows

Non missing response frequency for each item

	1	2	3	4	miss
altruism1	0.00	0.00	0.35	0.65	0
altruism2	0.00	0.07	0.33	0.60	0
altruism3	0.01	0.04	0.26	0.69	0

Raw Alpha beträgt 0.571: Die interne Konsistenz der Items ist schlecht. Durch das Weglassen von *altruism3*, das möglicherweise ein anderes Konstrukt misst als beabsichtigt, lässt sich ein sehr guter Wert von 0.738 erreichen, weshalb dieses Item entfernt wird. Die verbleibenden Items werden per Mittelwertbildung kombiniert:

Hide

```
df$altruism <- rowMeans(subset(df, select = c(altruism1, altruism2)), na.rm = TRUE)
```

4 Aufbereitetes Data Frame

Vorschau auf das aufbereitete Data Frame:

Hide

```
head(df)
```

prestige1 <dbl>	career1 <dbl>	diligence1 <dbl>	identification1 <dbl>	identification2 <dbl>	trustworthiness <dbl>
2	4	4	2	2	
2	3	3	2	2	
3	3	2	1	1	
3	4	4	3	3	
4	4	4	4	4	
3	3	3	4	4	

6 rows | 1-7 of 32 columns

5 Deskriptive Statistik

Alle Items, ausgenommen der demografischen Daten, wurden mithilfe vierstufiger Skalen erfasst, sodass die Werte zwischen 1.0 und 4.0 rangieren können. Im Folgenden werden lediglich besonders auffällige Werte beschrieben.

5.1 Kontrolle der Stichprobe

Die Ausgewogenheit der Stichprobe (n=94) wird anhand soziodemografischer Merkmale überprüft.

5.1.1 Geschlecht

Vollständigkeit der Antworten überprüfen:

Hide

```
sum(is.na(df$gender))
```

```
[1] 0
```

Alle Teilnehmenden haben ihr Geschlecht angegeben, sodass n=94 beträgt.

Anteile (in %) der Geschlechter ermitteln:

Hide

```
gender_count <- df %>%
  group_by(gender) %>%
  summarise(Anzahl = n()) %>%
  mutate(Prozent = round(Anzahl / sum(Anzahl), 3))

gender_count$Prozent <- (gender_count$Prozent * 100)
```

Daten für die Visualisierung vorbereiten:

Hide

```
levels(gender_count$gender) <- c("weiblich", "männlich", "divers")
gender_count <- rename(gender_count, Geschlecht = gender)
```

Plot rendern:

Hide

```
plot_gender <- ggplot(gender_count, aes(x = "", y = Prozent, fill = Geschlecht)) +
  geom_bar(stat = "identity", width = 1, color = "white") +
  coord_polar("y", start = 0.5) +
  geom_label(aes(label = paste0(Prozent, "%")),
             color = "white",
             position = position_stack(vjust = 0.5),
             show.legend = FALSE) +
  theme_void()

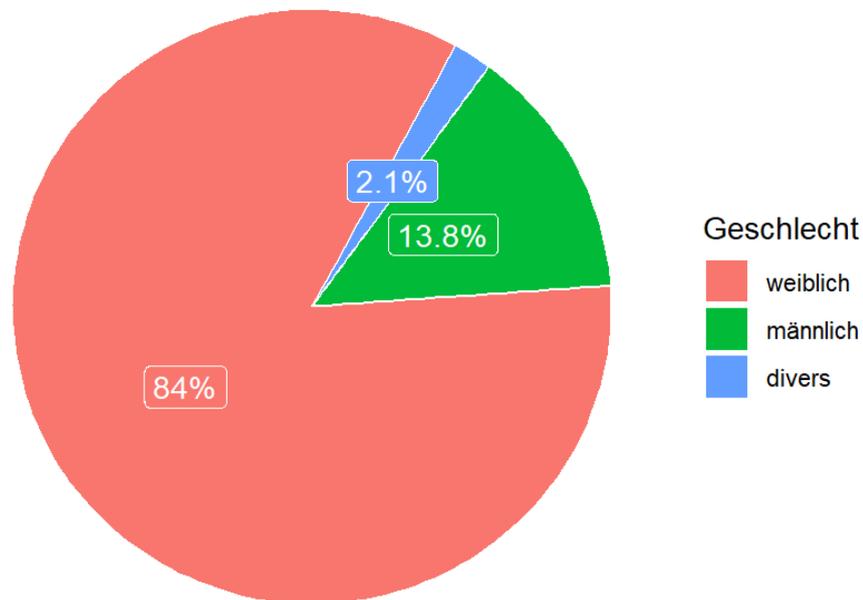
# Plot abspeichern
ggsave(plot_gender, file = "plot_gender.png")
```

```
Saving 7 x 7 in image
```

Hide

```
plot_gender + ggtitle("Anteile der Geschlechter in % (n=94)")
```

Anteile der Geschlechter in % (n=94)



(Wegen Rundungsdifferenzen ergibt die Summe der Beschriftungen nicht 100%.)

84,0% der Teilnehmenden sind weiblich, sodass die Stichprobe in Hinsicht auf das Merkmal Geschlecht nicht repräsentativ ist.

5.1.2 Alter

Vollständigkeit der Antworten überprüfen:

Hide

```
sum(is.na(df$age))
```

```
[1] 2
```

Zwei Teilnehmende haben ihr Alter nicht angegeben, sodass n=92 beträgt.

Altersgruppen (17-30 und 30+) bilden:

Hide

```
age_group <- df %>%  
  mutate(Altersgruppe = case_when(  
    age > 16 & age <= 30 ~ "17-30",  
    age > 30 ~ "30+")  
  )  
age_group <- age_group %>%  
  select("Altersgruppe") %>%  
  na.omit()
```

Anteile in Prozent ermitteln:

Hide

```
age_group <- age_group %>%
  group_by(Altersgruppe) %>%
  summarise(Anzahl = n()) %>%
  mutate(Prozent = round(Anzahl / sum(Anzahl), 3))
age_group$Prozent <- (age_group$Prozent * 100)
age_group
```

Altersgruppe <chr>	Anzahl <int>	Prozent <dbl>
17-30	85	92.4
30+	7	7.6

2 rows

92,4% der Teilnehmenden sind zwischen 17 und 30 Jahre alt. Sieben Ausreißer (7,6%) sind älter als 30 Jahre.

Histogramm der Altersverteilung:

Hide

```
plot_age_hist <- ggplot(df, aes(x = age)) +
  geom_histogram(binwidth = 1, color = "darkblue", fill = "lightblue") +
  labs(title = "Altersverteilung (n=92)", x = "Alter", y = "Absolute Häufigkeit")

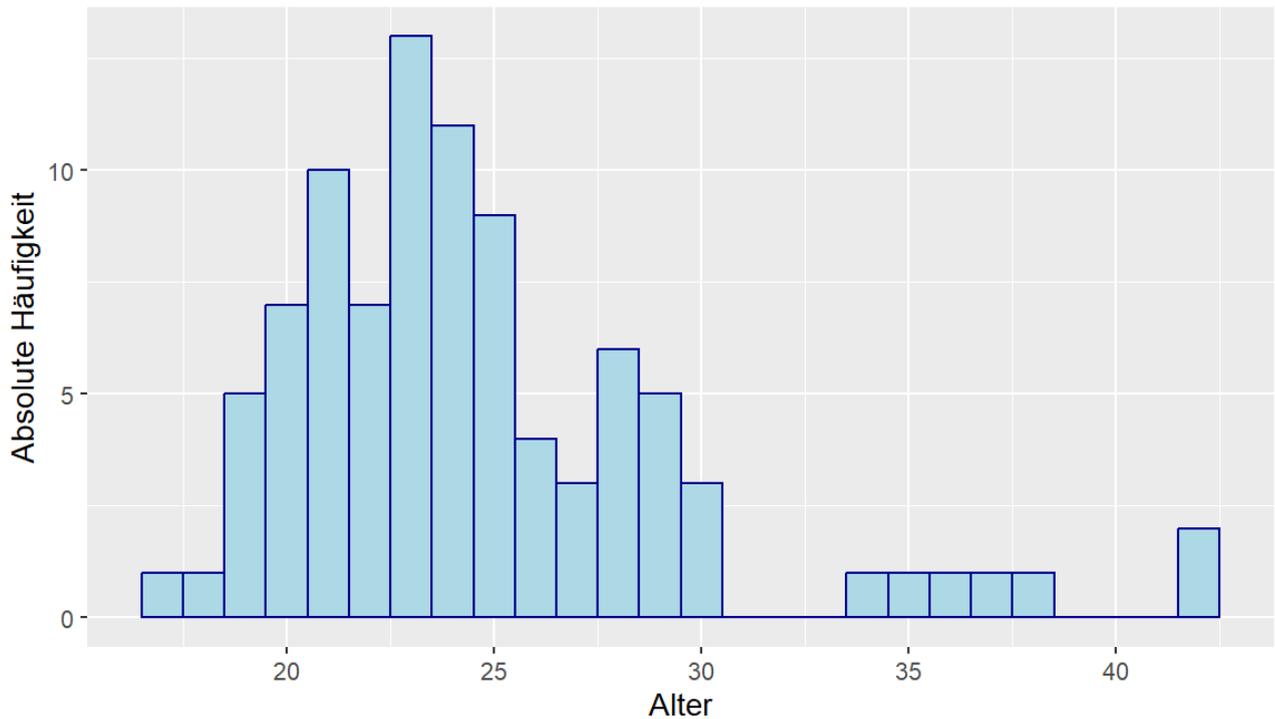
# Plot abspeichern
ggsave(plot_age_hist, file = "plot_age_hist.png")
```

Saving 7 x 7 in image

Hide

plot_age_hist

Altersverteilung (n=92)



In Relation zur Gesamtbevölkerung ist die Stichprobe in Hinsicht auf das Merkmal Alter nicht repräsentativ, was jedoch auch nicht zu erwarten war, da vor dem Hintergrund des Untersuchungsgegenstands bewusst nur Studierende befragt wurden. In der Gruppe der 17- bis 30-Jährigen ähnelt die Verteilung einer Glockenform, sodass die studentische Zielgruppe insgesamt gut abgedeckt wurde. Die sieben älteren Ausreißer führen jedoch zu einer linkssteilen Verteilung.

5.1.3 Studienfach

Vollständigkeit der Antworten überprüfen:

Hide

```
sum(is.na(df$discipline))
```

```
[1] 2
```

Zwei Teilnehmende haben ihr Studienfach nicht angegeben, sodass n=92 beträgt.

Anzahl der jeweiligen Studierenden eines Faches ermitteln:

Hide

```
discipline_count <- df %>%
  group_by(discipline) %>%
  summarise(Anzahl = n()) %>%
  rename(Studienfach = "discipline") %>%
  arrange (desc(Anzahl)) %>%
  na.omit()
discipline_count
```

Studienfach

<chr>

Anz...

<int>

Studienfach <chr>	Anz... <int>
Bibliotheks- und Informationsmanagement	36
Medien und Kommunikation	23
Digitale Transformation der Informations- und Medienwirtschaft	10
Klassische Archäologie	5
Digitale Kommunikation	4
Ethnologie	3
Germanistik	3
Geschichte	3
Kunstgeschichte	3
Empirische Kulturwissenschaft	1

1-10 of 11 rows Previous 1 2 Next

Fachgruppen bilden:

Hide

```
discipline_group <- df %>%
  mutate(Fachgruppe = case_when(discipline %in% c("Bibliotheks- und Informationsmanagem
ent", "Medien und Kommunikation", "Digitale Transformation der Informations- und Medien
wirtschaft", "Digitale Kommunikation") ~ "Informations- und Kommunikationswissenschafte
n",
  discipline %in% c("Klassische Archäologie", "Ethnologie", "Kunstgeschichte", "Empir
ische Kulturwissenschaft") ~ "Kulturwissenschaften",
  discipline %in% c("Germanistik", "Geschichte") ~ "Geisteswissenschaften",
  discipline %in% c("Physik") ~ "Naturwissenschaften"
  ))

discipline_group <- discipline_group %>%
  select("Fachgruppe") %>%
  na.omit()
```

Anteile in Prozent ermitteln:

Hide

```
discipline_group <- discipline_group %>%
  group_by(Fachgruppe) %>%
  summarise(Anzahl = n()) %>%
  mutate(Prozent = round(Anzahl / sum(Anzahl), 3))
discipline_group$Prozent <- (discipline_group$Prozent * 100)
discipline_group %>%
  arrange (desc(Prozent))
```

Fachgruppe <chr>	Anzahl <int>	Prozent <dbl>
----------------------------	------------------------	-------------------------

Fachgruppe <chr>	Anzahl <int>	Prozent <dbl>
Informations- und Kommunikationswissenschaften	73	79.3
Kulturwissenschaften	12	13.0
Geisteswissenschaften	6	6.5
Naturwissenschaften	1	1.1

4 rows

(Wegen Rundungsdifferenzen ergibt die Summe der Prozenspalte nicht 100%.)

73 der Teilnehmenden (79.3%) studieren ein informations- oder kommunikationswissenschaftliches Fach. 12 Teilnehmende (13.0%) sind Studierende der Kulturwissenschaften und sechs Teilnehmende (6.5%) studieren ein geisteswissenschaftliches Fach. Lediglich eine Person studiert ein naturwissenschaftliches Fach.

Insgesamt ist die Stichprobe in Hinsicht auf das Merkmal der Studiendisziplin nicht repräsentativ.

5.1.4 Angestrebter Abschluss

Vollständigkeit der Antworten überprüfen:

Hide

```
sum(is.na(df$degree))
```

```
[1] 0
```

Alle Teilnehmenden haben ihren aktuell angestrebten Abschluss angegeben, sodass n=94 beträgt.

Anteile (in %) der angestrebten Abschlüsse ermitteln:

Hide

```
degree_count <- df %>%
  group_by(degree) %>%
  summarise(Anzahl = n()) %>%
  mutate(Prozent = round(Anzahl / sum(Anzahl), 3))

degree_count$Prozent <- (degree_count$Prozent * 100)
```

Daten für die Visualisierung vorbereiten:

Hide

```
levels(degree_count$degree) <- c("Bachelor", "Master", "Promotion")
degree_count <- rename(degree_count, Abschluss = degree)
```

Plot rendern:

Hide

```

plot_degree <- ggplot(degree_count, aes(x = "", y = Prozent, fill = Abschluss)) +
  geom_bar(stat = "identity", width = 1, color = "white") +
  coord_polar("y", start = 0.5) +
  geom_label(aes(label = paste0(Prozent, "%")),
             color = "white",
             position = position_stack(vjust = 0.5),
             show.legend = FALSE) +
  theme_void()

# Plot abspeichern
ggsave(plot_degree, file = "plot_degree.png")

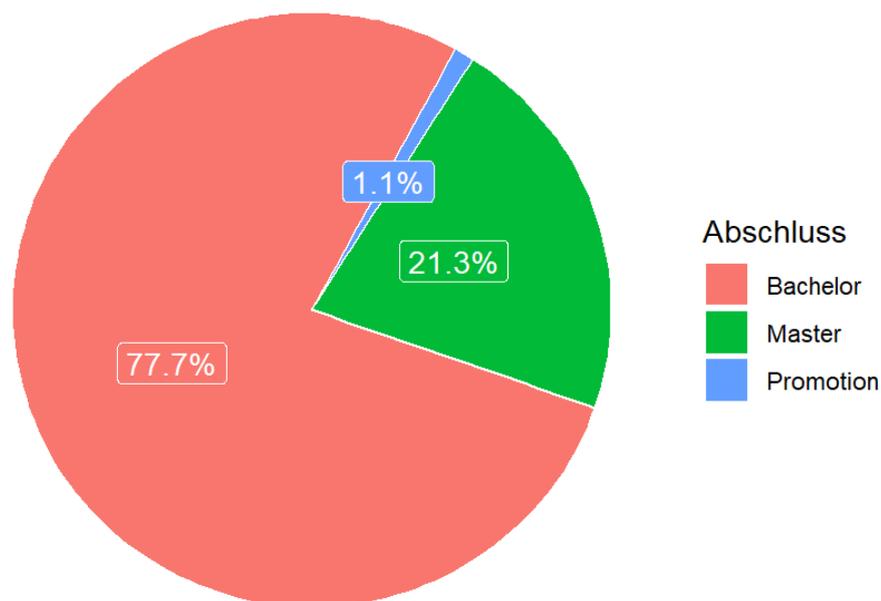
```

Saving 7 x 7 in image

Hide

```
plot_degree + ggtitle("Anteile der angestrebten Abschlüsse in % (n=94)")
```

Anteile der angestrebten Abschlüsse in % (n=94)



(Wegen Rundungsdifferenzen ergibt die Summe der Beschriftungen nicht 100%.)

77.7% der Teilnehmenden befinden sich aktuell in einem Bachelorstudiengang. Nur 21.3% der Teilnehmenden befinden sich im Master und lediglich 1.1% ($\hat{=}$ 1 Person) sind Promotionsstudierende. Verglichen mit der Anzahl der Bachelor- und Masterabschlüsse sowie der Promotionen an deutschen Hochschulen im Prüfungsjahr 2023 wird deutlich, dass Bachelorstudierende in der Stichprobe überrepräsentiert sind.

