

4 Ergebnisse

Dieses Kapitel beschäftigt sich mit der Präsentation der Ergebnisse der Vor- und Hauptstudie. Der erste Teil dieses Kapitels befasst sich mit dem Abschnitt der deskriptiven Statistik, in welchem die Stichprobenbeschreibung durchgeführt wird. Der zweite Teil beschreibt die Ergebnisse der Vorstudie, welche aus der thematischen Analyse hervorgingen und neben dem OTPP-Layout des BAföG-Formular als Befunde erschlossen werden konnten. Der dritten Teil befasst sich mit den Ergebnissen der Hypothesentestungen der Hauptstudie, welche zur Beantwortung der Forschungsfrage relevant sind.

4.1 Stichproben

Insgesamt nahmen 9 Versuchspersonen in 3 iterativen Durchgängen an dieser Vorstudie teil, da danach das nominelle Usability Goal erreicht wurde. Die Versuchspersonen teilten sich in 5 weiblich und 4 männlich Proband:innen. Bei 7 der 9 Versuchspersonen war Deutsch eine der Muttersprachen und alle gaben an, die Aussagen des Formulars verstanden zu haben (mit Ausnahmen einiger Fachbegriffe). Alle Versuchspersonen waren vertraut mit dem Versuchsleiter und wurden aufgrund dessen kontaktiert. Das Alter der Teilnehmer lag im Mittel bei 27,3 Jahren und reichte von 24 bis 30 Jahre. Alle gaben an in der Vergangenheit regelmäßig Formulare ausgefüllt zu haben, sowie mindestens einmal selbst BAföG beantragt zu haben. Lediglich 2 Teilnehmer:innen waren mit dem Begriff des deutschen eGovernment vertraut.

Insgesamt begonnen 77 Personen die Hauptstudie aus freien Stücken, wobei insgesamt 47 der Personen diese auch beendeten. Die Studienteilnehmenden wurden über das Versuchspersonenportal „Sona“ der TU Berlin, über einen Aufruf des StudierendenWERK Berlin und aus dem direkten Bekanntenkreis des Masteranden akquiriert. Von den 47

Teilnehmenden, die die Studie beendeten, mussten weitere zwei Personen ausgeschlossen werden, da die Teilnahmevoraussetzungen, zum Zielpersonenkreis des BAföG-Formulars zu gehören, nicht erfüllten wurden mit 62, respektive 42 Jahren. Eine weitere Teilnehmende musste ausgeschlossen werden, da sie entgegen der Anweisung das Experiment auf einem Handy ausführte. Die übrigen Teilnehmer:innen erfüllten alle relevanten Teilnahmevoraussetzungen und bestätigten die Formulare aufmerksam und zu Ihrer Zufriedenheit ausgefüllt zu haben. Von den 44 verbleibenden Teilnehmenden gaben 21 an sich dem männlich Geschlecht, 22 sich dem weiblichen Geschlecht und 1 Personen sich dem intersexuellen Geschlecht zuzuordnen. Das Alter der Teilnehmer:innen lag im Mittel bei $M = 27,23$ ($SD = 3,26$) Jahren und ist normalverteilt, entgegen der sonstigen Erwartung an das Alter von Student:innen. Im Schnitt gaben die Teilnehmenden an 2,8-mal ($SD = 2,43$) ein BAföG-Formular ausgefüllt zu haben, was auf eine eher gewisse Vorerfahrung in der Population schließen lässt. Von allen Teilnehmenden gaben 42 Personen an Deutsch-Muttersprachler:in zu sein bzw. 2 gaben an ein Deutschsprachniveau von C2 aufzuweisen. Außerdem gaben alle Proband:innen an zu dem Zielpersonenkreis des BAföG-Formulars zugehören oder vor kurzem noch gehört zu haben. Damit implizierten sie mindestens einen Hochschulabschluss zu besitzen oder angestrebt zu haben.

4.2 Ergebnisse der Vorstudie

In dieser Vorstudie ist ein hauptsächlich positives Meinungsbild der Proband:innen entstanden bezüglich der Reduzierung der Seiteninhalte und der dadurch entstandenen fein aufgegliederte Form des Formulars im Sinn des OTPP-Layouts. Positive Seiten, die von den Testpersonen unaufgefordert genannt wurden, waren, dass dieser Formulartyp „wesentlich überschaubarer und kompakter wirk[e]“, da „man jede einzelne Frage im Blick hat, ohne dass zig Kästchen [...] zu viel von einem verlangen“. Ebenso wie, dass „der Ausfüllprozess

entspannter [sei] als sonst“ und „nie das Gefühl aufkommt zu viel auf einmal vor sich zu haben“. Andere Formulartypen „mit großen Formularblättern schrecken ab“ oder wirken sogar stressreizend, was sich dann in einem Gefühl von Überforderung oder Ärger niederschlagen kann nach Aussage der Teilnehmenden. Wohingegen „dieses [OTPP artige] Formular angenehm wirk[e]“ und es „sich gut ausfüllen lässt“. Da man sich nach Aussage mehrerer Testpersonen durch die kurzen Formularseiten „geleitet fühlt“. Auch deshalb, aber nicht ausschließlich deshalb, äußerten zwei Proband:innen, dass die „feingegliederte Form des Formulars [...] motivierend“ auf sie wirke. Lediglich eine Testperson äußerte sich wenig positiv über das BAföG-Formular im OTPP-Layout, indem sie sagte, dass es „nervig [sei] andauernd scrollen zu müssen“, um auf die nächste Seite zu kommen und damit die nächsten Formularfelder zu offenbaren.

Der themenspezifische Schwerpunkt dieser Vorstudie, nämlich die Änderungen der Seitenaufteilung des Formulars, wurde ebenfalls mit einigen Kommentaren seitens der Proband:innen bedacht. Regelmäßigkeiten, die sich aus den Äußerungen der Testpersonen ableiten lassen, sind zum einen, dass personenbezogene, einfach abrufbare Informationen gerne zusammen auf derselben Seite stehen können. Dies wurde oft gesagt in Bezug auf Formularfelder wie „Straße und Hausnummer“, „Ort und Postleitzahl“ oder „Vor- und Nachname“. Aber auch in Bezug auf Informationen der Eltern wie eben deren Name, Geschlecht und Geburtsdatum. Auch Formularfelder, die für eine korrekte Eingabe miteinander verglichen werden müssen, sollten auf demselben Formularblatt stehen. Dies war bspw. der Fall bei der Einkommensfeststellung oder bei den Vermögensgegenstände (vgl. Anhang A3). Allerdings äußerten sich einige Testpersonen auch gegenteilig in Bezug auf eben diese Kategorien. Indem sie forderten, dass Fragen, welche nachgeschlagen werden müssen (wie bspw. die genau Aufteilung der Vermögenssituations) doch „lieber getrennt voneinander“ stehen sollen, um sodann fokussiert beantwortet zu werden. In anderen Beispielen, wie bei der Steueridentifikationsnummer, die stets nachgeschlagen werden muss, war es eindeutiger, dass

diese allein auf einem Formularblatt abgefragt werden sollte. Ebenfalls gaben verschiedene Proband:innen zu verstehen, dass Abfragen von Schlagwörtern ohne direkte inhaltliche Ergänzung, wie bspw. die Staatsangehörigkeit, alleinstehen sollten.