5.2 Statistische Maßzahlen der Verteilung

Hide

```
describe(df)
```

	vars	n	mean	sd	median	trimmed	mad	min	max
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
prestige1	1	94	2.78	0.94	3.0	2.84	1.48	1.0	4
career1	2	94	3.16	0.79	3.0	3.24	1.48	1.0	4
diligence1	3	94	3.20	0.84	3.0	3.29	1.48	1.0	4
identification1	4	94	2.64	0.94	3.0	2.67	1.48	1.0	4
identification2	5	94	2.61	0.94	3.0	2.63	1.48	1.0	4
trustworthiness1	6	94	2.98	0.62	3.0	2.99	0.00	1.0	4
trustworthiness2	7	94	3.06	0.60	3.0	3.08	0.00	2.0	4
time_effort2	8	94	1.90	0.87	2.0	1.79	1.48	1.0	4
time_effort3	9	94	1.88	0.89	2.0	1.78	1.48	1.0	4
plagiarism1	10	94	1.71	0.90	1.0	1.58	0.00	1.0	4

1-10 of 32 rows | 1-10 of 13 columns

Previous 1 2 3 4 Next

diligence2_recode (*Ich lege keinen Wert auf konstruktives Feedback.*): Der Median beträgt 4.0. Mit einer Skewness von -2.22 liegt im Vergleich zur symmetrischen Normalverteilung eine auffällig starke linksschiefe Verteilung vor. Möglicherweise hat soziale Erwünschtheit eine erhöhte Rolle gespielt. Das Item wurde im Zuge der Reliabilitätsanalyse entfernt.

time_effort1_recode (*Es ist leicht, das von meiner Hochschule geforderte Upload-Formular auszufüllen.*): Der Median beträgt 1.0. Mit einer Skewness von 1.27 liegt im Vergleich zur symmetrischen Normalverteilung eine stark rechtsschiefe Verteilung vor. Das Item wurde im Zuge der Reliabilitätsanalyse entfernt.

time_effort3 (*Es ist aufwendig, meine Arbeitsergebnisse auf der Website meiner Hochschule zu veröffentlichen.*): Der Median beträgt 2.0. Mit einer Skewness von 0.77 liegt eine rechtsschiefe Verteilung vor, die jedoch geringer ausfällt als beim Item *time_effort1_recode*.

time_effort (*kombinierte Skala*): Der Median beträgt 2.0. Mit einer Skewness von 0.72 liegt eine rechtsschiefe Verteilung vor, was zeigt, dass die Studierenden in der Stichprobe den Zeit- und Arbeitsaufwand des Selbstarchivierens eher gering einschätzen.

Plot rendern:

Hide

```
plot_time_effort <- ggplot(df, aes(x = time_effort)) +
  geom_histogram(binwidth = 0.5, color = "darkblue", fill = "lightblue") +
  labs(title = "Verteilung der Variable 'time_effort'", x = "time_effort", y = "Absolute Häufigkeit")

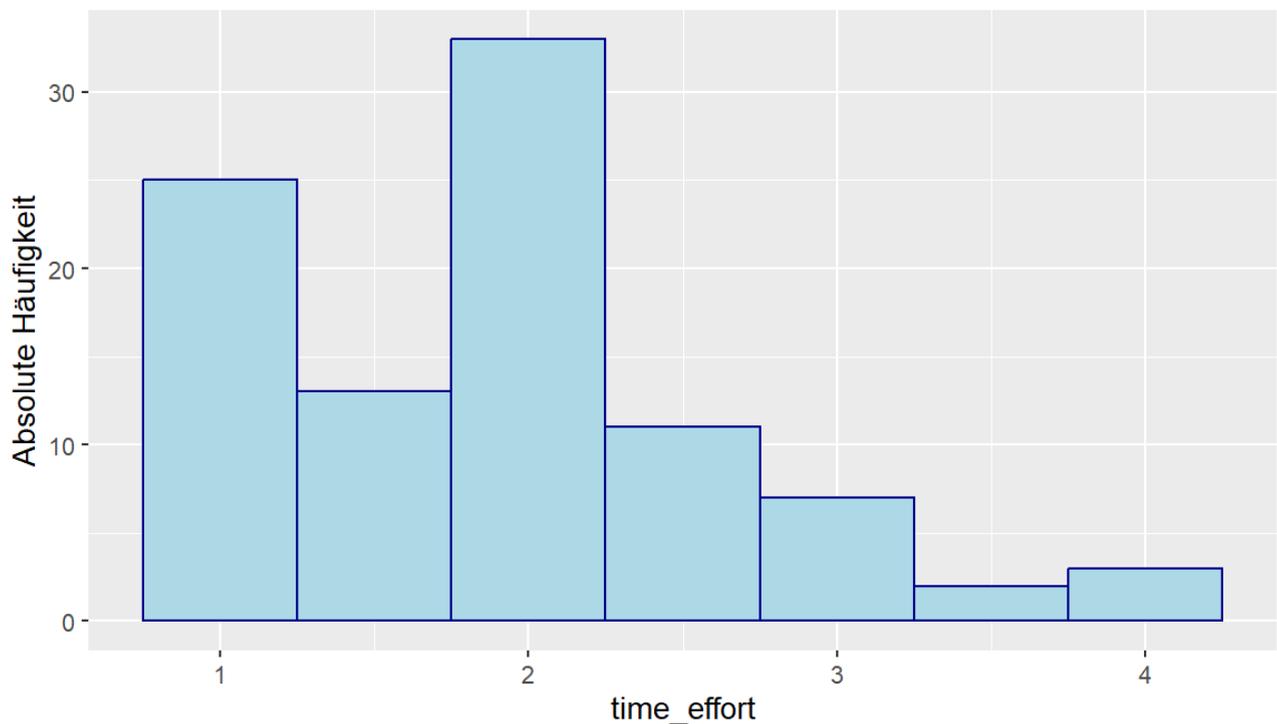
# Plot abspeichern
ggsave(plot_time_effort, file = "plot_time_effort.png")
```

Saving 7 x 7 in image

Hide

```
plot_time_effort
```

Verteilung der Variable 'time_effort'



altruism (kombinierte Skala): Der Median beträgt 4.0. Mit einer Skewness von -0.78 liegt eine linksschiefe Verteilung vor, was zeigt, dass die Studierenden in der Stichprobe eher altruistisch veranlagt sind. Hierbei könnte allerdings auch soziale Erwünschtheit eine Rolle gespielt haben.

career (kombinierte Skala): Der Median beträgt 3.0. Mit einer Skewness von -0.44 liegt eine linksschiefe Verteilung vor, was zeigt, dass die Studierenden in der Stichprobe eher mehr bereits im Studium die Grundlage für ihren künftigen beruflichen Erfolg legen möchten.

diligence (kombinierte Skala): Der Median beträgt 3.0. Mit einer Skewness von -0.71 liegt eine linksschiefe Verteilung vor, was zeigt, dass die Studierenden in der Stichprobe eher mehr Wert auf die eigene schulische Leistung legen.

prestige (kombinierte Skala): Der Median beträgt 2.5. Mit einer Skewness von 0.03 liegt eine minimal rechtsschiefe Verteilung vor, die jedoch sehr nahe an einer symmetrischen Normalverteilung liegt. Das Verhältnis zwischen Studierenden, die eher mehr bzw. weniger nach persönlicher Anerkennung suchen, ist in der Stichprobe sehr ausgewogen.

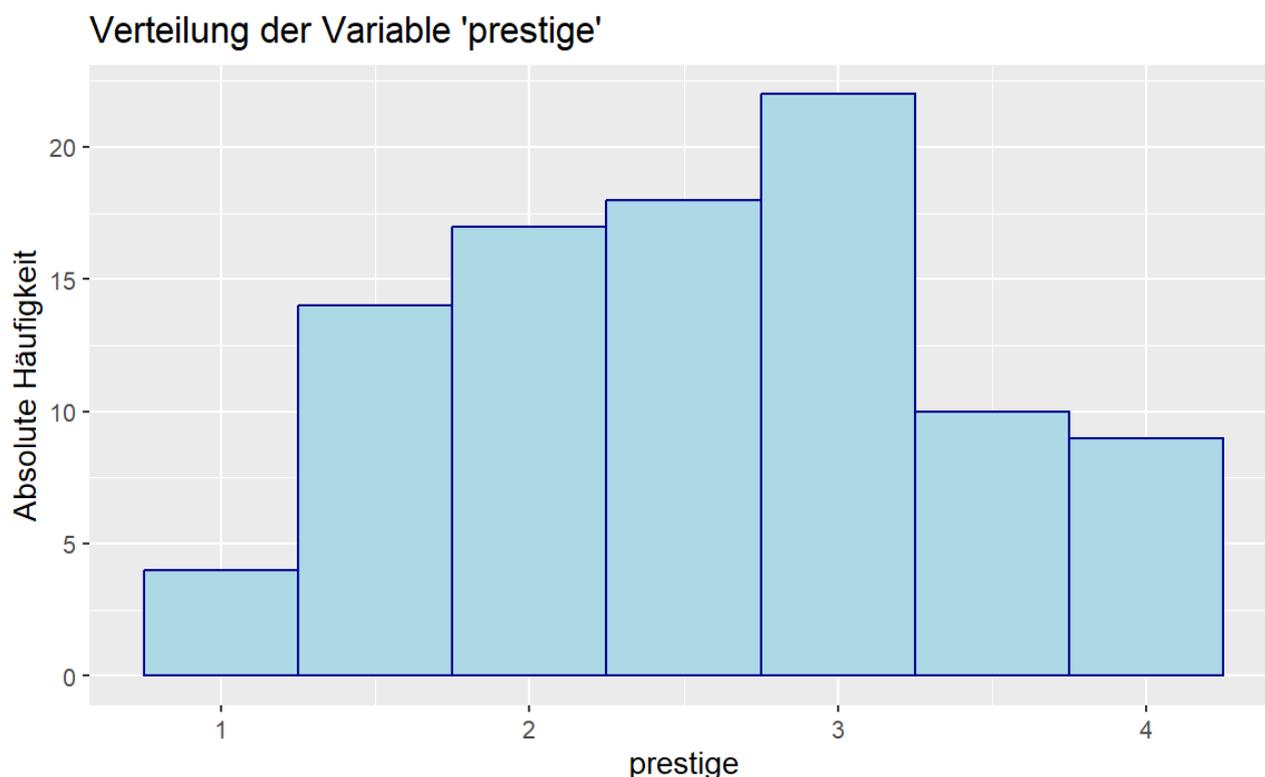
Plot rendern:

Hide

```
plot_prestige <- ggplot(df, aes(x = prestige)) +  
  geom_histogram(binwidth = 0.5, color = "darkblue", fill = "lightblue") +  
  labs(title = "Verteilung der Variable 'prestige'", x = "prestige", y = "Absolute Häufigkeit")  
  
# Plot abspeichern  
ggsave(plot_prestige, file = "plot_prestige.png")
```

Saving 7 x 7 in image

plot_prestige



plagiarism (*kombinierte Skala*): Der Median beträgt 1.5. Mit einer Skewness von 0.68 liegt eine rechtsschiefe Verteilung vor, was zeigt, dass die Studierenden in der Stichprobe eher weniger Furcht vor Plagiaten verspüren.

trustworthiness (*kombinierte Skala*): Der Median beträgt 3.0. Mit einer Skewness von -0.19 liegt eine leicht linksschiefe Verteilung vor, was zeigt, dass die Studierenden in der Stichprobe eher mehr Vertrauen in die qualitätssichernden Prozesse ihrer Institution haben.

identification (*kombinierte Skala*): Der Median beträgt 2.5. Mit einer Skewness von -0.20 liegt eine leicht linksschiefe Verteilung vor, was zeigt, dass die Studierenden in der Stichprobe sich eher mehr mit ihrer Hochschule identifizieren.

acad_career: Der Median beträgt 2.0. Mit einer Skewness von 0.77 liegt eine rechtsschiefe Verteilung vor, was zeigt, dass die Studierenden in der Stichprobe eher weniger den Wunsch nach einer wissenschaftlichen Karriere verspüren.

selfarchive: Der Median beträgt 3.0. Mit einer Skewness von -0.84 liegt eine linksschiefe Verteilung vor, was zeigt, dass die Studierenden in der Stichprobe eher mehr zur Selbstarchivierung bereit sind.

6 Überprüfung der Einzelhypothesen mittels einfacher linearer Regressionen

Die Hypothesen werden isoliert mittels einfacher linearer Regressionen überprüft.

6.1 H1 (*time_effort* → *selfarchive*)

Lineares Modell erstellen:

```
fit1 <- lm(selfarchive ~ time_effort, data = df)
```

Scatterplot rendern:

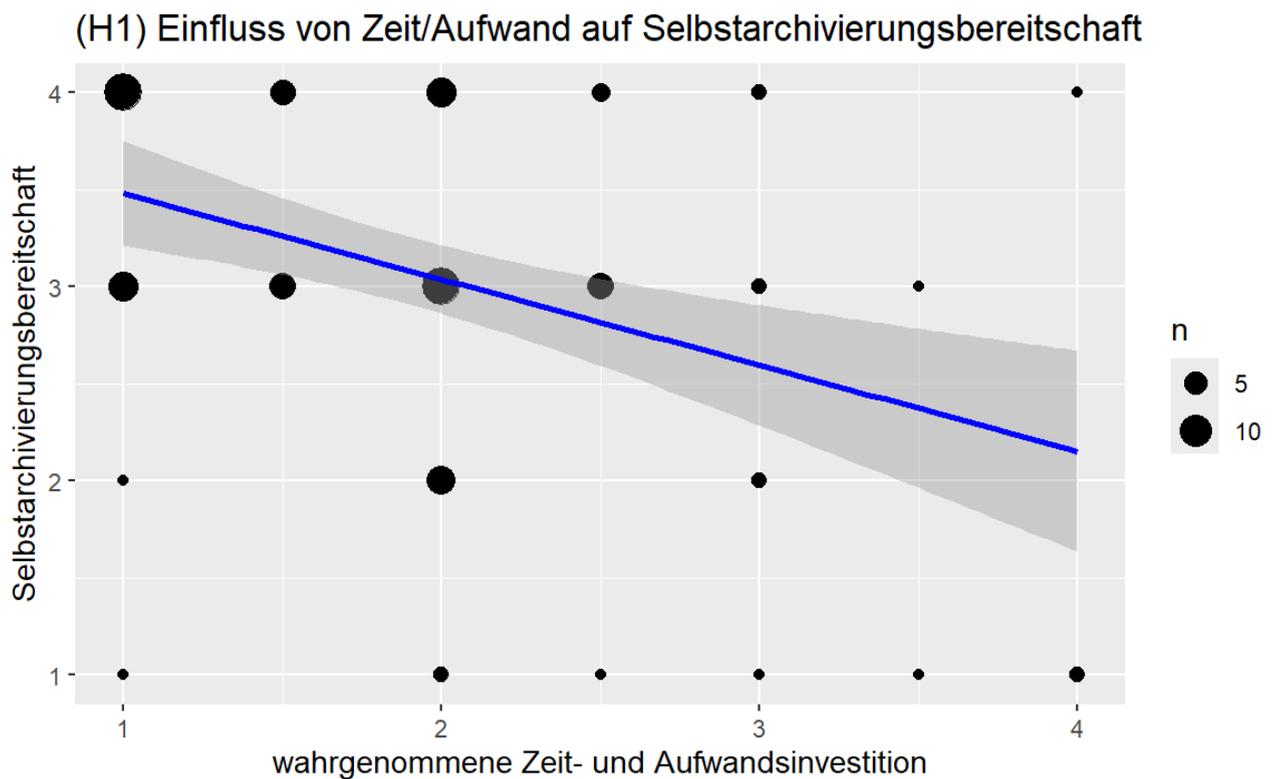
Hide

```
plot_h1 <- ggplot(df, aes(x = time_effort, y = selfarchive)) +  
  geom_count() +  
  scale_size_area() +  
  stat_smooth(method = "lm", col = "blue") +  
  labs(title = "(H1) Einfluss von Zeit/Aufwand auf Selbstarchivierungsbereitschaft", x  
= "wahrgenommene Zeit- und Aufwandsinvestition", y = "Selbstarchivierungsbereitschaft")  
  
# Plot abspeichern  
ggsave(plot_h1, file = "plot_h1.png")
```

Saving 7 x 7 in image

Hide

plot_h1



Summary des linearen Modells:

Hide

```
summary(fit1)
```

```

Call:
lm(formula = selfarchive ~ time_effort, data = df)

Residuals:
    Min       1Q   Median       3Q      Max
-2.48059 -0.48059 -0.03803  0.51941  1.84710

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.9232     0.2377  16.506 < 2e-16 ***
time_effort -0.4426     0.1166  -3.795 0.000264 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8524 on 92 degrees of freedom
Multiple R-squared:  0.1354,    Adjusted R-squared:  0.126
F-statistic: 14.4 on 1 and 92 DF,  p-value: 0.0002641

```

R-squared beträgt 0.135, sodass das lineare Modell die Varianz in der abhängigen Variablen zu 14% erklären kann.

Überprüfung auf Heteroskedastizität:

Hide

```

# Plot abspeichern
png(file = "h1_hs.png")
plot(fit1, 1)
dev.off()

```

```

null device
      1

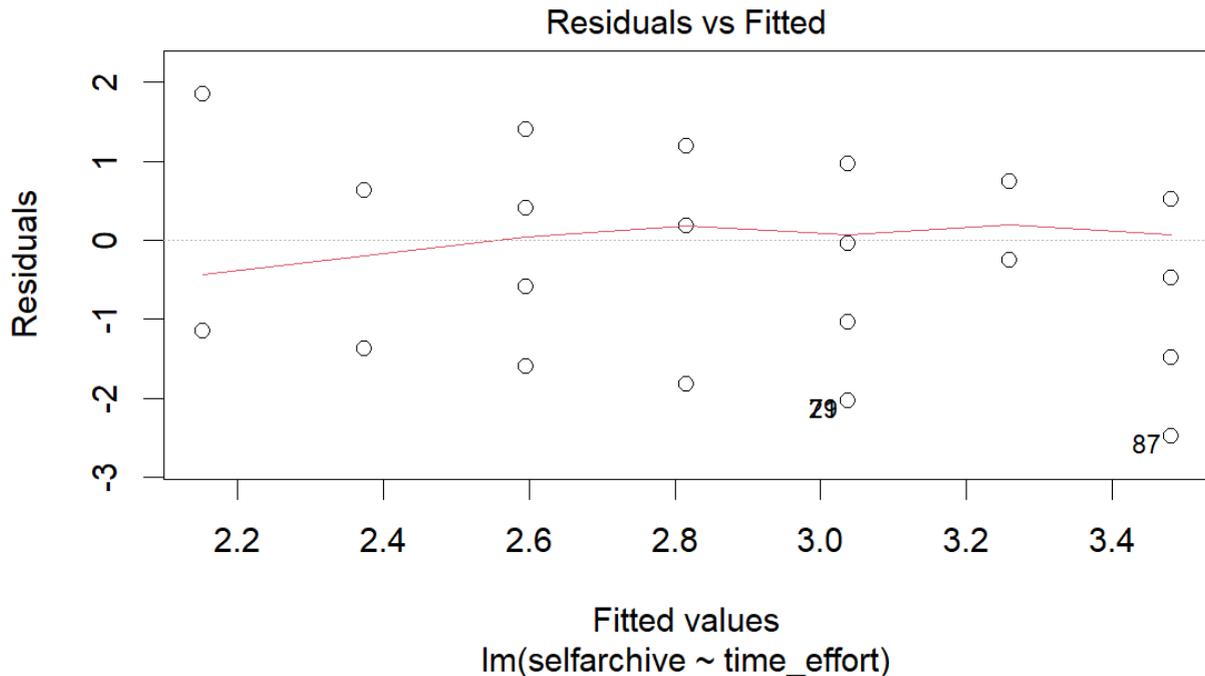
```

Hide

```

h1_hs <- plot(fit1, 1)

```



Die Streuung der Residuen ist nicht linear, was die Berechnung robuster Standardfehler erforderlich macht, um einer Verzerrung des p-Wertes des Koeffizienten vorzubeugen:

Hide

```
coefTest(fit1, vcov = vcovHC(fit1))
```

t test of coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	3.92315	0.27136	14.4575	< 2.2e-16 ***
time_effort	-0.44256	0.15097	-2.9316	0.004255 **

 Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Der robuste p-Wert beträgt 0.004 und liegt damit unter dem Signifikanzniveau von 0.05, sodass die Irrtumswahrscheinlichkeit gering ist. Die Nullhypothese wird verworfen und die Alternativhypothese angenommen.

Der Koeffizient beträgt -0.443, sodass ein negativer Zusammenhang zwischen den Variablen besteht: Wenn die wahrgenommene Zeit- und Aufwandsinvestition um eine Einheit steigt, dann sinkt die Bereitschaft zur Selbstarchivierung in IR um 0.443.

Visualisierung der Selbstarchivierungsbereitschaft nach Aufwandswahrnehmung:

Aufwandswahrnehmungsgruppen erstellen:

Hide

```
df <- df %>% mutate(Aufwandswahrnehmung = ifelse(time_effort <= 2, "geringer Aufwand", "hoher Aufwand"))
```

Aufwandswahrnehmungsgruppen in absoluten Zahlen:

Hide

```
time_effort_absolut <- df %>%  
  group_by(Aufwandswahrnehmung) %>%  
  summarise(Anzahl = n())  
time_effort_absolut
```

Aufwandswahrnehmung	Anzahl
<chr>	<int>
geringer Aufwand	71
hoher Aufwand	23

2 rows

Anteile (in %) der Aufwandswahrnehmung ermitteln:

Hide

```
time_effort_count <- df %>%  
  count(Aufwandswahrnehmung, selfarchive) %>%  
  group_by(Aufwandswahrnehmung) %>%  
  mutate(Prozent = n / sum(n)) %>%  
  mutate(Prozent = round(n / sum(n), 3))  
  
time_effort_count$Prozent <- (time_effort_count$Prozent * 100)
```

Daten für die Visualisierung vorbereiten:

Hide

```
time_effort_count <- time_effort_count %>%  
  mutate(selfarchive = as.factor(selfarchive)) %>%  
  mutate(selfarchive = str_replace(selfarchive, "1", "gar nicht")) %>%  
  mutate(selfarchive = str_replace(selfarchive, "2", "weniger")) %>%  
  mutate(selfarchive = str_replace(selfarchive, "3", "weitgehend")) %>%  
  mutate(selfarchive = str_replace(selfarchive, "4", "voll und ganz"))  
  
time_effort_count$selfarchive <- ordered(time_effort_count$selfarchive, levels = c("gar  
nicht", "weniger", "weitgehend", "voll und ganz"))
```

Gestapeltes Balkendiagramm plotten:

Hide

```

plot_time_effort2 <- ggplot(time_effort_count, aes(x = Aufwandswahrnehmung, y = Prozent
, fill = selfarchive)) +
  geom_col(width = 0.7) +
  geom_text(aes(label = paste0(Prozent, "%")),
            color = "white",
            position = position_stack(vjust = 0.5)) +
  labs(x = "Aufwandswahrnehmung", y = "Prozent") +
  scale_fill_hue(l = 65, c = 100)

# Plot abspeichern
ggsave(plot_time_effort2, file = "plot_time_effort2.png")

```

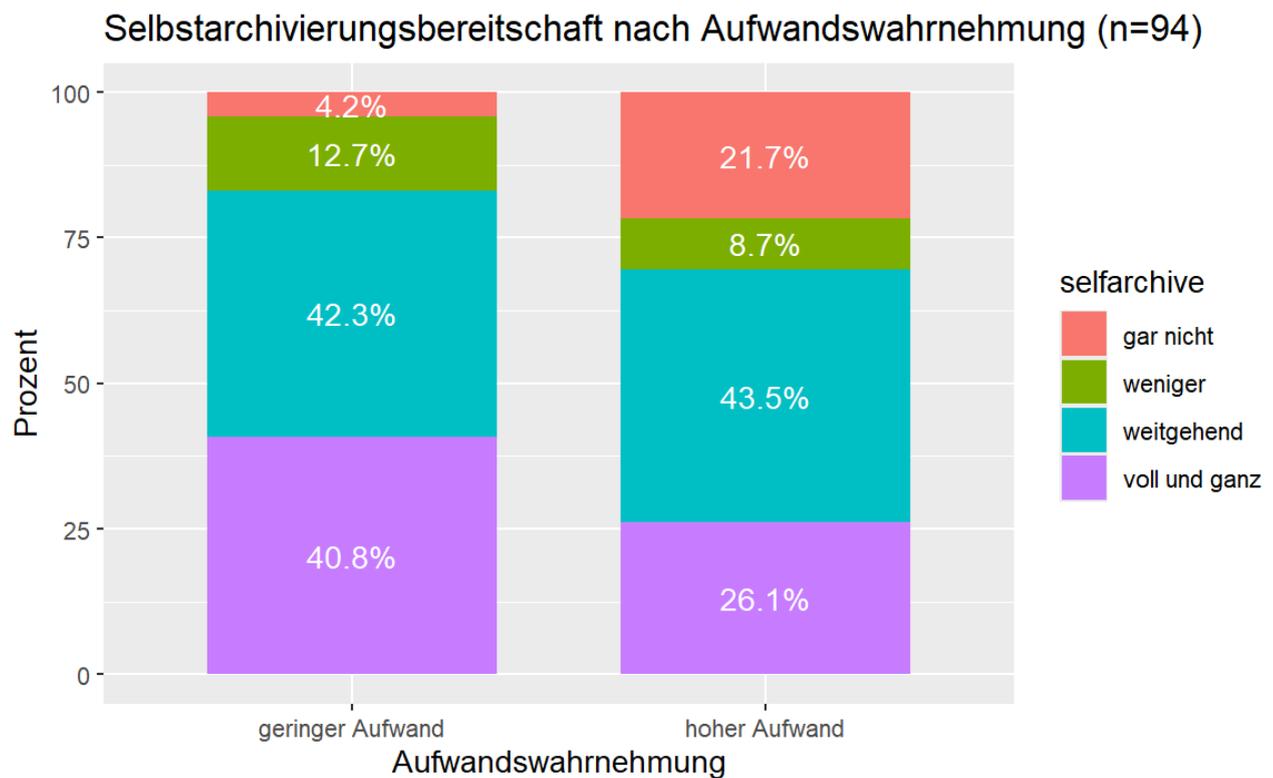
Saving 7 x 7 in image

Hide

```

plot_time_effort2 + ggtitle("Selbstarchivierungsbereitschaft nach Aufwandswahrnehmung
(n=94)")

```



6.2 H2 (plagiarism → selfarchive)

Lineares Modell erstellen:

Hide

```

fit2 <- lm(selfarchive ~ plagiarism, data = df)

```

Scatterplot rendern:

Hide

```

plot_h2 <- ggplot(df, aes(x = plagiarism, y = selfarchive)) +
  geom_count() +
  scale_size_area() +
  stat_smooth(method = "lm", col = "blue") +
  labs(title = "(H2) Einfluss von Plagiatsfurcht auf Selbstarchivierungsbereitschaft",
x = "Plagiatsfurcht", y = "Selbstarchivierungsbereitschaft")

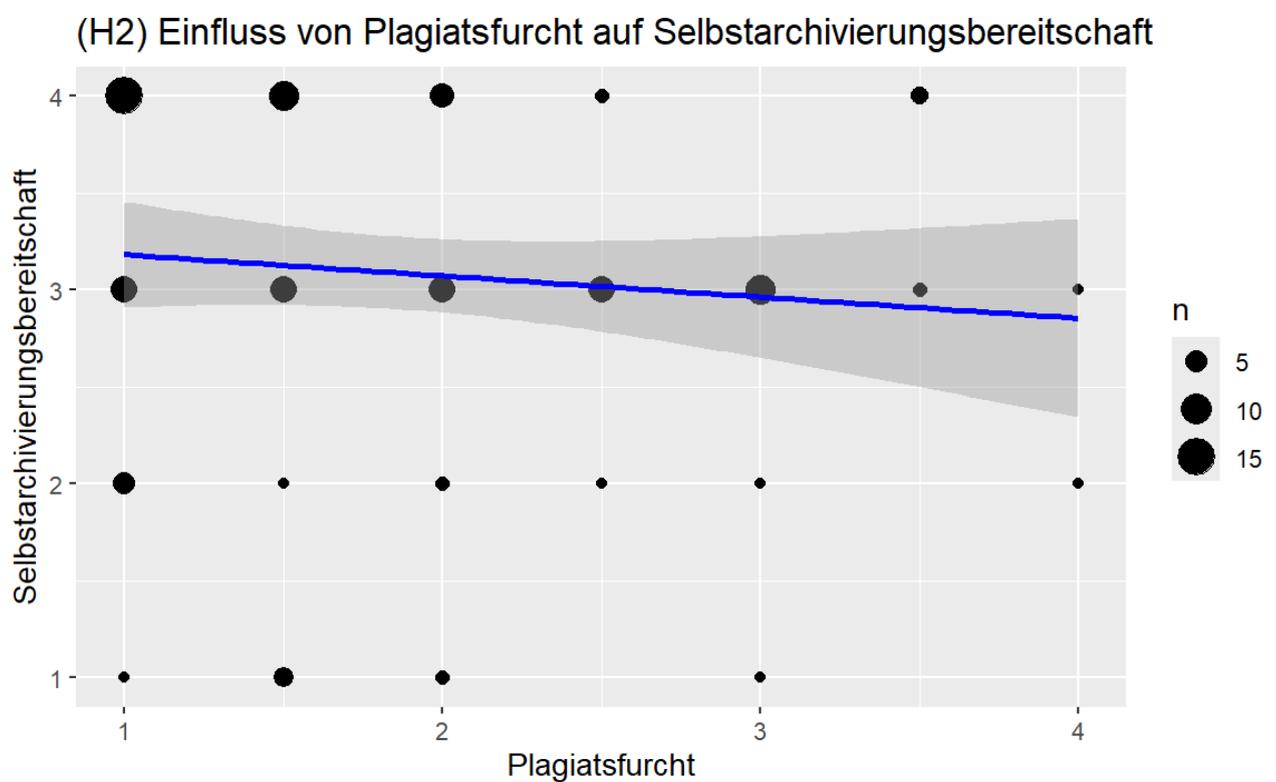
# Plot abspeichern
ggsave(plot_h2, file = "plot_h2.png")

```

Saving 7 x 7 in image

Hide

plot_h2



Summary des linearen Modells:

Hide

```
summary(fit2)
```

```

Call:
lm(formula = selfarchive ~ plagiarism, data = df)

Residuals:
    Min       1Q   Median       3Q      Max
-2.18149 -0.18149  0.00953  0.81851  1.09140

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.2906     0.2333  14.103  <2e-16 ***
plagiarism   -0.1092     0.1134  -0.963   0.338
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9122 on 92 degrees of freedom
Multiple R-squared:  0.009972, Adjusted R-squared:  -0.0007887
F-statistic: 0.9267 on 1 and 92 DF,  p-value: 0.3382

```

R-squared beträgt 0.010, sodass das lineare Modell die Varianz in der abhängigen Variablen zu lediglich 1% erklären kann.

Überprüfung auf Heteroskedastizität:

Hide

```

# Plot abspeichern
png(file = "h2_hs.png")
plot(fit2, 1)
dev.off()

```

```

null device
      1

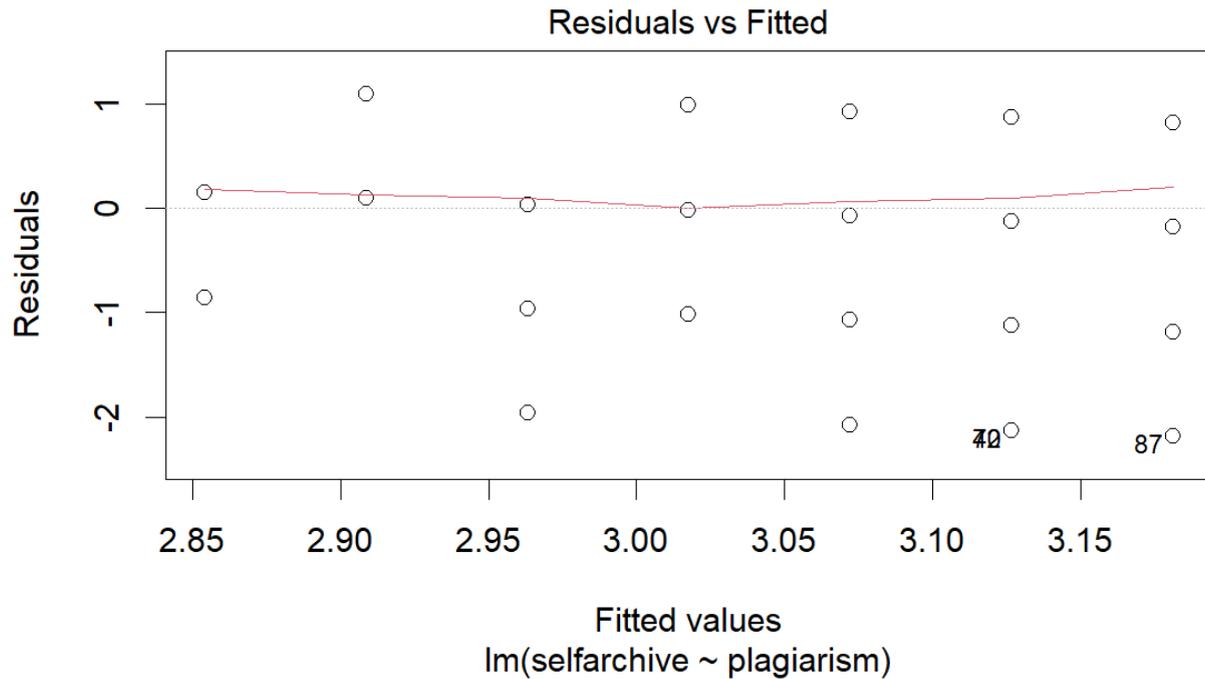
```

Hide

```

h2_hs <- plot(fit2, 1)

```



Die Streuung der Residuen ist nicht linear, was die Berechnung robuster Standardfehler erforderlich macht, um einer Verzerrung des p-Wertes des Koeffizienten vorzubeugen:

Hide

```
coefTest(fit2, vcov = vcovHC(fit2))
```

t test of coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	3.29065	0.23303	14.1213	<2e-16 ***
plagiarism	-0.10916	0.10260	-1.0639	0.2902

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Der robuste p-Wert beträgt 0.290 und liegt damit deutlich über dem Signifikanzniveau von 0.05, sodass die Irrtumswahrscheinlichkeit sehr hoch ist. Die Nullhypothese kann nicht verworfen und die Interpretation nicht sinnvoll fortgesetzt werden.