Äußerungen, die über die reine Formatebene hinaus gehen, aber ebenfalls sehr hilfreich waren für weitere Formularerstellungszwecke, waren zum einen, dass mehrere Personen erkannten, dass Branching in dieser fein aufgeteilten Form gut umzusetzen war und somit nicht alle Fragen überhaupt gesehen werden müssen von allen Formularausfüllenden. Dies wurde klar am Beispiel der Angaben zu den Eltern, aber auch mehrfach bei der Einkommensfeststellung. In zuletzt genannter BAföG-Formulkategorie wurde u.a. ausgeführt, dass „man die Kategorien vorher auswählen kann und dann nur diese erscheinen“. Ferner sind beispielhafte Vorgaben („Muster“), wie sie bei den Angaben des Lebenslauf verwendet werden, „sehr hilfreich“ und können das Gefühl von Sicherheit im Ausfüllprozess geben. Auch wurde mehrfach geäußert, dass Zwischenspeicherfunktion bei derart langen Anträgen „wichtig und gut“ sei, sodass man zur exakten Stelle im Formular zurückkehren kann, falls Fragen im Verlauf des Ausfüllprozess aufkommen, die mit Dritten besprochen werden müssen. Für lange Formulare mit einer dem OTPP-Prinzip entsprechende feingliedrige Aufteilung sollen außerdem Fortschrittsbalken wie auch Seitenzahlangaben weggelassen werden, da diese derart kleinschrittig sind, dass sie „nicht motivierend [wirken]“ nach Aussage einiger Teilnehmender. Stattdessen soll lieber eine Gliederung des Antrags gegeben sein, „worauf die Überthemen [des Formulars] genannt werden“ oder auch eine „Checkliste, die aufzeigt, welche Unterlagen gebraucht werden“ zum Ausfüllen dieser Formularteile.

Um abschließend eine gewisse Form der Quantifizierung zu geben, soll erwähnt sein, dass von ursprünglich 136 Formularseiten, noch 93 Seiten im finalen OTPP-Formular übriggeblieben sind, wobei 36 Fragen weiterhin einzelnen auf jeweils einer Seiten stehen. Ebenfalls ist interessant, dass es nicht dazu gekommen ist, dass mehr als 4 Fragen auf einer Seite zusammenstehen. In der ersten Iterationsrunde wurde das feingliedrige Design zwar schon

als gut bezeichnet, aber auch kam der Wunsch auf wie oben genannt die einfach abrufbare Informationen wie Name und Vorname, Straße und Hausnummer, Ort und Postleitzahl, Universität und Ort, Anfangs- und Enddatum einer Station im Lebenslauf zusammenzustellen. Außerdem wurde die fehlende Abgleichbarkeit bei den Daten des Einkommens oder des Vermögens kritisiert, woraufhin diese ebenfalls in der nächsten Iterationsrunde zusammengezogen wurden. In der zweiten Iterationsrunde kamen mehr positive Kommentare, wie dass es motivierend, übersichtlicher und „allgemein besser wirk[e] wie der herkömmliche BAföG Antrag“. Allerdings wurden die großen, zusammengezogenen Seiten des Einkommens oder des Vermögens kritisiert. Diesmal, da die Seiten viel zu groß wirken, was durch Kontrast zu den anderen sehr kurzen Seiten verstärkt wurde. Auf Grundlage dessen wurden diese BAföG-Formularkategorien wieder etwas entschlackt und nach Unterkategorien zusammengezogen. Dabei waren Unterkategorien des Vermögens bspw. Sparguthaben, Sachwerte, Eigentum oder Schulden. Außerdem gab es in dieser zweiten Iterationsrunde eine Testperson, die eher kritisch dem feingliedrigen Design gegenüberstand, wie zuvor erwähnt. In der dritten Iterationsrunde kamen sehr viele positive Kommentare, wie dass „die Aufteilung nie zu viel [sei]“ und sehr „verständlich“ wirke. Außerdem dass es „entspannt [sei] durchzuklicken“ dennoch „kompakt“ erscheint und es sich dadurch „nicht so lang wie sonst“ anfühlt. Nicht zuletzt wurde auch von zwei der drei Teilnehmenden geäußert, dass das Formular in der bestehenden Form ideal aufgeteilt ist, und auch die dritte Person äußerte sich positiv über das OTPP-Formular. Besonders in dieser letzten Iterationsrunde sind die Verschiedenheiten zwischen den ausfüllenden Personen ausgefallen, wie bspw. deren Bearbeitungsgeschwindigkeit, aber auch wie viel sie sich über die Plausibilität ihrer Eingaben und über potenzielle Probleme mit den Formularfragen nachdachten.

Zusammenfassend wurden überwiegend positive Seiten des OTPP-Designs genannt, wie dass es übersichtlicher und kompakter und dadurch angenehmer wirke. Auch fühlten Testpersonen sich durch die Formularaufteilung geleitet. Darüber welche Abfragen

zusammenstehen sollten, gab es in Teilen allerdings keinen Konsens und so wurden die letzten Formularkategorien (Einkommensfeststellung und Vermögensgegenstände) zu einem besonderen Fall. Denn hier konnte ein Kompromiss erst über mehrere Iterationsrunden hinweg gefunden werden durch die Einordnung in themenspezifische Unterkategorien. Dies zeigt die Unterschiede zwischen verschiedenen Formularausfüllenden auf und unterstreicht damit die Notwendigkeit eines nutzer:innenzentrierten, formativen Usabilitytest bei der Erstellung eines solchen OTPP-Formulars.