6.3 H3 (*altruism* → *selfarchive*)

Lineares Modell erstellen:

Hide

```
fit3 <- lm(selfarchive ~ altruism, data = df)
```

Scatterplot rendern:

Hide

```

plot_h3 <- ggplot(df, aes(x = altruism, y = selfarchive)) +
  geom_count() +
  scale_size_area() +
  stat_smooth(method = "lm", col = "blue") +
  labs(title = "(H3) Einfluss von Altruismus auf Selbstarchivierungsbereitschaft", x =
"Altruismus", y = "Selbstarchivierungsbereitschaft")

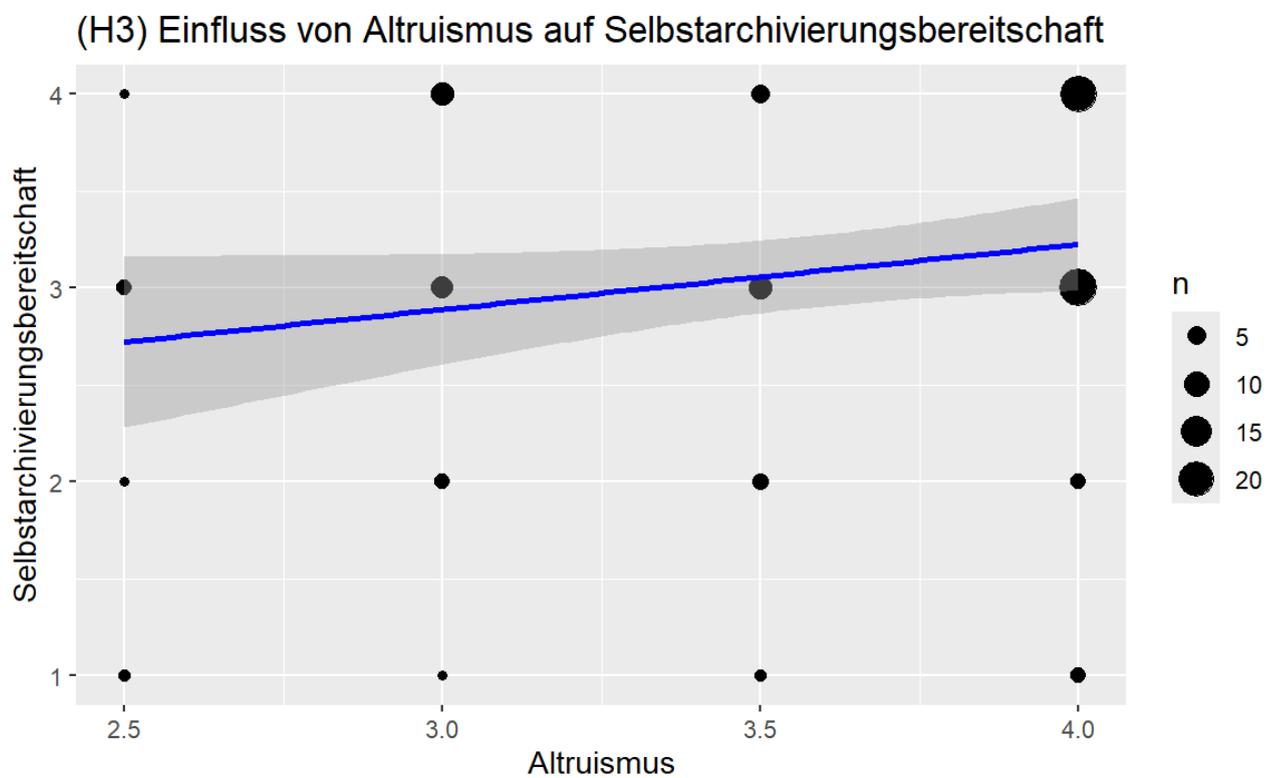
# Plot abspeichern
ggsave(plot_h3, file = "plot_h3.png")

```

Saving 7 x 7 in image

Hide

plot_h3



Summary des linearen Modells:

Hide

```
summary(fit3)
```

```

Call:
lm(formula = selfarchive ~ altruism, data = df)

Residuals:
    Min       1Q   Median       3Q      Max
-2.22422 -0.22422 -0.05657  0.77578  1.27874

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.8830     0.6754   2.788  0.00644 **
altruism     0.3353     0.1866   1.797  0.07562 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9011 on 92 degrees of freedom
Multiple R-squared:  0.03391, Adjusted R-squared:  0.02341
F-statistic: 3.229 on 1 and 92 DF, p-value: 0.07562

```

R-squared beträgt 0.034, sodass das lineare Modell die Varianz in der abhängigen Variablen zu lediglich 3% erklären kann.

Überprüfung auf Heteroskedastizität:

Hide

```

# Plot abspeichern
png(file = "h3_hs.png")
plot(fit3, 1)
dev.off()

```

```

null device
      1

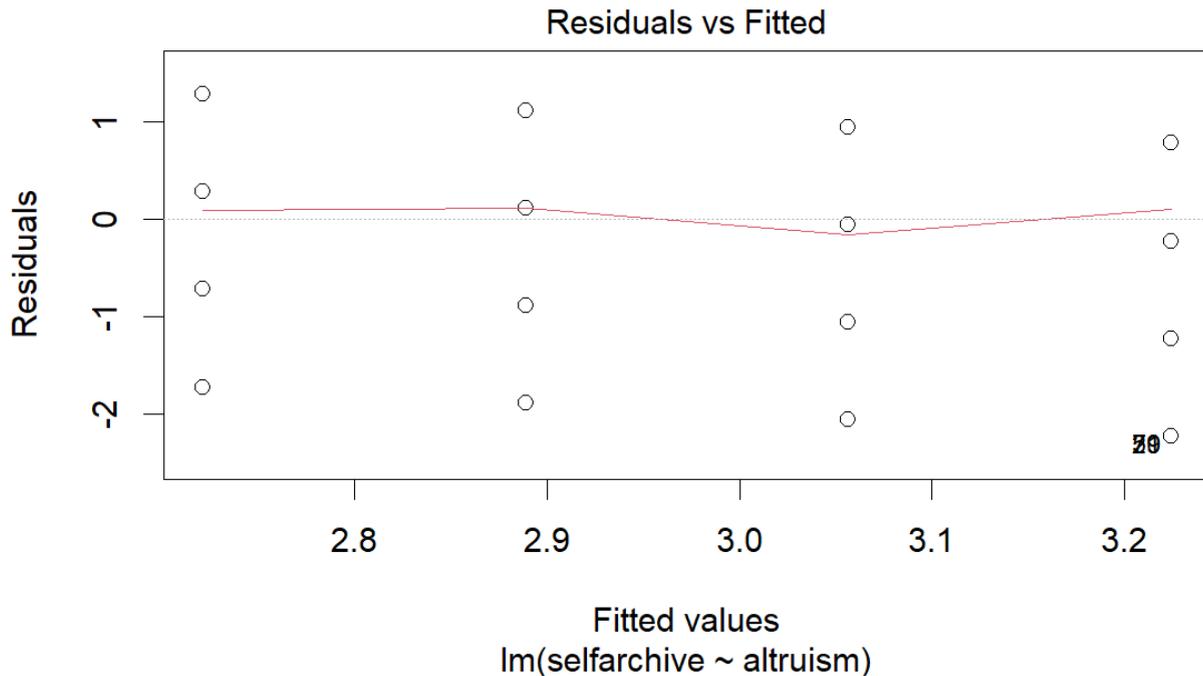
```

Hide

```

h3_hs <- plot(fit3, 1)

```



Die Streuung der Residuen ist nicht linear, was die Berechnung robuster Standardfehler erforderlich macht, um einer Verzerrung des p-Wertes des Koeffizienten vorzubeugen:

Hide

```
coefTest(fit3, vcov = vcovHC(fit3))
```

t test of coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.88298	0.75718	2.4868	0.01469 *
altruism	0.33531	0.20508	1.6350	0.10546

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Der robuste p-Wert beträgt 0.105 und liegt damit deutlich über dem Signifikanzniveau von 0.05, sodass die Irrtumswahrscheinlichkeit hoch ist. Die Nullhypothese kann nicht verworfen und die Interpretation nicht sinnvoll fortgesetzt werden.

6.4 H4 (*diligence* → *selfarchive*)

Lineares Modell erstellen:

Hide

```
fit4 <- lm(selfarchive ~ diligence, data = df)
```

Scatterplot rendern:

Hide

```

plot_h4 <- ggplot(df, aes(x = diligence, y = selfarchive)) +
  geom_count() +
  scale_size_area() +
  stat_smooth(method = "lm", col = "blue") +
  labs(title = "(H4) Einfluss von schulischer Leistung auf Selbstarchivierungsbereitsch
aft", x = "Wert auf schulische Leistung", y = "Selbstarchivierungsbereitschaft")

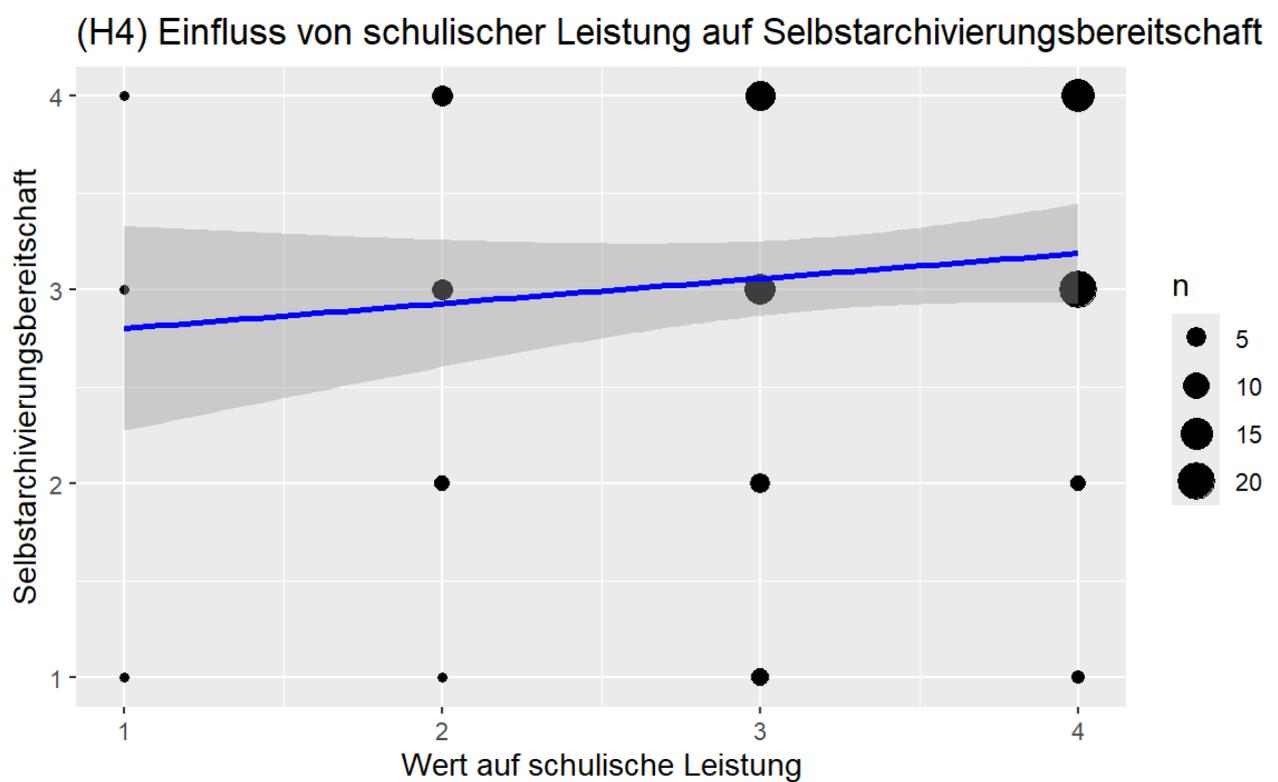
# Plot abspeichern
ggsave(plot_h4, file = "plot_h4.png")

```

Saving 7 x 7 in image

Hide

plot_h4



Summary des linearen Modells:

Hide

```
summary(fit4)
```

```

Call:
lm(formula = selfarchive ~ diligence, data = df)

Residuals:
    Min       1Q   Median       3Q      Max
-2.1878 -0.1878 -0.0591  0.8122  1.1982

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.6731     0.3731   7.164 1.89e-10 ***
diligence    0.1287     0.1128   1.141  0.257
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9103 on 92 degrees of freedom
Multiple R-squared:  0.01395,    Adjusted R-squared:  0.003231
F-statistic: 1.301 on 1 and 92 DF,  p-value: 0.2569

```

R-squared beträgt 0.014, sodass das lineare Modell die Varianz in der abhängigen Variablen zu lediglich 1% erklären kann.

Überprüfung auf Heteroskedastizität:

Hide

```

# Plot abspeichern
png(file = "h4_hs.png")
plot(fit4, 1)
dev.off()

```

```

null device
      1

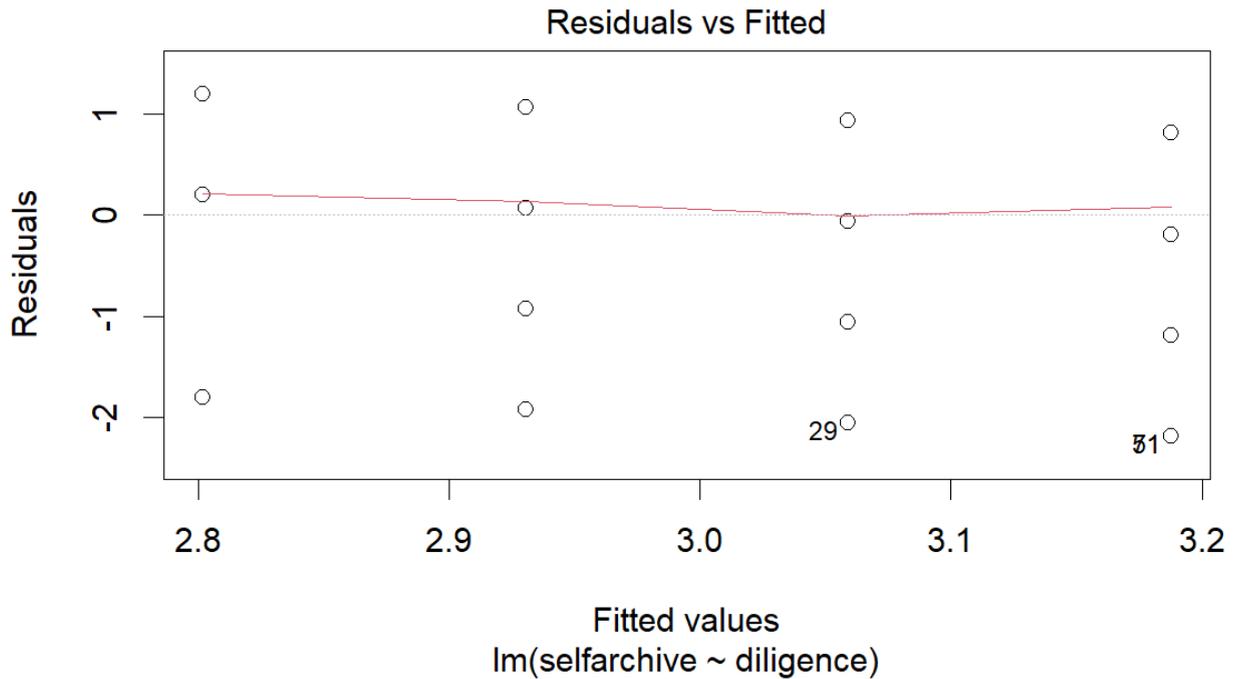
```

Hide

```

h4_hs <- plot(fit4, 1)

```



Die Streuung der Residuen ist nicht linear, was die Berechnung robuster Standardfehler erforderlich macht, um einer Verzerrung des p-Wertes des Koeffizienten vorzubeugen:

Hide

```
coefTest(fit4, vcov = vcovHC(fit4))
```

t test of coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	2.67314	0.42902	6.2309	1.377e-08	***
diligence	0.12865	0.12448	1.0335	0.3041	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Der robuste p-Wert beträgt 0.304 und liegt damit deutlich über dem Signifikanzniveau von 0.05, sodass die Irrtumswahrscheinlichkeit sehr hoch ist. Die Nullhypothese kann nicht verworfen und die Interpretation nicht sinnvoll fortgesetzt werden.