4.3 Ergebnisse der Hauptstudie

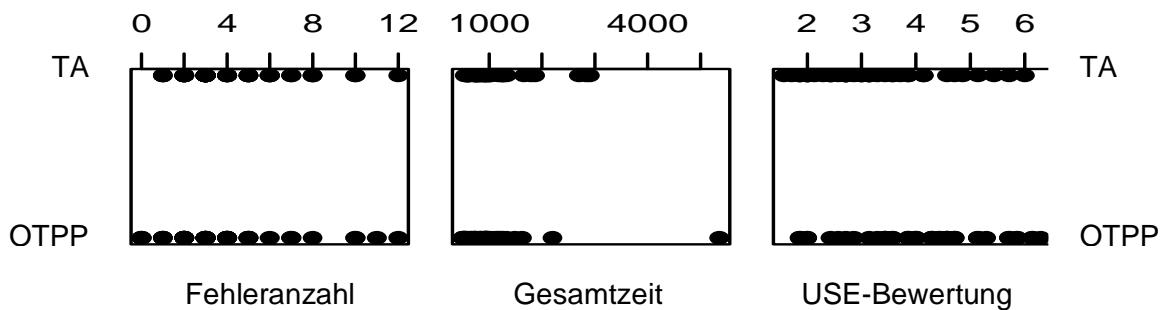
In diesem Unterkapitel werden die Ergebnisse des Online-Experiments vorgestellt. Dafür erfolgt eine Beleuchtung der Ergebnisse zum Gesamtmodell, wie auch einzeln zu den gemessenen abhängigen Variablen, die Rückschlüsse auf die Hypothesen erlauben. Zudem erfolgt eine weiterführende Erklärung von zusätzlich aufgetretenen Korrelationen und am Schluss eine qualitative Auswertung der freien Kommentare, die die Teilnehmenden am Ende eines jeden Formulars geben konnten.

4.3.1 Varianzanalyse des Gesamtmodells

Für die Ergründung des Einflusses, welches das OTPP-Layout auf die Kennwerte der Usability und der subjektiven Beanspruchung hat, wurden die Messwerte des Hauptexperimentes in einer einfaktoriellen MANOVA mit Messwiederholung berechnet. Für die Durchführung einer solchen Multivariaten Varianzanalyse gelten verschiedene Voraussetzungen. Zum einen wurde geprüft, ob Linearität zwischen der unabhängigen Variablen für jede Gruppe der abhängigen Variablen gegeben ist. Dies wurde durch eine grafischen Überprüfung anhand der Streudiagramme in Abbildung 3 bestätigt. Dies liegt auch in der Testform und den kategorialen UV begründet.

Abbildung 1

grafische Prüfung der Linearität zwischen der AV für jede Gruppe der UV



Daraufhin wurden univariante und multivariate Ausreiser aussortiert. Um die univariaten Ausreiser zu entfernen, wurden die Datensätze mit einem oder mehrere Werte von über einem 1,5-fachen Interquartilsabstand oberhalb des 75-Prozent Quartils ausgeschlossen wurden (vgl. Bortz & Schuster, 2010, S. 44). In diesem Zuge wurden 5 Datensätze aufgrund extremer Werte der Gesamtbearbeitungszeit und weitere 5 aufgrund der überdurchschnittlich hohen Anzahl an gemachten Fehlern ausgeschlossen. Die anderen Kategorien wiesen keine univariaten Ausreiser auf. Um die multivariaten Ausreiser zu identifizieren, wurden die Mahalanobis-Distanz berechnet (De Maesschalck et al., 2000). Auf deren Grundlage musste ein weiterer Datensatz ausgeschlossen werden. Daraufhin wurde die Daten auf Homogenität der Varianz-Kovarianz Matrizen getestet mittels des Fligner-Killeen Test (Algina et al., 1989). Die Ergebnisse des Test ($\chi^2 = 0,12$; $df = 1$; $p = .72$) zeigen auf, dass die Varianzen homogen sind mit $p > .05$. Außerdem sollen die Datenreihen auf Multikolinearität geprüft werden, wofür die VIF_i Werte berechnet und interpretiert wurden. Gemäß der Einordnung von Neter et al. (2005) liegt keine Multikolinearität vor bei Werten unter 10. Die für die relevanten Messgrößen errechneten Werte sind alle kleiner als 1,2 und deuten somit keinesfalls auf Multikolinearität hin. Zur Überprüfung der Homoskedastizität der Messgrößen eignet sich der Breusch-Pagan Test (Breusch & Pagan, 1979). Die Homogenität der Fehlervarianzen zwischen den Messwerten der AVs war gemäß dem Test erfüllt ($BP = 4.63$; $df = 2$; $p = .1$). Zur Überprüfung der univariaten Normalverteilungsannahme bietet sich der Shapiro-Wilk-Test an (Shapiro & Francia, 1972).

Ein solcher wurde über alle zu betrachtenden Bedingungen der Hauptfaktoren nicht signifikant ($n = 33$; $W = 0.96$; $p = .49$), woraus zu schließen war, dass die Daten univariat & multivariat normalverteilt sind. Die Ergebnisse der MANOVA wurden signifikant $F(1,64) = 3,77$; Wilks $\Lambda = .845$; partielle $\eta^2 = .056$; $p = .015$ *. Die Ergebnisse der anschließenden Post-Hoc Tests sind in Tabelle 4 aufgelistet und durch das Effektstärkenmaß η_p^2 nach Hemmerich (2018) ergänzt. Dies zeigt auf, dass unter den 4 untersuchten AV nur die Zufriedenstellung einen signifikanten Unterschied mit hoher Effektstärke zwischen den Formulartyp aufweist. Ferner wird deutlich, dass ein Fehler bei den zu Grunde liegenden Daten der subjektiven Beanspruchung vorlag, der im Folgenden noch näher thematisiert werden soll.

Tabelle 1*Ergebnisse der Post-Hoc Tests der Varianzanalyse des Gesamtmodells*

Abhängige Variable	F-Wert	p-Wert	η_p^2
Fehleranzahl	1.36	.25	.02
Gesamtzeit	0.69	.40	.01
Zufriedenstellung	10.99	.0015 **	.15
Subjektiv Beanspruchung	-	-	-

4.3.2 Ergebnisse der Effektivität

Es konnte zum einen gezeigt werden, dass die Hauptmessgröße der Effektivität, nämlich die Fehleranzahl, keinen signifikanten Einfluss erfährt von der Umgestaltung des Layouts zum OTPP-Formular. Denn während in den OTPP-Formularen durchschnittlich $M = 3,96$ ($SD = 2,9$) Fehler gemacht wurden, wurden bei den Formularen in der thematischen Aufteilung im Schnitt $M = 4,36$ ($SD = 2,52$) Fehler gemacht.

Darüber hinaus wurde die Effizienz auch anhand der Abbruchrate gemessen. Insgesamt kam es zu 3 Abbrüchen während des Aufgabenteils des OTPP-Formulars und zu 4 Abbrüchen während des Aufgabenteils des TA-Formulars. Somit wurde mittels dieser Daten gezeigt, dass es zu mehr Abbrüchen kam bei dem TA-Formular als bei dem OTPP-Formular. Diese Ergebnisse müssen allerdings im Kontext von 47 vollendeten Experimentdurchgängen und 23 Abbrüchen an anderen Stellen des Experiments (zumeist am Anfang) gesetzt werden.

Hypothesenprüfung:

H1: Die Präsentationsform gemäß des OTPP-Prinzips hat einen Einfluss auf die Effektivität.

Die Hypothese H1 kann auf einem Signifikanzniveau $p > .05$ nicht bestätigt werden

4.3.3 Ergebnisse der Effizienz

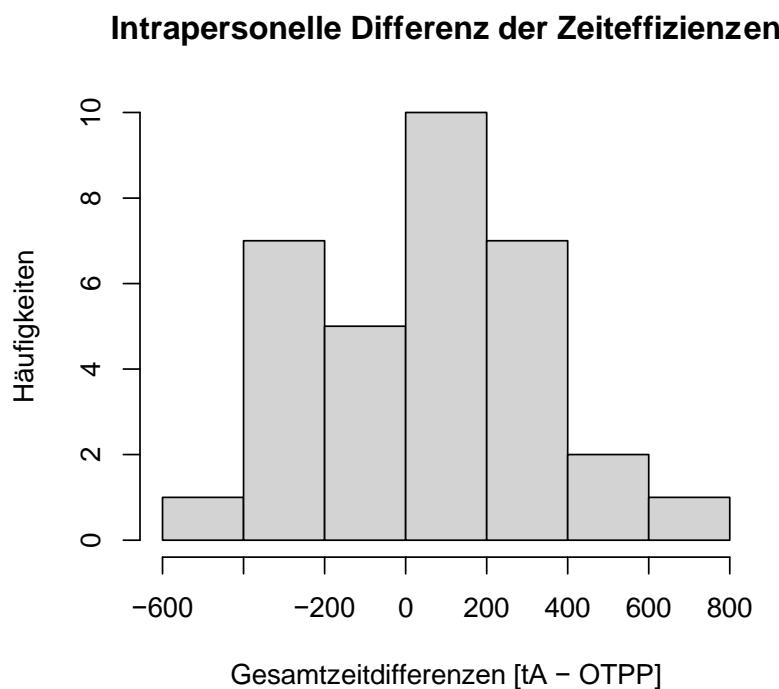
Es konnte gezeigt werden, dass die Hauptmessgröße der Effizienz, nämlich die zum Ausfüllen des Formulars benötigte Gesamtzeit, ebenfalls keinen signifikanten Einfluss erfuhr von der Umgestaltung des Layouts zum OTPP-Prinzip. Denn während der Ausfüllprozess eines OTPP-Formulars durchschnittlich $M = 1011s$ ($SD = 725$) dauerte, wurden die Formulare in der thematischen Aufteilung im Schnitt in $M = 1059s$ ($SD = 556s$) von den Teilnehmenden des Experiments ausgefüllt. Auch hierbei war das OTPP-Formular leicht, aber nicht signifikant anders als der bestehende Status-Quo in der thematischen Aufteilung. Außerdem fiel dabei auf, dass die Streuung bei beiden Formulartypen verhältnismäßig hoch ist ($SD > 0,5M$). Allerdings ist die Streuung beim OTPP-Formular nochmals höher ($SD > 0,7M$).

Ergänzend wurden die intrapersonellen Differenzen der Gesamtbearbeitungszeit pro Teilnehmenden betrachtet, was in Abbildung 4 verdeutlicht ist. Dies entsteht wenn für jede:n Teilnehmenden die benötigte Bearbeitungszeit im TA-Formulars von der benötigte

Bearbeitungszeit im OTPP-Formulars subtrahiert wird. In der Abbildung 4 fällt ein ausgeglichenes Verhältnis auf, das einer Standardnormalverteilung nahekommt. Die Normalverteilungsannahme kann in einem Shapiro-Wilk-Test ($n = 33$; $W = 0.97$; $p = .45$) bestätigt werden.

Abbildung 2

Intrapersonelle Differenz der Gesamtzeitbearbeitungszeiten



In dieser Messgröße ergab sich außerdem die Möglichkeit die neun verschiedenen Formulkategorien und deren jeweilige durchschnittliche Bearbeitungsdauer einzeln zu bewerten. Hierfür wurde eine weitere einfaktorielle MANOVA mit Messwiederholung gerechnet, denn die Bearbeitungszeit der einzelnen Kategorien musste unabhängig der Gesamtbearbeitungszeit berechnet werden, um Multikollinearität zu vermeiden. Denn die Gesamtzeit und die einzelnen Zeitkategorien weisen eine sehr hohe Kollinearität auf, da sie durch einfache Rechenoperationen ineinander überführt werden können. Dadurch könnte in einer MANOVA keine zusätzliche Varianz aufgeklärt werden, wodurch die Berechnung

redundant wäre (Hemmerich, 2023). Somit sind in dieser zweiten und detaillierten Untersuchung der Zeiteffizienz die neuen AVs die Bearbeitungszeiten der einzelnen Formularkategorien wohingegen die UV weiterhin durch die zwei Formulartypen OTPP & TA gegeben ist.

Ebenso wie zuvor wurden die weiteren Vorbedingungen getestet. Die Linearität zwischen der unabhängigen Variablen für jede Gruppe der abhängigen Variablen kann angenommen werden aufgrund der Testform und der kategorialen UV, wie zuvor dargelegt. Somit wurden zuerst die Daten auf univariate und multivariate Ausreiser untersucht. Um univariate Ausreiser aus den Datensätze auszuschließen, wurde alle Werte von über einem 1,5-fachen Interquartilsabstand oberhalb des 75-Prozent Quartils ausgeschlossen (vgl. Bortz & Schuster, 2010, S. 44). Um die multivariaten Ausreiser zu identifizieren, wurden die Mahalanobis-Distanz berechnet (De Maesschalck et al., 2000). Auf dieser Grundlage mussten 26 Datensätze ausgeschlossen werden. Daraufhin wurde die Daten auf Homogenität der Varianz-Kovarianz Matrizen getestet mittels des Fligner-Killeen Test (Algina et al., 1989). Die Ergebnisse des Test ($\chi^2 = 3,06$; $df = 1$, $p = .08$) zeigen auf, dass die Varianzen homogen sind mit $p > .05$. Außerdem sollen die Datenreihen auf Multikolinearität geprüft werden, wofür die VIF_i Werte berechnet und interpretiert werden. Gemäß den Berechnungen liegt keine Multikollinearität vor (Neter et al. 2005), da alle $VIF_i << 10$ waren. Die Homogenität der Fehlervarianzen zwischen den Messwerten der Zeitkategorien war gemäß dem Breusch-Pagan Test (1979) gegeben (BP = 12.23; df = 8, p = .14). Zur Überprüfung der Normalverteilungsannahme wurde abermals ein Shapiro-Wilk-Test gerechnet (Shapiro & Francia, 1972). Der Test wurde signifikant ($n = 18$; $W = 0,96$; $p = .05$) womit die Normalverteilungsannahme verworfen werden musste. Allerdings ist nach Finch (2005) die einfaktorielle MANOVA robust gegenüber Verletzung der Normalverteilungsannahme. Die Ergebnisse der MANOVA $F(1,34) = 0.58$; Wilks $\Lambda = .83$; partielles $\eta^2 = .02$; $p = .8$ wurden, entsprechend den Ergebnissen der Gesamtzeit, nicht signifikant. Darüber hinaus sind die

Ergebnisse der Post-Hoc-Tests in Tabelle 6 aufgelistet und durch das Effektstärkemaß η_p^2 nach

Hemmerich (2018) ergänzt.