6.5 H5 (*prestige* → *selfarchive*)

Lineares Modell erstellen:

Hide

```
fit5 <- lm(selfarchive ~ prestige, data = df)
```

Scatterplot rendern:

Hide

```

plot_h5 <- ggplot(df, aes(x = prestige, y = selfarchive)) +
  geom_count() +
  scale_size_area() +
  stat_smooth(method = "lm", col = "blue") +
  labs(title = "(H5) Einfluss von Prestigestreben auf Selbstarchivierungsbereitschaft",
x = "Streben nach persönlicher Anerkennung", y = "Selbstarchivierungsbereitschaft")

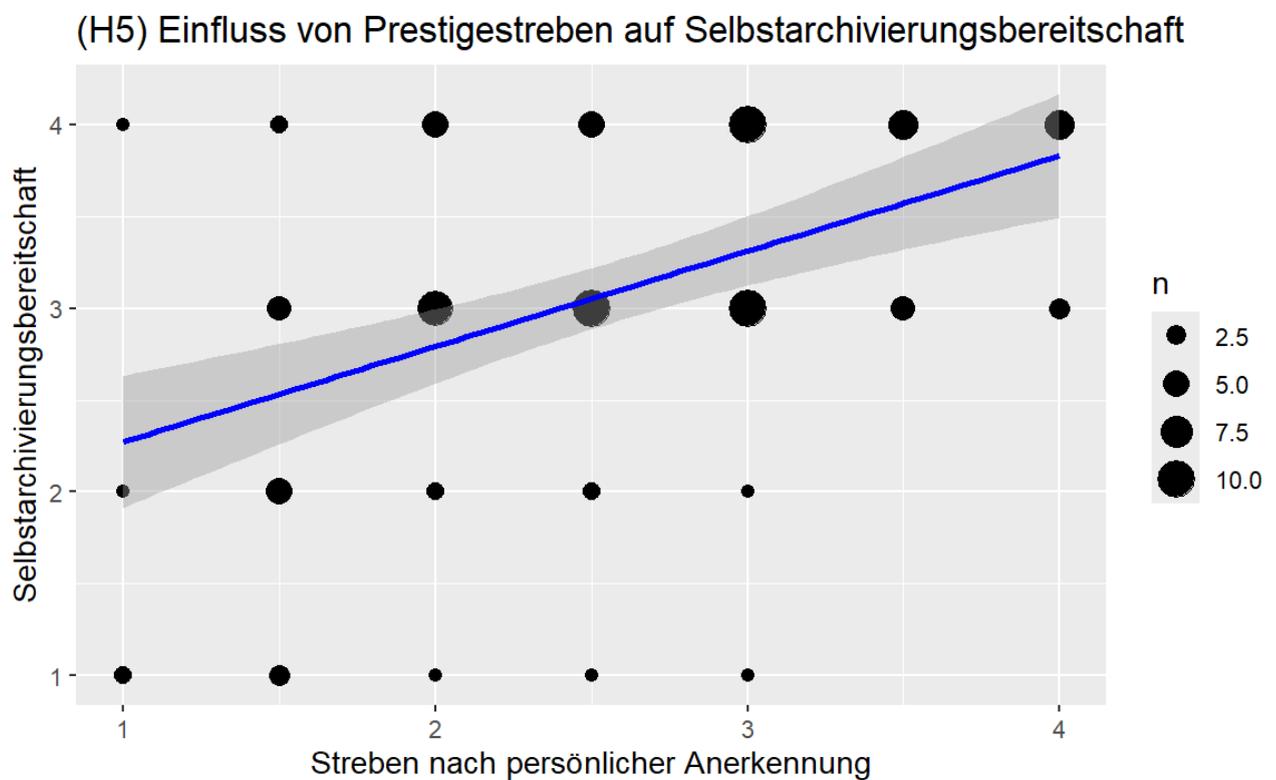
# Plot abspeichern
ggsave(plot_h5, file = "plot_h5.png")

```

Saving 7 x 7 in image

Hide

plot_h5



Summary des linearen Modells:

Hide

```
summary(fit5)
```

```

Call:
lm(formula = selfarchive ~ prestige, data = df)

Residuals:
    Min       1Q   Median       3Q      Max
-2.3114 -0.5331  0.1697  0.4669  1.7263

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.7548     0.2756   6.366 7.48e-09 ***
prestige     0.5189     0.1024   5.065 2.10e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8107 on 92 degrees of freedom
Multiple R-squared:  0.218, Adjusted R-squared:  0.2095
F-statistic: 25.65 on 1 and 92 DF,  p-value: 2.095e-06

```

R-squared beträgt 0.218, sodass das lineare Modell die Varianz in der abhängigen Variablen zu 22% erklären kann.

Überprüfung auf Heteroskedastizität:

Hide

```

# Plot abspeichern
png(file = "h5_hs.png")
plot(fit5, 1)
dev.off()

```

```

null device
      1

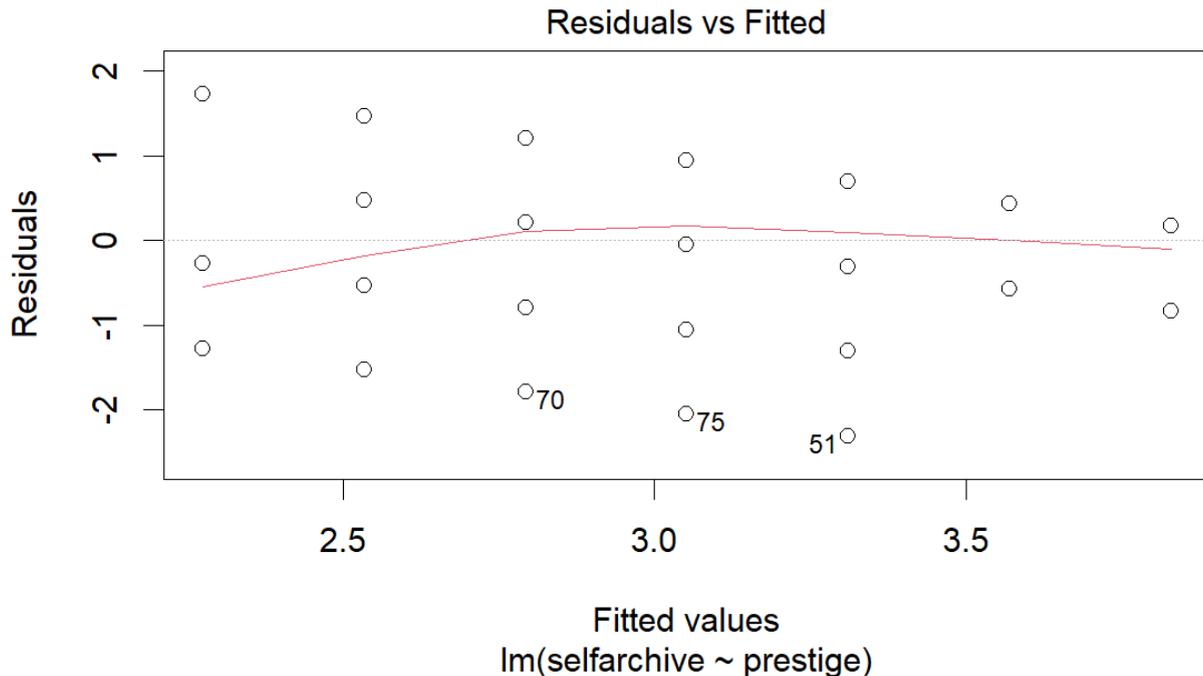
```

Hide

```

h5_hs <- plot(fit5, 1)

```



Die Streuung der Residuen ist nicht linear, was die Berechnung robuster Standardfehler erforderlich macht, um einer Verzerrung des p-Wertes des Koeffizienten vorzubeugen:

Hide

```
coefTest(fit5, vcov = vcovHC(fit5))
```

t test of coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	1.75484	0.32545	5.3921	5.357e-07	***
prestige	0.51886	0.10976	4.7274	8.162e-06	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Der robuste p-Wert beträgt 8.162×10^{-6} (0.000008162) und liegt damit deutlich unter dem Signifikanzniveau von 0.05, sodass die Irrtumswahrscheinlichkeit gering ist. Die Nullhypothese wird verworfen und die Alternativhypothese angenommen.

Der Koeffizient beträgt 0.519, sodass ein positiver Zusammenhang zwischen den Variablen besteht: Wenn das Streben nach persönlicher Anerkennung um eine Einheit steigt, dann steigt die Bereitschaft zur Selbstarchivierung in IR um 0.519.

Visualisierung der Selbstarchivierungsbereitschaft nach Prestigestreben:

Prestigestrebensgruppen erstellen:

Hide

```
df <- df %>% mutate(Prestigestreben = ifelse(prestige <= 2, "geringes Streben", "hohes Streben"))
```

Prestigestrebensgruppen in absoluten Zahlen:

Hide

```
prestige_absolut <- df %>%  
  group_by(Prestigestreben) %>%  
  summarise(Anzahl = n())  
prestige_absolut
```

Prestigestreben	Anzahl
<chr>	<int>
geringes Streben	35
hohes Streben	59

2 rows

Anteile (in %) des Prestigestrebens ermitteln:

Hide

```
prestige_count <- df %>%  
  count(Prestigestreben, selfarchive) %>%  
  group_by(Prestigestreben) %>%  
  mutate(Prozent = n / sum(n)) %>%  
  mutate(Prozent = round(n / sum(n), 3))  
  
prestige_count$Prozent <- (prestige_count$Prozent * 100)
```

Daten für die Visualisierung vorbereiten:

Hide

```
prestige_count <- prestige_count %>%  
  mutate(selfarchive = as.factor(selfarchive)) %>%  
  mutate(selfarchive = str_replace(selfarchive, "1", "gar nicht")) %>%  
  mutate(selfarchive = str_replace(selfarchive, "2", "weniger")) %>%  
  mutate(selfarchive = str_replace(selfarchive, "3", "weitgehend")) %>%  
  mutate(selfarchive = str_replace(selfarchive, "4", "voll und ganz"))  
  
prestige_count$selfarchive <- ordered(prestige_count$selfarchive, levels = c("gar nicht", "weniger", "weitgehend", "voll und ganz"))
```

Gestapeltes Balkendiagramm plotten:

Hide

```

plot_prestige2 <- ggplot(prestige_count, aes(x = Prestigestreben, y = Prozent, fill = selfarchive)) +
  geom_col(width = 0.7) +
  geom_text(aes(label = paste0(Prozent, "%"),
    color = "white",
    position = position_stack(vjust = 0.5)) +
  labs(x = "Prestigestreben", y = "Prozent") +
  scale_fill_hue(l = 65, c = 100)

# Plot abspeichern
ggsave(plot_prestige2, file = "plot_prestige2.png")

```

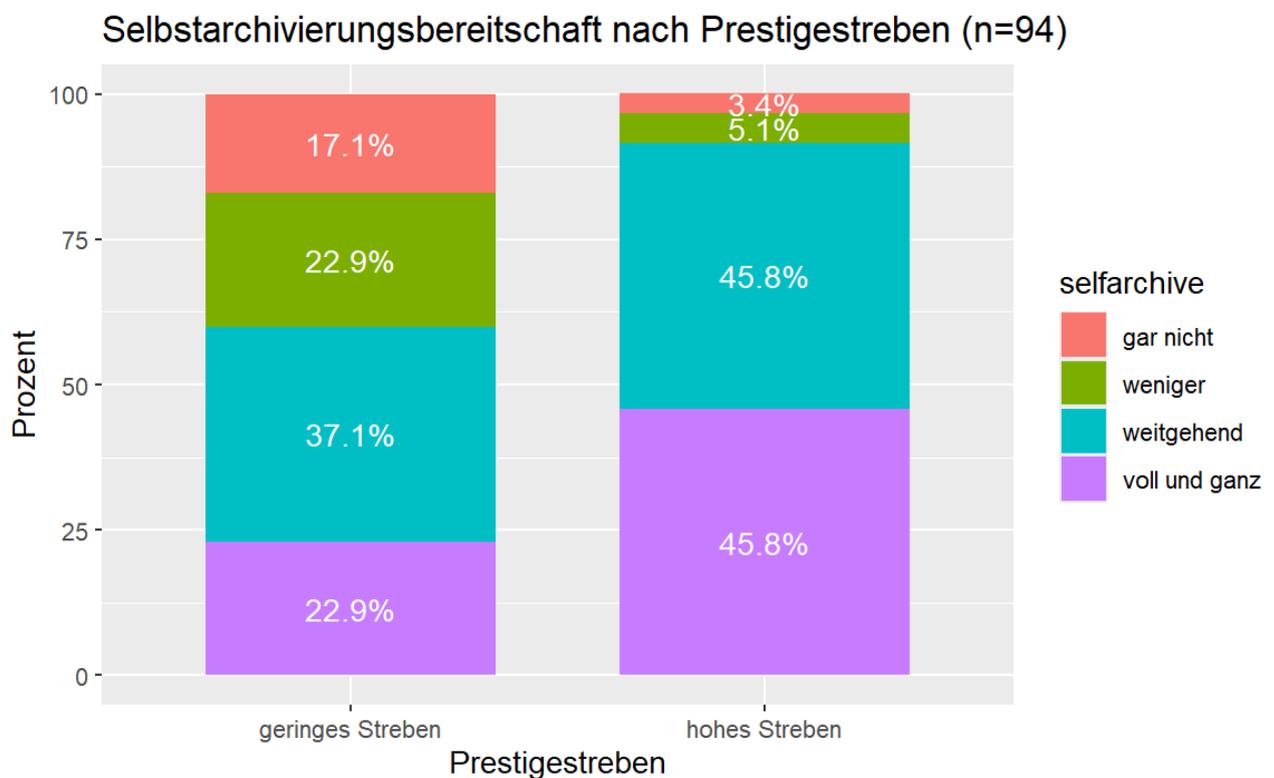
Saving 7 x 7 in image

Hide

```

plot_prestige2 + ggtitle("Selbstarchivierungsbereitschaft nach Prestigestreben (n=94)")

```



(Wegen Rundungsdifferenzen ergibt die Summe der Beschriftungen nicht 100%.)

6.6 H6 (*trustworthiness* → *selfarchive*)

Lineares Modell erstellen:

Hide

```

fit6 <- lm(selfarchive ~ trustworthiness, data = df)

```

Scatterplot rendern:

Hide

```

plot_h6 <- ggplot(df, aes(x = trustworthiness, y = selfarchive)) +
  geom_count() +
  scale_size_area() +
  stat_smooth(method = "lm", col = "blue") +
  labs(title = "(H6) Einfluss von Qualitätssicherung auf Selbstarchivierungsbereitschaf
t", x = "Vertrauen in qualitätssichernde Prozesse", y = "Selbstarchivierungsbereitschaf
t")

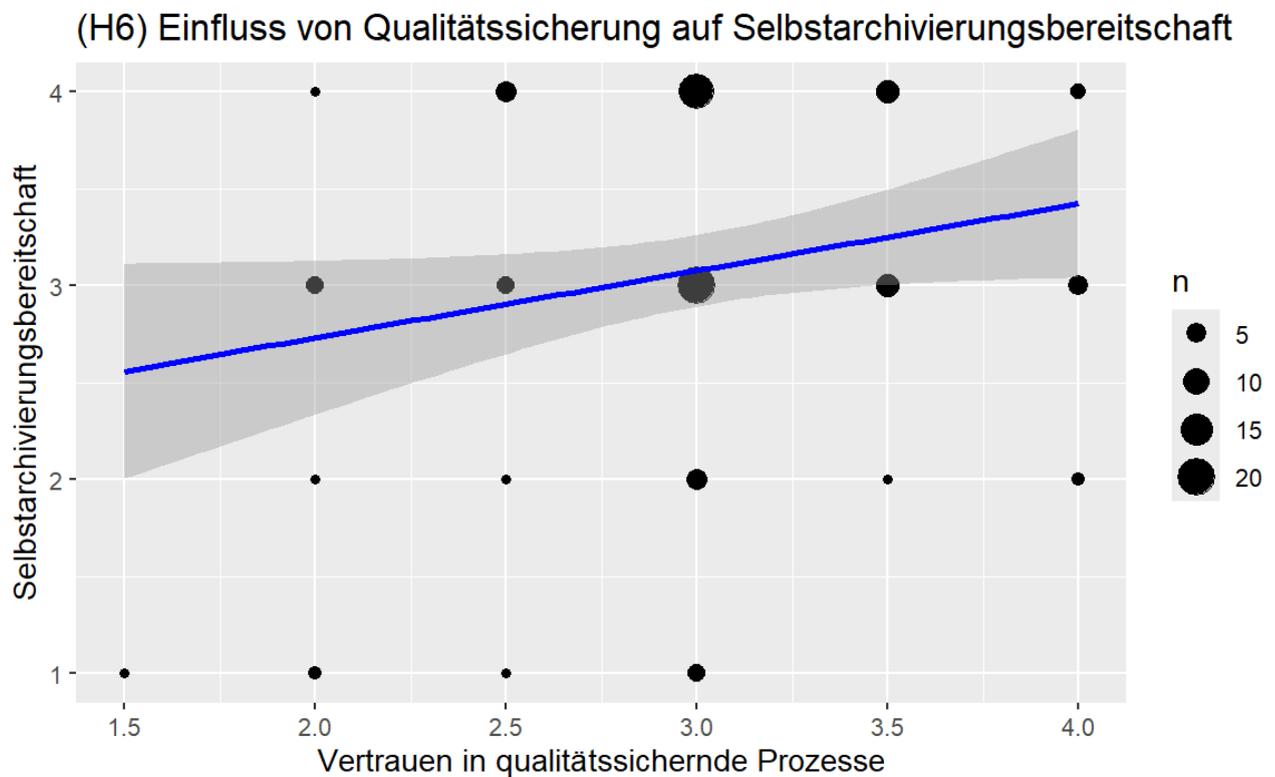
# Plot abspeichern
ggsave(plot_h6, file = "plot_h6.png")

```

Saving 7 x 7 in image

Hide

plot_h6



Summary des linearen Modells:

Hide

```
summary(fit6)
```

```

Call:
lm(formula = selfarchive ~ trustworthiness, data = df)

Residuals:
    Min       1Q   Median       3Q      Max
-2.07774 -0.38057 -0.07774  0.92226  1.26835

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)    2.0395     0.5303   3.846 0.000221 ***
trustworthiness  0.3461     0.1728   2.002 0.048188 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8974 on 92 degrees of freedom
Multiple R-squared:  0.04176,    Adjusted R-squared:  0.03135
F-statistic:  4.01 on 1 and 92 DF,  p-value: 0.04819

```

R-squared beträgt 0.042, sodass das lineare Modell die Varianz in der abhängigen Variablen zu lediglich 4% erklären kann.