Tabelle 2

Ergebnisse der Post-Hoc Tests der Varianzanalyse der benötigten Zeit pro Kategorie

Abhängige Variable	F-Wert	p-Wert	η_p^2
Zeit: Kategorie 1 (Angaben zur Person)	1.1	.3	.03
Zeit: Kategorie 2 (Anschrift)	1.05	.31	.03
Zeit: Kategorie 3 (Bankverbindung)	0.52	.47	.02
Zeit: Kategorie 4 (Versicherung)	0.1	.76	.002
Zeit: Kategorie 5 (Ausbildung)	0.81	.37	.02
Zeit: Kategorie 6 (Lebenslauf)	0.27	.6	.008
Zeit: Kategorie 7 (Eltern)	0.05	.82	.002
Zeit: Kategorie 8 (Einkommen)	0.76	.39	.02
Zeit: Kategorie 9 (Vermögen)	0.03	.87	.0008

Hypothesenprüfung:

H2: Die Präsentationsform gemäß des OTPP-Prinzips hat einen Einfluss auf die Effizienz.

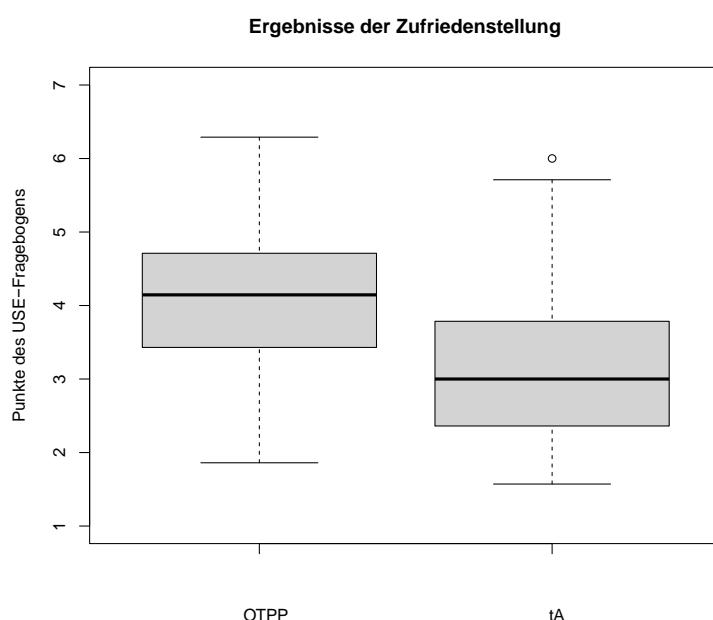
Die Hypothese H2 kann auf einem Signifikanzniveau $p > .05$ nicht bestätigt werden

4.3.4 Ergebnisse der Zufriedenstellung

Darüber hinaus ist in Tabelle 4 ersichtlich, dass die Formulargestaltung gemäß des OTPP-Layouts einen hochsignifikanten Einfluss auf die Zufriedenstellung der Teilnehmenden im Experiment hatte. Zudem deutet die Varianzanalyse mittels des partiellen Eta-Quadrats auch auf eine hohe Effektstärke ($\eta^2_p > .14$) hin. Bei der Gegenüberstellung der Ergebnisse des USE-Fragebogens in Abbildung 5, wird ersichtlich, dass die Bewertungen in Bezug auf das OTPP-Formular ($M = 4,1$; $SD = 1,3$) durchschnittlich höher ausfielen als die Bewertung in Bezug auf das Formular in der thematische Aufteilung ($M = 3,3$; $SD = 1,2$). Diese Werte sind im Kontext der USE-Fragebogenskala zu sehen, welche 1 als schlechteste Wertung und 7 als beste Wertung zuließ. Weiterhin sind die für beide Gruppen vergleichbare Standardabweichungen, ebenso wie die vergleichbare Range beider Verteilungen (Range OTPP: 4,3; Range TA: 4,7) aufgetreten.

Abbildung 3

Ergebnisse des Zufriedenstellungsmoduls des USE-Fragebogens geordnet nach Formulartyp

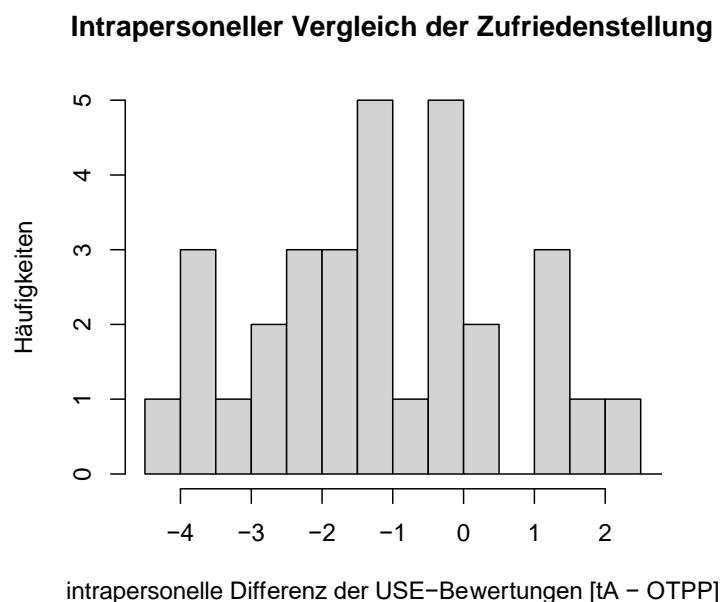


Werden die Ergebnisse des USE-Fragebogens in einer intrapersonellen Bedingungen subtrahiert, entsteht die Abbildung 6. Diesem kann entnommen werden, dass lediglich 7

Proband:innen im positiven Bereich der Grafik befinden und damit ausdrücken, dass sie das TA-Formular als zufriedenstellender empfanden. Wohingegen mehr als dreimal so viele Teilnehmende sich im negativen Bereich der Abbildung befinden und somit aufzeigen, dass das OTPP als zufriedenstellender empfanden im direkten Vergleich. Auch fällt in der Abbildung 6 auf, dass die Extremwerte der Differenzen im negativen Bereich betragsmäßig deutlich höher ausfielen

Abbildung 4

Intrapersonelle Differenz der Ergebnisse des Zufriedenstellungsmoduls des USE-Fragebogens



Hypothesenprüfung:

H3: Die Präsentationsform gemäß des OTPP-Prinzips hat einen Einfluss auf die Nutzer:innenzufriedenstellung.