Überprüfung auf Heteroskedastizität:

Hide

```

# Plot abspeichern
png(file = "h6_hs.png")
plot(fit6, 1)
dev.off()

```

```

null device
      1

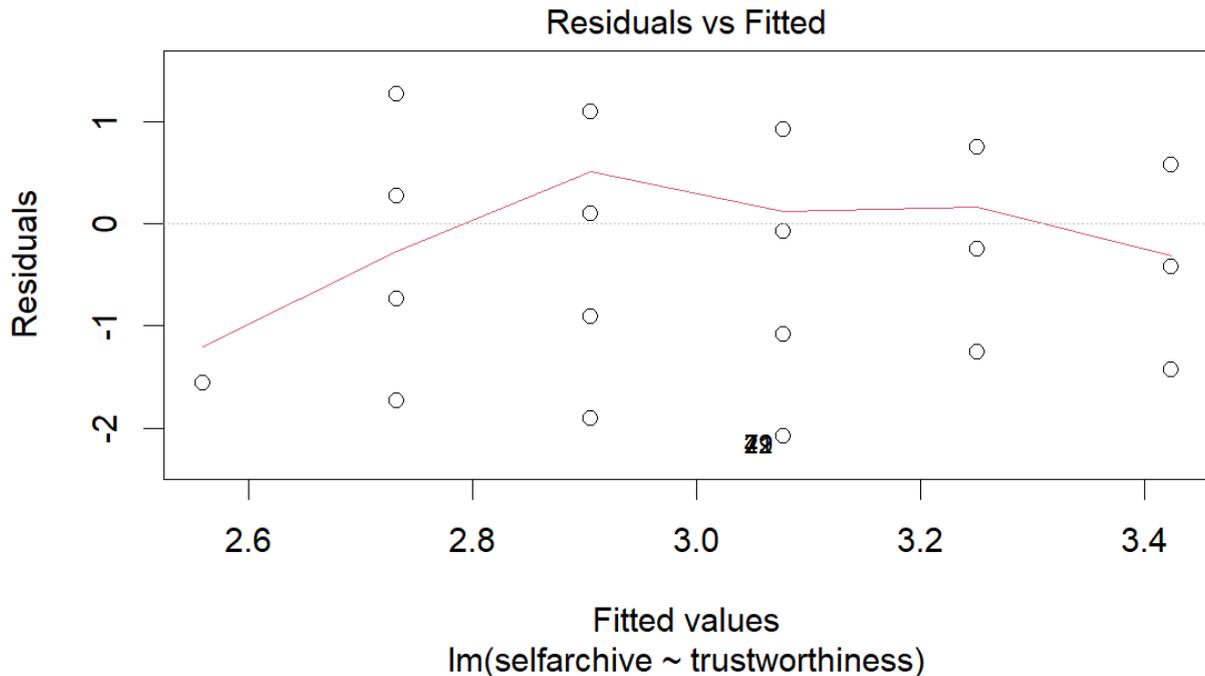
```

Hide

```

h6_hs <- plot(fit6, 1)

```



Die Streuung der Residuen ist nicht linear, was die Berechnung robuster Standardfehler erforderlich macht, um einer Verzerrung des p-Wertes des Koeffizienten vorzubeugen:

Hide

```
coefTest(fit6, vcov = vcovHC(fit6))
```

t test of coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	2.03946	0.61978	3.2906	0.001418	**
trustworthiness	0.34609	0.19495	1.7753	0.079161	.

 Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Der robuste p-Wert beträgt 0.079 und liegt damit über dem Signifikanzniveau von 0.05, sodass die Irrtumswahrscheinlichkeit bei über 5% liegt. Die Nullhypothese kann nicht verworfen und die Interpretation nicht sinnvoll fortgesetzt werden.

6.7 H7a (*career* → *selfarchive*)

Lineares Modell erstellen:

Hide

```
fit7a <- lm(selfarchive ~ career, data = df)
```

Scatterplot rendern:

Hide

```

plot_h7a <- ggplot(df, aes(x = career, y = selfarchive)) +
  geom_count() +
  scale_size_area() +
  stat_smooth(method = "lm", col = "blue") +
  labs(title = "(H7a) Einfluss von Karrierestreben auf Selbstarchivierungsbereitschaft", x = "Karrierestreben im Studium", y = "Selbstarchivierungsbereitschaft")

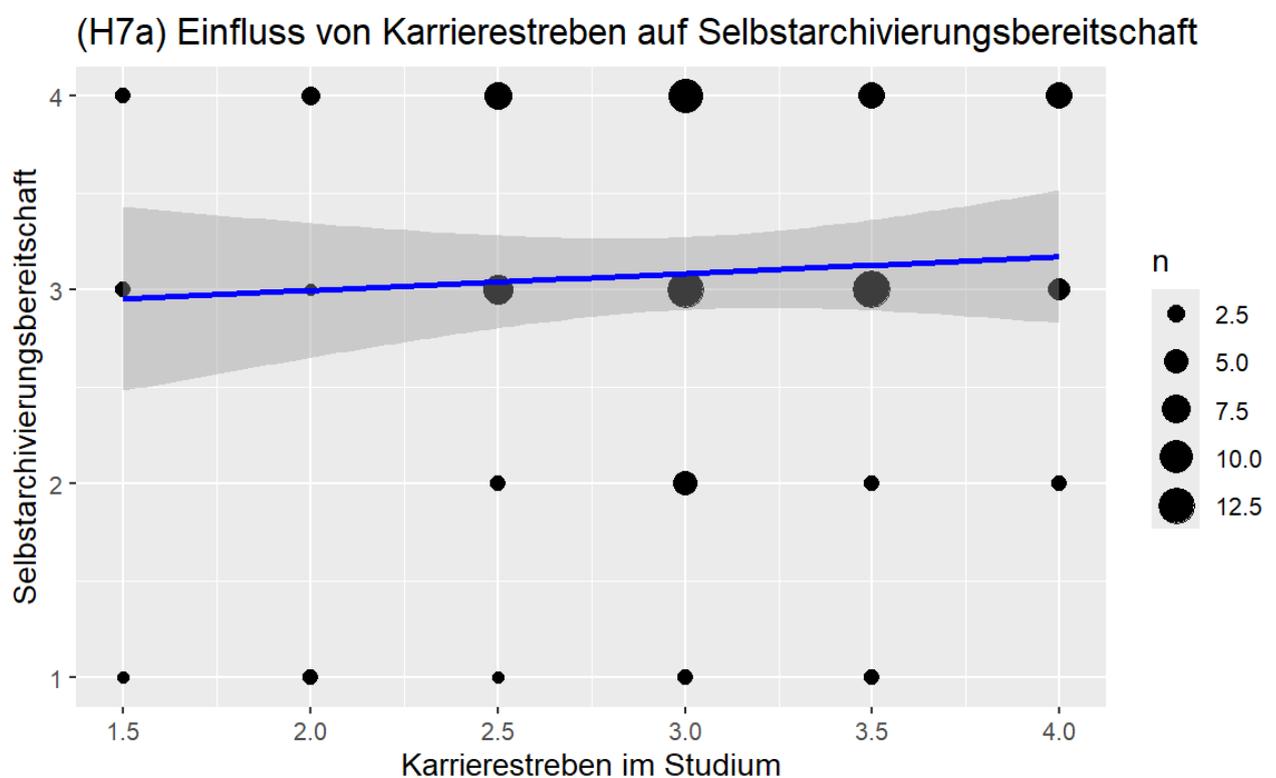
# Plot abspeichern
ggsave(plot_h7a, file = "plot_h7a.png")

```

Saving 7 x 7 in image

Hide

plot_h7a



Summary des linearen Modells:

Hide

```
summary(fit7a)
```

```

Call:
lm(formula = selfarchive ~ career, data = df)

Residuals:
    Min       1Q   Median       3Q      Max
-2.12742 -0.12742 -0.08419  0.87258  1.04553

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.82476    0.44841   6.299 1.01e-08 ***
career       0.08648    0.14561   0.594  0.554
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.915 on 92 degrees of freedom
Multiple R-squared:  0.003819, Adjusted R-squared:  -0.007009
F-statistic: 0.3527 on 1 and 92 DF,  p-value: 0.554

```

R-squared beträgt 0.004, sodass das lineare Modell die Varianz in der abhängigen Variablen zu lediglich 0.4% erklären kann.

Überprüfung auf Heteroskedastizität:

Hide

```

# Plot abspeichern
png(file = "h7a_hs.png")
plot(fit7a, 1)
dev.off()

```

```

null device
      1

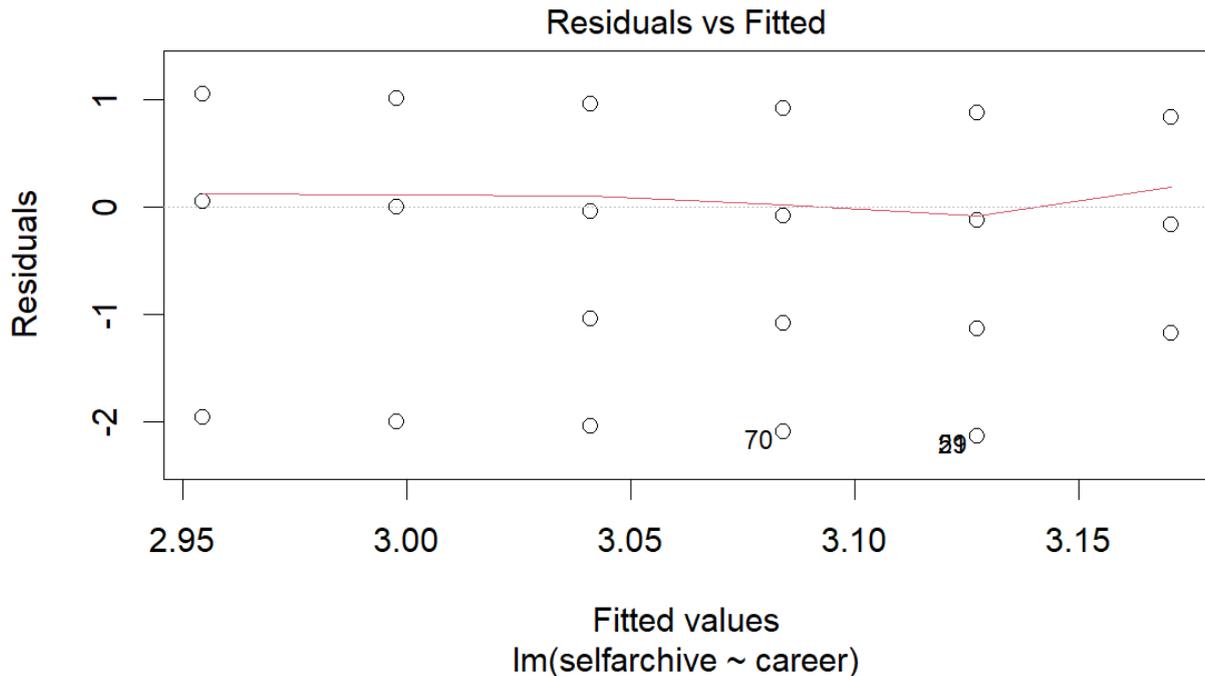
```

Hide

```

h7a_hs <- plot(fit7a, 1)

```



Die Streuung der Residuen ist nicht linear, was die Berechnung robuster Standardfehler erforderlich macht, um einer Verzerrung des p-Wertes des Koeffizienten vorzubeugen:

Hide

```
coefTest(fit7a, vcov = vcovHC(fit7a))
```

t test of coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	2.824758	0.529124	5.3386	6.713e-07	***
career	0.086476	0.165592	0.5222	0.6028	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Der robuste p-Wert beträgt 0.603 und liegt damit deutlich über dem Signifikanzniveau von 0.05, sodass die Irrtumswahrscheinlichkeit sehr hoch ist. Die Nullhypothese kann nicht verworfen und die Interpretation nicht sinnvoll fortgesetzt werden.

6.8 H8 (*identification* → *selfarchive*)

Lineares Modell erstellen:

Hide

```
fit8 <- lm(selfarchive ~ identification, data = df)
```

Scatterplot rendern:

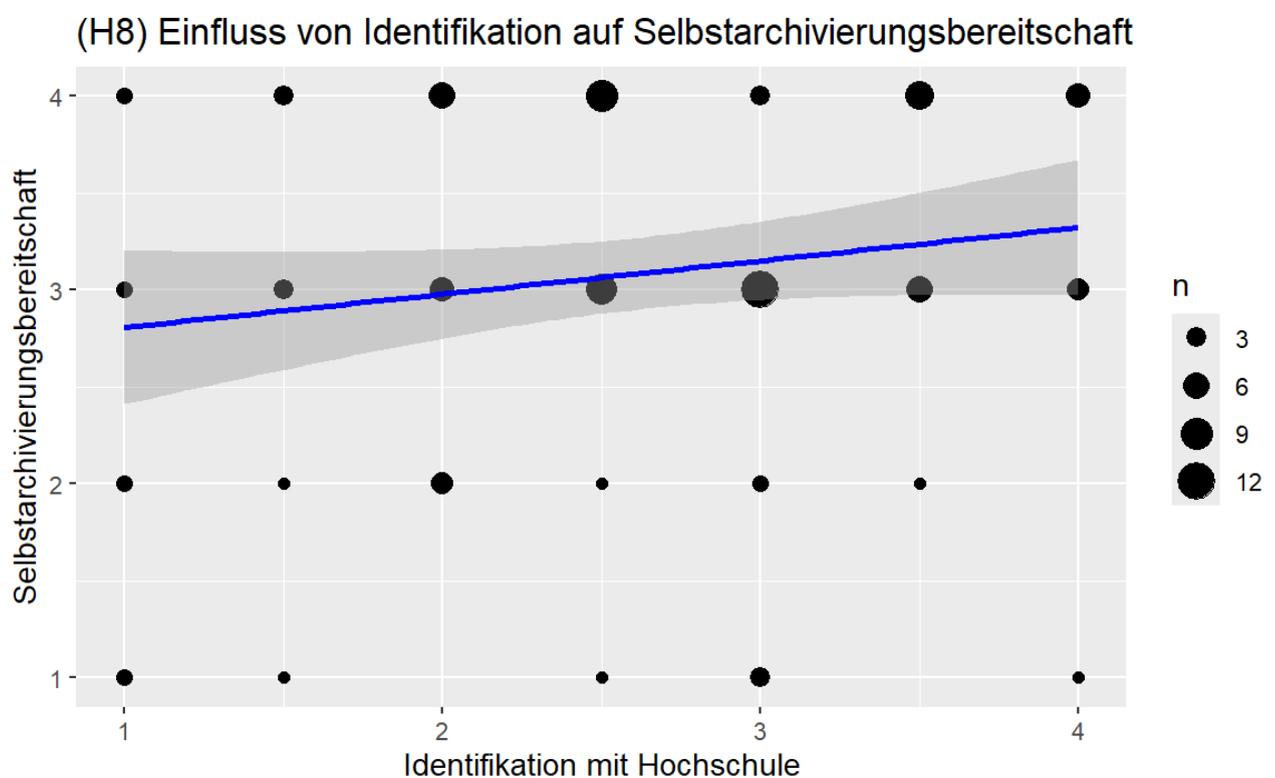
Hide

```
plot_h8 <- ggplot(df, aes(x = identification, y = selfarchive)) +  
  geom_count() +  
  scale_size_area() +  
  stat_smooth(method = "lm", col = "blue") +  
  labs(title = "(H8) Einfluss von Identifikation auf Selbstarchivierungsbereitschaft",  
x = "Identifikation mit Hochschule", y = "Selbstarchivierungsbereitschaft")  
  
# Plot abspeichern  
ggsave(plot_h8, file = "plot_h8.png")
```

Saving 7 x 7 in image

Hide

plot_h8



Summary des linearen Modells:

Hide

```
summary(fit8)
```

```

Call:
lm(formula = selfarchive ~ identification, data = df)

Residuals:
    Min       1Q   Median       3Q      Max
-2.32223 -0.23617 -0.06405  0.76383  1.19413

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   2.6338     0.2987   8.817 7.04e-14 ***
identification 0.1721     0.1082   1.590  0.115
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9044 on 92 degrees of freedom
Multiple R-squared:  0.02676,    Adjusted R-squared:  0.01618
F-statistic:  2.53 on 1 and 92 DF,  p-value: 0.1152

```

R-squared beträgt 0.027, sodass das lineare Modell die Varianz in der abhängigen Variablen lediglich zu 3% erklären kann.

Überprüfung auf Heteroskedastizität:

Hide

```

# Plot abspeichern
png(file = "h8_hs.png")
plot(fit8, 1)
dev.off()

```

```

null device
      1

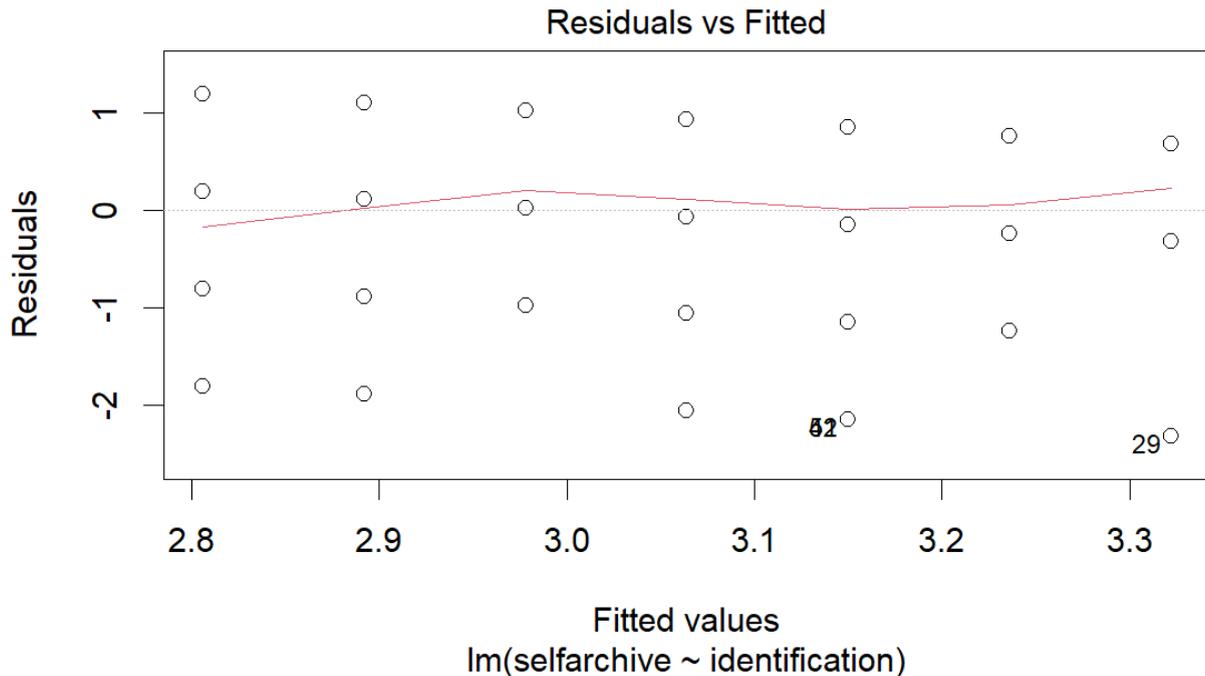
```

Hide

```

h8_hs <- plot(fit8, 1)

```



Die Streuung der Residuen ist nicht linear, was die Berechnung robuster Standardfehler erforderlich macht, um einer Verzerrung des p-Wertes des Koeffizienten vorzubeugen:

Hide

```
coefTest(fit8, vcov = vcovHC(fit8))
```

t test of coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	2.63375	0.34349	7.6676	1.759e-11 ***
identification	0.17212	0.12018	1.4321	0.1555

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Der robuste p-Wert beträgt 0.156 und liegt damit deutlich über dem Signifikanzniveau von 0.05, sodass die Irrtumswahrscheinlichkeit hoch ist. Die Nullhypothese kann nicht verworfen und die Interpretation nicht sinnvoll fortgesetzt werden.

6.9 Zusatzüberprüfung (*altruism3* → *selfarchive*)

Lineares Modell erstellen:

Hide

```
fit9 <- lm(selfarchive ~ altruism3, data = df)
```

Scatterplot rendern:

Hide

```

plot_h9 <- ggplot(df, aes(x = altruism3, y = selfarchive)) +
  geom_count() +
  scale_size_area() +
  stat_smooth(method = "lm", col = "blue") +
  labs(title = "(ZÜ) Einfluss von OA-Spirit auf Selbstarchivierungsbereitschaft", x =
"OA-Spirit", y = "Selbstarchivierungsbereitschaft")

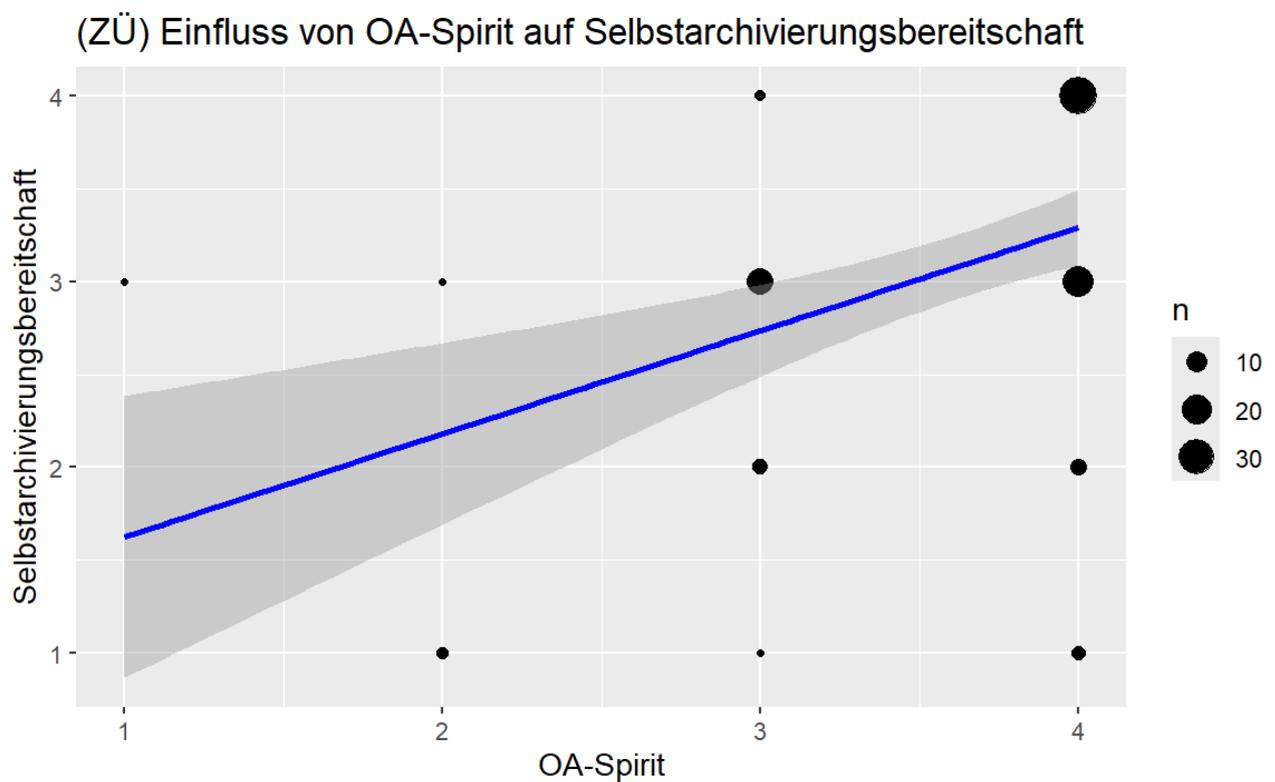
# Plot abspeichern
ggsave(plot_h9, file = "plot_h9.png")

```

Saving 7 x 7 in image

Hide

plot_h9



Summary des linearen Modells:

Hide

```
summary(fit9)
```

```

Call:
lm(formula = selfarchive ~ altruism3, data = df)

Residuals:
    Min       1Q   Median       3Q      Max
-2.2919 -0.2919  0.2635  0.7081  1.3744

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.0701     0.5206   2.056 0.042661 *
altruism3    0.5555     0.1415   3.926 0.000166 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8484 on 92 degrees of freedom
Multiple R-squared:  0.1435,    Adjusted R-squared:  0.1342
F-statistic: 15.42 on 1 and 92 DF,  p-value: 0.0001664

```

R-squared beträgt 0.144, sodass das lineare Modell die Varianz in der abhängigen Variablen zu 14% erklären kann.

Überprüfung auf Heteroskedastizität:

Hide

```

# Plot abspeichern
png(file = "h9_hs.png")
plot(fit9, 1)
dev.off()

```

```

null device
      1

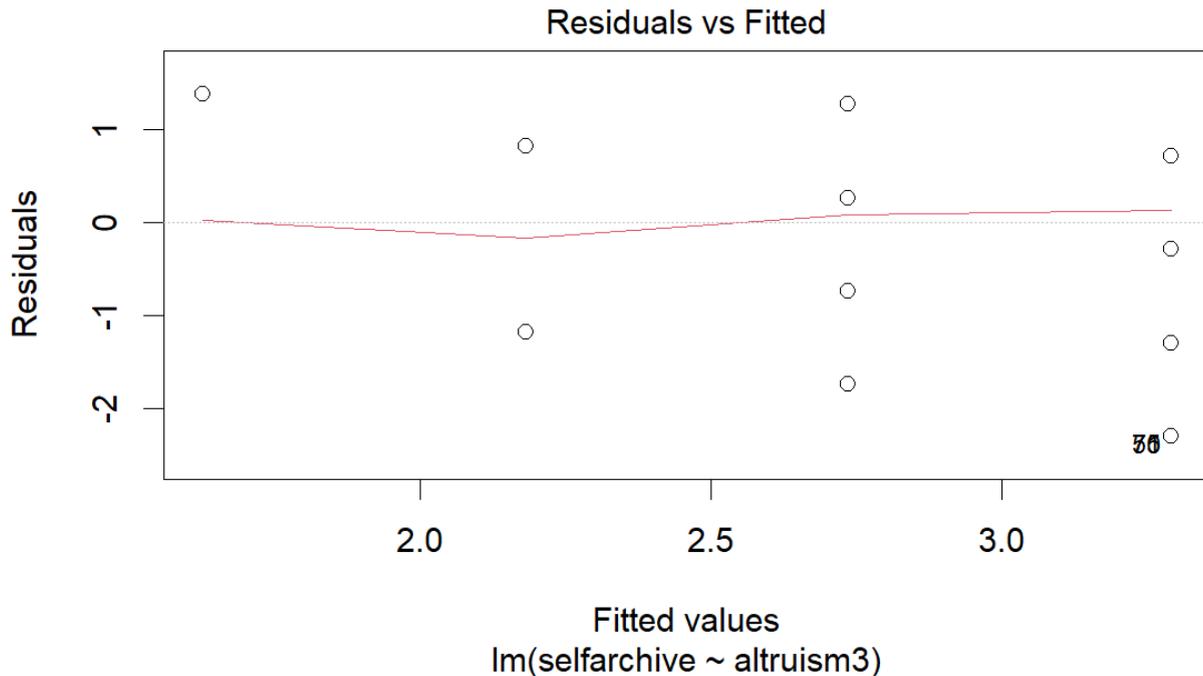
```

Hide

```

h9_hs <- plot(fit9, 1)

```



Die Streuung der Residuen ist nicht linear, was die Berechnung robuster Standardfehler erforderlich macht, um einer Verzerrung des p-Wertes des Koeffizienten vorzubeugen:

Hide

```
coefTest(fit9, vcov = vcovHC(fit9))
```

t test of coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.07010	0.70488	1.5181	0.132411
altruism3	0.55546	0.19068	2.9130	0.004493 **

 Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Der robuste p-Wert beträgt 0.004 und liegt damit unter dem Signifikanzniveau von 0.05, sodass die Irrtumswahrscheinlichkeit gering ist. Die Nullhypothese wird verworfen und die Alternativhypothese angenommen.

Der Koeffizient beträgt 0.555, sodass ein positiver Zusammenhang zwischen den Variablen besteht: Wenn der OA-Spirit um eine Einheit steigt, dann steigt die Bereitschaft zur Selbstarchivierung in IR um 0.555.

Visualisierung der Selbstarchivierungsbereitschaft nach OA-Spirit:

OA-Spirit-Gruppen erstellen:

Hide

```
df <- df %>% mutate(OA_Spirit = ifelse(altruism3 <= 2, "geringer OA-Spirit", "hoher OA-Spirit"))
```

Anteile (in %) des OA-Spirits ermitteln:

Hide

```
spirit_count <- df %>%  
  count(OA_Spirit, selfarchive) %>%  
  group_by(OA_Spirit) %>%  
  mutate(Prozent = n / sum(n)) %>%  
  mutate(Prozent = round(n / sum(n), 3))  
  
spirit_count$Prozent <- (spirit_count$Prozent * 100)
```

Daten für die Visualisierung vorbereiten:

Hide

```
spirit_count <- spirit_count %>%  
  mutate(selfarchive = as.factor(selfarchive)) %>%  
  mutate(selfarchive = str_replace(selfarchive, "1", "gar nicht")) %>%  
  mutate(selfarchive = str_replace(selfarchive, "2", "weniger")) %>%  
  mutate(selfarchive = str_replace(selfarchive, "3", "weitgehend")) %>%  
  mutate(selfarchive = str_replace(selfarchive, "4", "voll und ganz"))  
  
spirit_count$selfarchive <- ordered(spirit_count$selfarchive, levels = c("gar nicht",  
"weniger", "weitgehend", "voll und ganz"))
```

Gestapeltes Balkendiagramm plotten:

Hide

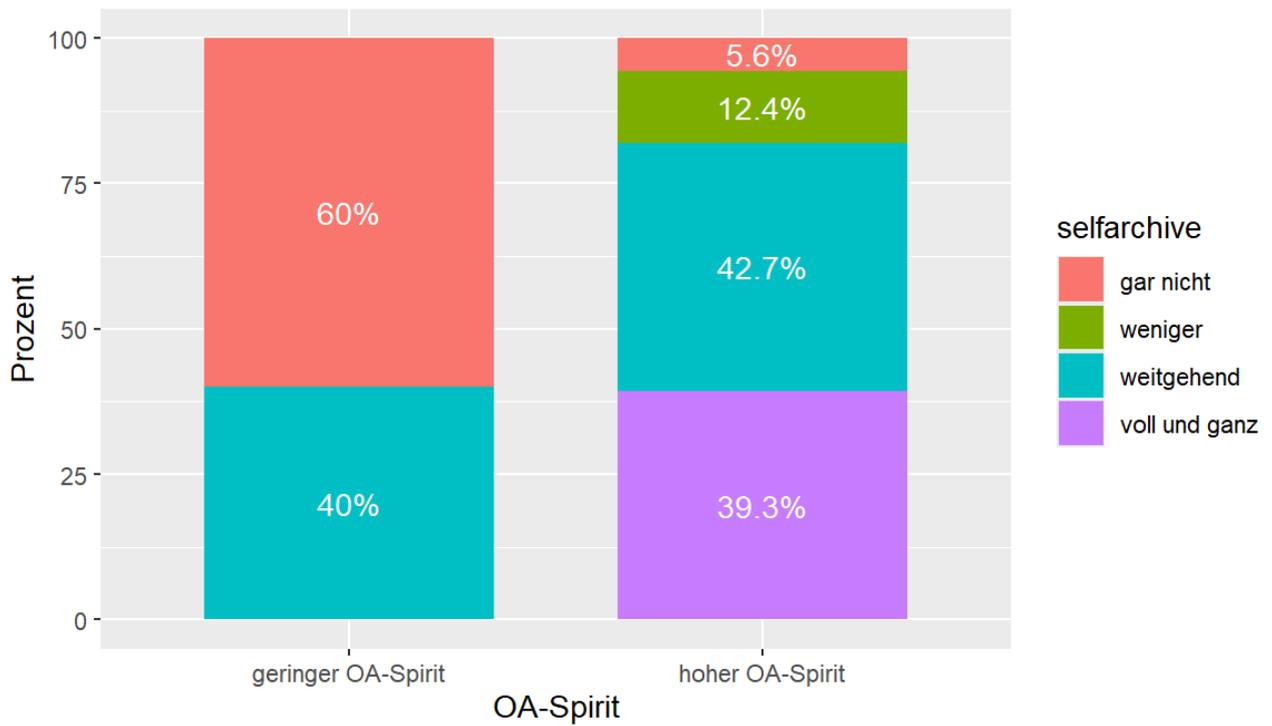
```
plot_spirit <- ggplot(spirit_count, aes(x = OA_Spirit, y = Prozent, fill = selfarchive)) +  
  geom_col(width = 0.7) +  
  geom_text(aes(label = paste0(Prozent, "%")),  
            color = "white",  
            position = position_stack(vjust = 0.5)) +  
  labs(x = "OA-Spirit", y = "Prozent") +  
  scale_fill_hue(l = 65, c = 100)  
  
# Plot abspeichern  
ggsave(plot_spirit, file = "plot_spirit.png")
```