Die Hypothese H3 kann auf einem Signifikanzniveau $p > .05$ bestätigt werden

4.3.5 Ergebnisse der subjektiven Beanspruchung

In Bezug auf die letzte relevante Messgröße der subjektiv wahrgenommenen Beanspruchung, kann keine Aussage getroffen werden. Denn bei der Messung auf den Skalen des TLX-Fragebogens kam es im Online-Experiment zu Fehlzuordnungen und dadurch zu einer Vielzahl von fehlenden Werten. Diese äußerten sich insofern in den Daten, als dass lediglich der schlechteste - als auch der beste Wert einer Skala aufgenommen wurde, ebenso wie die semantische Mitte der Skala. Alle Werte dazwischen wurden als 0 kodiert. Infolgedessen konnten keine statisch validen Rechnungen mit dem vorliegenden Datensatz ausgeführt werden. Dennoch sind Tendenzen in den gegebenen Werten zu erkennen, die für eine Vorteil vom OTPP-Formular im Vergleich zum TA-Formular sprechen. Besonders in den Kategorien der Leistung und der Frustration.

Hypothesenprüfung:

H4: Die Präsentationsform gemäß des OTPP-Prinzips hat einen Einfluss auf die subjektive Beanspruchung.

In Bezug auf die Hypothese H4 kann keine Aussage auf Grundlage der Ergebnisse des Experiments getroffen werden

4.3.6 Weitere Beobachtungen

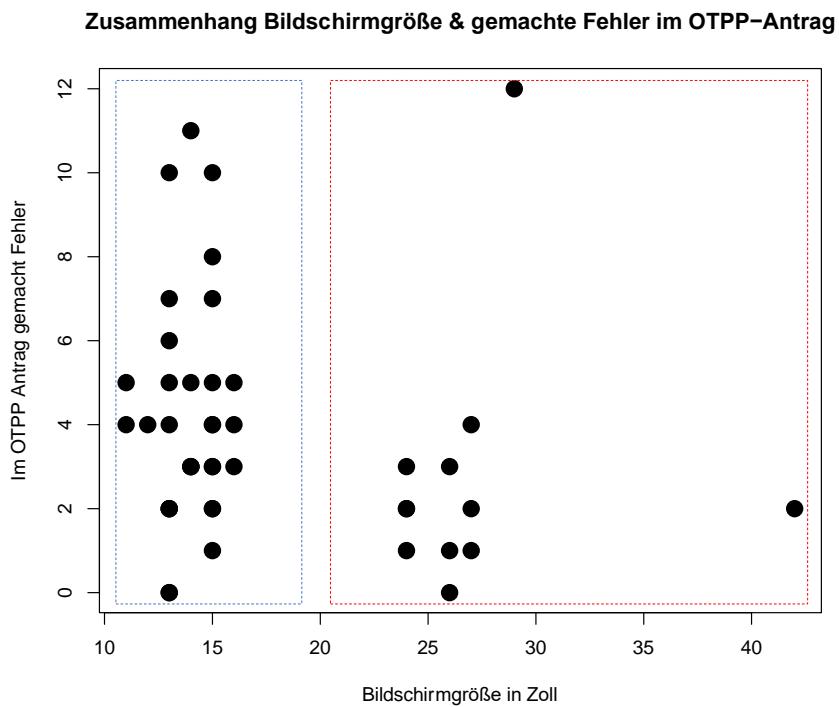
Die Daten des Experiments wurden auf weitere Auffälligkeiten und aufgetretene Gruppenunterschiede hin untersucht. Zum einen darauf, ob die gemessenen Werte abhängig sind von der Größe des verwendeten Bildschirms. Ebenso wie darauf, ob die Vorerfahrung mit dem Formular einen Einfluss auf die Messgrößen beim OTPP-Formular hat.

In Bezug auf die gemachten Fehleranzahl im Antrag konnte festgestellt werden, dass die Anzahl der Fehler im OTPP-Formular geringer wurde, wenn ein größerer Bildschirm verwendet wurde (vgl. Abbildung 7). Dieser Effekt war bei dem Formular der thematischen Aufteilung nicht in entsprechendem Maße zu erkennen (vgl. Abbildung 8). Dies führte zu einer zusätzlichen Betrachtung der Gruppenunterschiede, wobei die UV weiterhin die unterschiedlichen Formularkategorien und die AV die Fehleranzahl bei Bildschirmgrößen von über 20 Zoll mit $n = 12$ für beide Gruppen waren. Die relevanten Werte beider Gruppen wurden durch die rot gestrichelte Markierungen in Abbildung 7 & 8 verdeutlicht. Die durchschnittliche Fehleranzahl dieser Teilstichprobe betrug beim OTPP-Formular $M = 2,75$ ($SD = 3,1$), während beim TA-Formular $M = 4,75$ ($SD = 2,8$) betrug. Gegen die gegebene Verletzung der Normalverteilungsannahme ist der Welch-Test robust nach Zimmermann und Zumbo (1993). So wurde ein Welch-Test für gepaarte Stichproben ($t(11) = 1,7$; $p = .07$) gerechnet, der allerdings keinen signifikanten Unterschied zwischen in der Population aufzeigte. Anschließend wurde ein T-Test auf Unterschiede zwischen den mittleren Fehleranzahlen des OTPP-Formulars bei Bildschirmen von unter 20 Zoll ($n=32$) und über 20 Zoll ($n=12$) durchgeführt. Diese beiden Bereiche der relevanten Werte wurden durch die blau bzw. rot gestrichelten Markierungen in Abbildung 7 verdeutlicht. Gemäß der Monte-Carlo-Studien bleibt der T-Test robust bei unterschiedlich großen Stichproben, wenn die Varianz annähernd gleich sind (Sawilowsky & Blair, 1992). Bei der Teilstichprobe von Bildschirmen unter 20 Zoll betrug die durchschnittliche Fehleranzahl $M = 4,28$ ($SD = 2,71$) und bei der Teilstichprobe von Bildschirmen über 20 Zoll betrug $M = 2,75$ ($SD = 3,1$). Der T-Test ergab aber ebenfalls keine Signifikanz ($t(17.7) = 1,5$; $p = .075$).

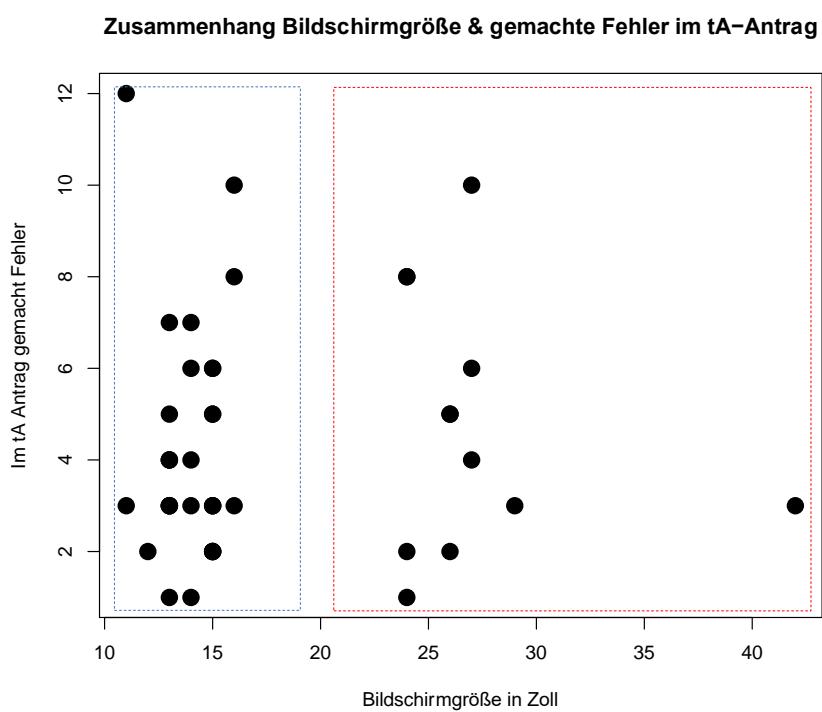
Zur Vollständigkeit werden auch die Kennwerte der Fehleranzahl des TA-Formulars bei Bildschirmdiagonalen unter 20 Zoll genannt. Diese ist durch den blau gestrichelten Kasten in Abbildung 8 verdeutlicht. Hierbei betrug die durchschnittliche Fehleranzahl $M = 4,22$ ($SD = 2,51$), was den Werten des OTPP-Formulars sehr ähnlich ist.

Abbildung 5

Streudiagramm der Bildschirmgröße und der im OTPP-Formular gemachten Fehler

**Abbildung 6:**

Streudiagramm der Bildschirmgröße und der im TA-Formular gemachten Fehler



Weitere bedeutungsvolle Unterschiede in Bezug auf die Bildschirmgröße konnte weder bei der Messgröße der Effizienz noch bei der Messgröße der Zufriedenstellung festgestellt werden.

Bei einer weiterführenden Betrachtung konnte ein positiver Zusammenhang zwischen der Vorerfahrung und der Zufriedenstellung festgestellt werden. Die Vorerfahrung wurde dabei quantifiziert über die Male, die die Teilnehmenden bereits ein BAföG-Formular ausgefüllt haben. Im Vorfeld zu dieser Berechnung wurden eine A-Priori Analyse durchgeführt wurde, um zu bestätigen, dass eine ausreichende Populationsgröße gegeben ist, für einen mittleren Effekt von $f = 0.25$ (Cohen, 2016) auf dem α -Niveau von 0.05. Die hierfür errechnete Mindestanzahl für eine angestrebten statistische Power von $1 - \beta = 0.9$ entspricht genau 44,04 Teilnehmende. Außerdem wurden die relevanten Voraussetzungen für eine lineare Regression geprüft. Dass ein linearer Zusammenhang zwischen UV und AVs besteht, sowie dass die AVs metrisch skaliert sind ist gegeben und ergibt sich aus dem Methodikteil beschriebenen Vorgehen der Datenmessung. Dass die Residuen voneinander unabhängig sind, kann mit dem Durbin-Watson Test überprüft werden (Savin & White, 1977). Für das vorliegende Model (DW = 2,39; $p = .91$) bestätigt der Test die Unabhängigkeit der Residuen. Außerdem wurde das Model auf annähernde Normalverteilung der Residuen getestet. Hierbei wurde allerdings anders als zuvor auf den Shapiro-Wilk-Test verzichtet, da dieser bei größeren Stichproben schon bei kleinsten Verletzungen der Normalverteilungsannahme signifikant wird (Field, 2018, S. 248f). Somit wurde ein QQ-Plot der Mediane zurückgegriffen, die in Abbildung 9 und Abbildung 10 zu sehen sind. Diese lassen eine annähernde Normalverteilung plausibel erscheinen.

Abbildung 7
QQ-Plot der Vorerfahrung

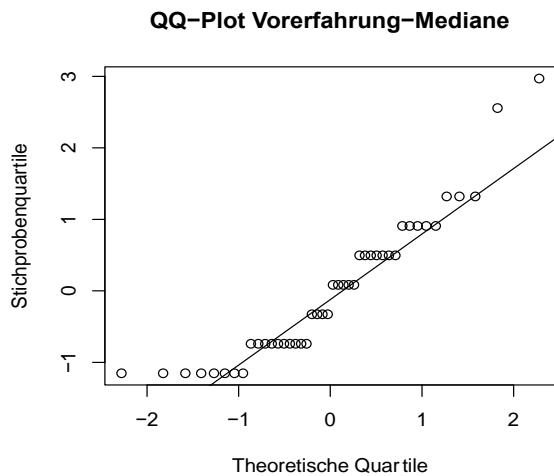
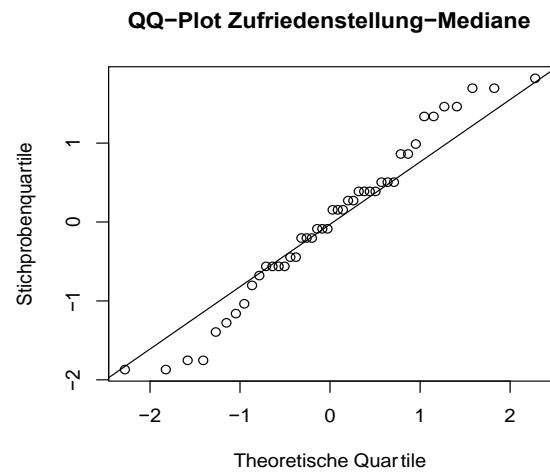


Abbildung 8
QQ-Plot der Zufriedenstellung



Zuletzt wurde auch überprüft, ob Homoskedastizität bei den Residuen vorliegt. Wie zuvor wird hierfür der Breusch-Pagan Test (1979) ausgeführt. Da dieser Test nicht signifikant wurde ($BP = 0.40; df = 1, p = .53$), wurde die Nullhypothese des Tests weiterhin angenommen und somit bestand ebenfalls die Annahme der Homoskedastizität weiter. Auf Grundlage dessen kann die lineare Regression wie vorgesehen durchgeführt werden. Diese kann mit der Gleichung $y = -1,0307 + 0,0323x$ beschrieben werden. Zusätzlich wird in Abbildung 11 der Zusammenhang der Vorerfahrung & der Ergebnisse des Zufriedenstellungsmoduls anhand einer roten Trendlinie, die sich aus einer linearen Regression ergibt, verdeutlicht. Weitere relevante Daten der ausgeführten linearen Regression wie die F-Statistik, die Standardabweichung, der P-Wert, und deren eher geringer Erklärungsbeitrag von R^2 werden in der Tabelle 7 aufgezeigt. Ein vergleichbarer, signifikanter Effekt konnte bei den Daten des TA-Formular nicht nachgewiesen werden und wird somit auch nicht weiter ausgeführt.