Saving 7 x 7 in image

Hide

```
plot_spirit + ggtitle("Selbstarchivierungsbereitschaft nach OA-Spirit (n=94)")
```

Selbstarchivierungsbereitschaft nach OA-Spirit (n=94)



Prestigestreben bei OA-Spirit == 3:

Hide

```
spirit_prestige <- df[df$selfarchive == 3 & df$altruism3 < 3,] %>%
  group_by(Prestigestreben) %>%
  summarise(Anzahl = n()) %>%
  rename("Prestigestreben bei OA-Spirit == 3" = "Prestigestreben")
spirit_prestige
```

Prestigestreben bei OA-Spirit == 3

Anzahl

<chr>

<int>

geringes Streben

1

hohes Streben

1

2 rows

7 H7b: Moderationsanalyse (*career x acad_career* → *selfarchive*)

Lineares Modell erstellen:

Hide

```
fit7b <- lm(selfarchive ~ career * acad_career, data = df)
```

Summary des linearen Modells:

Hide

```
summary(fit7b)
```

Call:

```
lm(formula = selfarchive ~ career * acad_career, data = df)
```

Residuals:

```
      Min       1Q   Median       3Q      Max
-2.27945 -0.30047 -0.04811  0.84630  1.07046
```

Coefficients:

```
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    3.00730    0.93693   3.210  0.00184 **
career         -0.07621    0.31935  -0.239  0.81193
acad_career    -0.06759    0.43167  -0.157  0.87594
career:acad_career 0.07330    0.14610   0.502  0.61709
```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9119 on 90 degrees of freedom

Multiple R-squared: 0.032, Adjusted R-squared: -0.0002643

F-statistic: 0.9918 on 3 and 90 DF, p-value: 0.4004

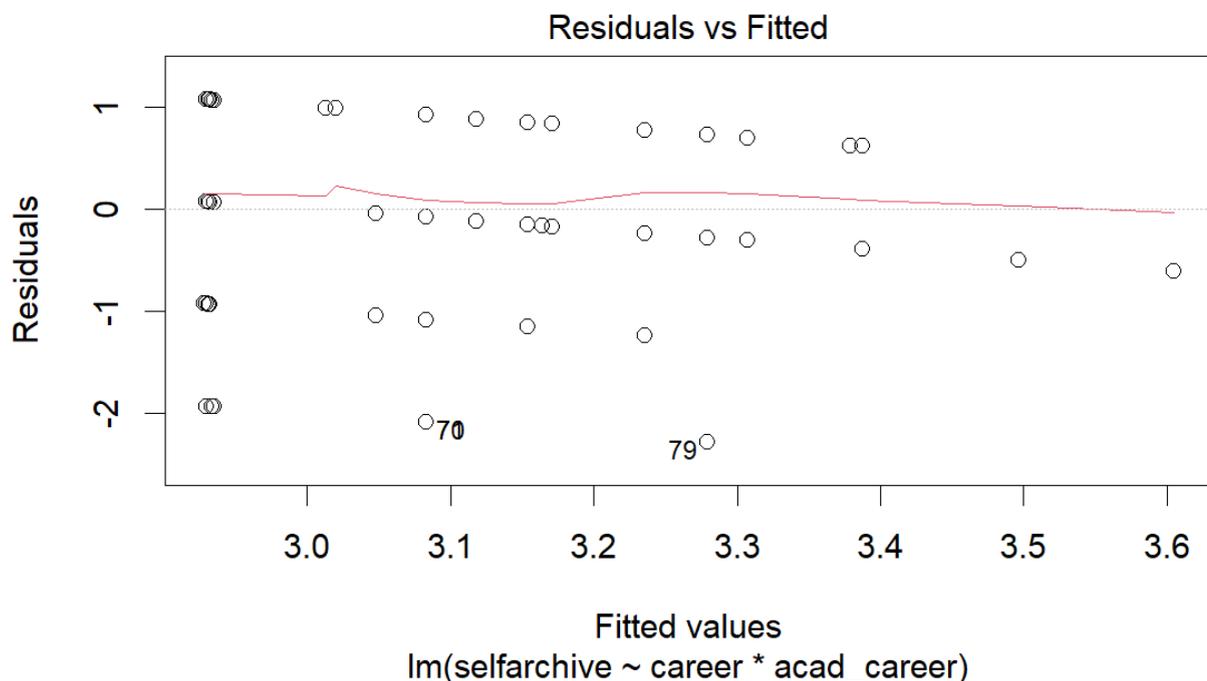
Der p-Wert des Modells beträgt 0.400 und liegt damit deutlich über dem Signifikanzniveau von 0.05, sodass die Irrtumswahrscheinlichkeit sehr hoch ist: Das Modell leistet keinen Beitrag zur Erklärung der abhängigen Variablen.

R-squared beträgt 0.032, sodass das lineare Modell die Varianz in der abhängigen Variablen zu 3% erklären kann.

Überprüfung auf Heteroskedastizität:

Hide

```
h7b_hs <- plot(fit7b, 1)
```



Die Streuung der Residuen ist nicht linear, was die Berechnung robuster Standardfehler erforderlich macht, um einer Verzerrung der p-Werte der Koeffizienten vorzubeugen:

Hide

```
coefTest(fit7b, vcov = vcovHC(fit7b))
```

t test of coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	3.007304	1.243096	2.4192	0.01757 *
career	-0.076209	0.405866	-0.1878	0.85148
acad_career	-0.067587	0.508729	-0.1329	0.89461
career:acad_career	0.073301	0.164092	0.4467	0.65616

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

Sowohl der p-Wert des Interaktionseffekts (0.656) als auch die p-Werte der Haupteffekte (0.851 bzw. 0.895) liegen deutlich über dem Signifikanzniveau von 0.05, sodass die Irrtumswahrscheinlichkeit hoch ist. Die Nullhypothese kann nicht verworfen und die Interpretation nicht sinnvoll fortgesetzt werden.

8 Überprüfung des Gesamtmodells mittels multipler linearer Regression

Obwohl im ursprünglichen Modell keine direkte lineare Beziehung zwischen *altruism3* bzw. OA-Spirit und *selfarchive* angenommen wurde, wird *altruism3* als unabhängige Variable in die multiple Regression aufgenommen. (Weil das Item im Zuge der Reliabilitätsanalyse in Kapitel 3.8 aus der Altruismusskala entfernt wurde, tritt keine Dopplung auf.)

Modell berechnen:

Hide

```
fit10 <- lm(selfarchive ~ time_effort + plagiarism + altruism + diligence + prestige + trustworthiness + career * acad_career + identification + altruism3, data = df)
```

8.1 Überprüfung auf Multikollinearität

Um sicherzustellen, dass keine starken Korrelationen zwischen den unabhängigen Variablen bestehen, wird das Modell in Hinsicht auf Multikollinearität überprüft.

Generalisierte Varianzinflationsfaktoren berechnen:

Hide

```
vif(fit10, type = "predictor")
```

GVIFs computed for predictors

	GVIF <dbl>	Df <dbl>	GVIF^(1/(2*Df)) <dbl>	Interacts With <chr>
time_effort	1.233259	1	1.110522	--

	GVIF <dbl>	Df <dbl>	GVIF^{1/(2*Df)} <dbl>	Interacts With <chr>
plagiarism	1.230841	1	1.109433	--
altruism	1.288062	1	1.134928	--
diligence	1.495101	1	1.222743	--
prestige	1.208580	1	1.099354	--
trustworthiness	1.603029	1	1.266108	--
career	1.696595	3	1.092102	acad_career
acad_career	1.696595	3	1.092102	career
identification	1.647494	1	1.283547	--
altruism3	1.239625	1	1.113385	--

1-10 of 10 rows | 1-5 of 5 columns

Bei allen unabhängigen Variablen ist $GVIF^{1/(2 \cdot Df)} < 2$, sodass keine für die multiple Regression problematische Multikollinearität vorliegt.

8.2 Summary und Heteroskedastizität

Summary des Modells:

Hide

```
summary(fit10)
```

Call:

```
lm(formula = selfarchive ~ time_effort + plagiarism + altruism +  
  diligence + prestige + trustworthiness + career * acad_career +  
  identification + altruism3, data = df)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-2.12424	-0.36445	0.01268	0.51451	1.65960

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	0.668480	1.277947	0.523	0.602323	
time_effort	-0.370828	0.111536	-3.325	0.001325	**
plagiarism	-0.109697	0.101254	-1.083	0.281816	
altruism	0.073330	0.172552	0.425	0.671970	
diligence	0.025068	0.111212	0.225	0.822223	
prestige	0.406120	0.102000	3.982	0.000147	***
trustworthiness	-0.008622	0.179035	-0.048	0.961709	
career	0.051179	0.277799	0.184	0.854289	
acad_career	0.401177	0.367379	1.092	0.278031	
identification	0.071467	0.112761	0.634	0.527979	
altruism3	0.385398	0.136299	2.828	0.005892	**
career:acad_career	-0.094142	0.122737	-0.767	0.445273	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7342 on 82 degrees of freedom

Multiple R-squared: 0.4283, Adjusted R-squared: 0.3516

F-statistic: 5.586 on 11 and 82 DF, p-value: 1.303e-06

Der p-Wert des Modells beträgt 1.303e-06 (0,000001303) und liegt damit weit unter dem Signifikanzniveau von 0.05, sodass die Irrtumswahrscheinlichkeit sehr gering ist. Die Nullhypothese wird verworfen und die Alternativhypothese angenommen: Das Modell leistet einen Beitrag zur Erklärung der abhängigen Variablen; die Prädiktoren bilden linear auf die abhängige Variable ab.

R-squared beträgt 0.428, sodass das Gesamtmodell die Varianz in der abhängigen Variablen zu 43% erklären kann.

Überprüfung auf Heteroskedastizität:

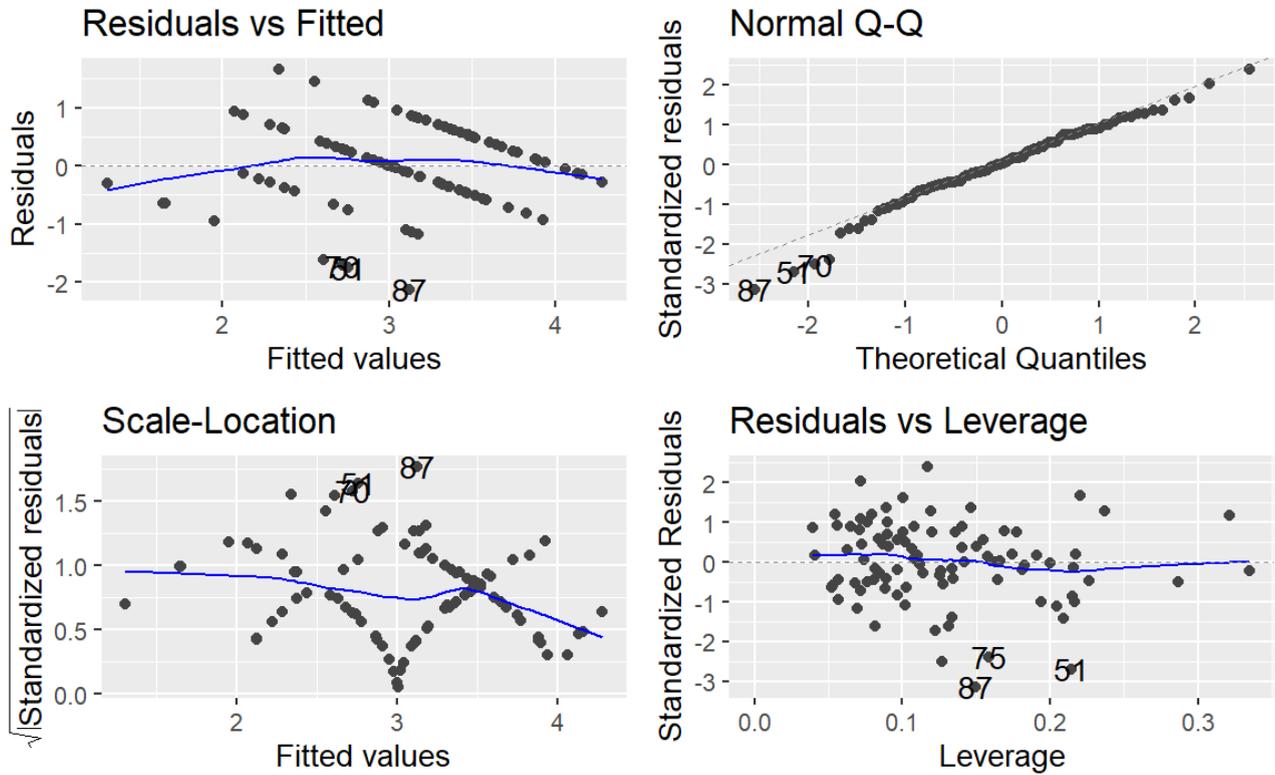
Hide

```
# Plot abspeichern  
png(file = "lm_hs.png")  
plot(fit10, 1)  
dev.off()
```

```
null device  
1
```

Hide

```
autoplot(fit10)
```



Die Streuung der Residuen ist nicht linear, was die Berechnung robuster Standardfehler erforderlich macht, um einer Verzerrung der p-Werte der Koeffizienten vorzubeugen:

Hide

```
coefTest(fit10, vcov = vcovHC(fit10))
```

t test of coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.6684802	1.7268815	0.3871	0.69968
time_effort	-0.3708283	0.1500500	-2.4714	0.01553 *
plagiarism	-0.1096967	0.0863477	-1.2704	0.20753
altruism	0.0733297	0.2056464	0.3566	0.72232
diligence	0.0250682	0.1306963	0.1918	0.84837
prestige	0.4061197	0.0970380	4.1852	7.116e-05 ***
trustworthiness	-0.0086216	0.2229941	-0.0387	0.96925
career	0.0511786	0.3515144	0.1456	0.88460
acad_career	0.4011774	0.4232347	0.9479	0.34597
identification	0.0714670	0.1174775	0.6083	0.54464
altruism3	0.3853979	0.1567983	2.4579	0.01608 *
career:acad_career	-0.0941418	0.1422418	-0.6618	0.50993

 Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

8.3 Interpretation signifikanter Koeffizienten

Weil alle Variablen einen Wertebereich von 1-4 haben, ist keine Z-Standardisierung (Mittelwert von 0, Standardabweichung von 1) erforderlich; die Estimates sind bereits vergleichbar.

prestige: Der robuste p-Wert beträgt 7.116e-05 (0,00007116). Der Koeffizient beträgt 0.406, sodass ein positiver Zusammenhang zwischen den Variablen besteht: Wenn das Streben nach persönlicher

Anerkennung um eine Einheit steigt, dann steigt die Selbstarchivierungsbereitschaft um 0.406. Die Variable übt den größten Einfluss auf die AV aus.

altruism3: Der robuste p-Wert beträgt 0.016. Der Koeffizient beträgt 0.385, sodass ein positiver Zusammenhang zwischen den Variablen besteht: Wenn der OA-Spirit um eine Einheit steigt, dann steigt die Selbstarchivierungsbereitschaft um 0.385. Die Variable übt den zweitgrößten Einfluss auf die AV aus.

time_effort: Der robuste p-Wert beträgt 0.016. Der Koeffizient beträgt -0.371, sodass ein negativer Zusammenhang zwischen den Variablen besteht: Wenn die wahrgenommene Zeit- und Aufwandsinvestition um eine Einheit steigt, dann sinkt die Selbstarchivierungsbereitschaft um 0.371. Die Variable übt den drittgrößten Einfluss auf die AV aus.

Eidesstattliche Erklärung

Ich versichere, die vorliegende Arbeit selbstständig ohne fremde Hilfe verfasst und keine anderen Quellen und Hilfsmittel als die angegebenen benutzt zu haben. Die aus anderen Werken wörtlich entnommenen Stellen oder dem Sinn nach entlehnten Passagen sind durch Quellenangabe kenntlich gemacht.

█, 05.05.2025

Ort, Datum

█

Unterschrift