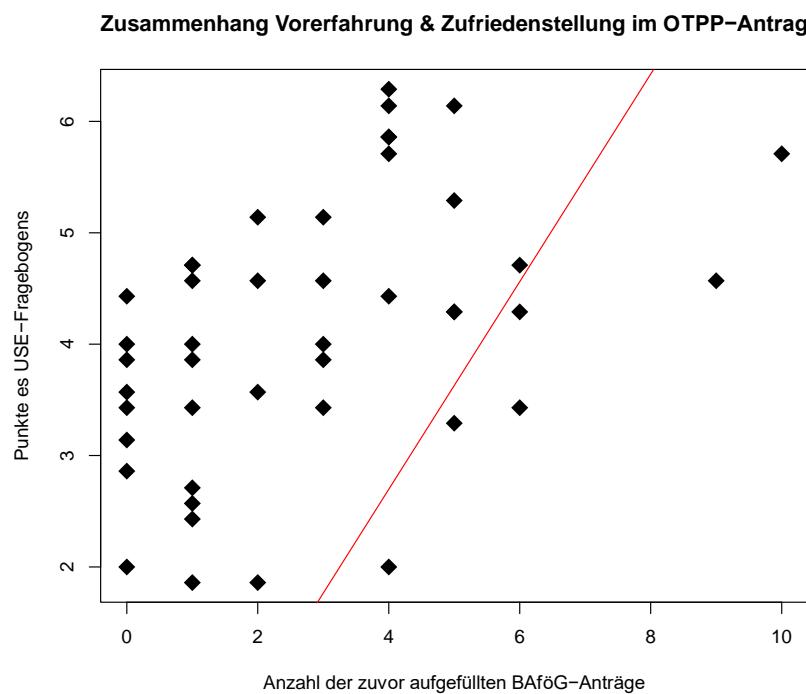
Tabelle 3

Daten der Lineare Regression: Einfluss der Vorerfahrung auf die Zufriedenstellung

R^2	Adjustiertes R^2	Standardfehler	F (df = 1; 42)	p-Wert
0,2128	0,1941	0,2766	11,36	.00162 **

Abbildung 9

Streudiagramm der Vorerfahrung und der Ergebnisse des Zufriedenstellungsmoduls des USE-Fragebogens des OTPP-Formulars mit Regressionsgerade



Weitere bedeutungsvolle Unterschiede in Bezug auf die Vorerfahrung wurden bei keiner weiteren Messgröße festgestellt.

4.3.7 Thematische Analyse der freien Kommentare

Im direkten Anschluss an die auszufüllenden Nachbauten des BAföG-Formulars, kam jeweils eine Seite auf der freie Kommentare geäußert werden konnten in Bezug auf das Erlebnis des Ausfüllprozesses (dargestellt in Anhang A9). Insgesamt wurden auf diese Art 247 Kommentare unterschiedlicher Qualität und Länge von den Teilnehmenden gesammelt, wovon sich 91 Kommentare (58 positiv, 33 negativ) direkt oder indirekt auf die OTPP-Gestaltung bezogen. Diese Kommentare dienen als Grundlage für die nun folgende qualitative Auswertung nach der thematischen Analyse von Braun und Clarke (2006). Diese Auswertung zeigt zuerst benannte positive Seiten des OTPP-Formulars, auch im Vergleich zum anderen Formulartypen. Anschließend werden Verbesserungswünsche und angesprochene negative Seiten des OTPP-

Formulars aufgezeigt. Abschließend wird dies in den Kontext des gesamten Ausfüllprozess eines Formulars gerückt. Zusammenfassend bietet die Auswertung eine Überblick über das zu größeren Teilen positives Meinungsbild in Bezug auf das OTPP-Layout und gibt außerdem auch Ansätze, wodurch dieses weiterhin verbessert werden kann.

Positive Seiten, die von den Testpersonen genannt wurden, waren, dass dieser Formulartyp eine „einfache Struktur“ mit „klaren Fragen“ habe und daher „einfach auszufüllen“ sei. Mehrfach wurde von den Proband:innen geäußert, dass dieser Formulartyp „übersichtlicher“ sei durch das „minimalistische Design“. Somit seien es „nie zu viele Fragen auf einmal“ schrieben mehrere Person in die Kommentarspalten. Durch die „häppchenweise Aufteilung“ oder „Schritt für Schritt Anleitung“, wie das OTPP-Prinzip des häufigeren genannt wurde, „ühlte[n] sich [Teilnehmende] geleitet“. Zudem wurde auch genannt, dass es durch die „Kleinschrittig[keit zu] größeren Erfolgserlebnissen“ kam und zudem „kein Stress / Zeitdruck“ und „weniger Frustration“ entstand. Eine genannte Begründung dafür war, dass diese „Formularverkürzung durch Unterkategorien“ suggeriert werde. Andere Personen sagten hingegen, dass das Formular im OTPP-Layout „angenehmer, aber auch mehr“ sei. Zwei Personen äußerten auch, dass der Ausfüllprozess dieses Formulartyps „schneller als der zuvor“ gewesen sei und sie „durch den Fragebogen durchrutschte[n]“. Zudem wurde positiv angemerkt, dass „man keine Scheu [hatte] vor den Fragen, die noch kommen sollten, weil diese eben nicht auf derselben Seite zu sehen waren und keinen Druck erzeugten“. Oftmals wurde auch von Teilnehmenden in Bezug auf das OTPP-Formular vermerkt, dass „Diese Art des Formulars [ihnen] kürzer“ vorkam als das andere Formular im TA-Layout. So wurde vereinzelt auch klar geäußert, dass der OTPP- „Aufbau [...] besser“ sei. Dies geschah in Worten wie „Es war viel besser als beim ersten Antrag: Weniger Fragen pro Screen, somit weniger Scrollen nötig und das Gefühl, insgesamt viel schneller voranzukommen und weniger Frustration“. Denn das Formular im OTPP-Stil „war ohne große Überwindung leicht auszufüllen, was sonst eher mit Stress und viel Zeit in Anspruch genommen hat, war hier leicht zu bewältigen“. Auch

im Vergleich zum „originalen Formular“ seien die „Eintragung schneller möglich“ gewesen, gaben Teilnehmende zu verstehen. Im direkten Vergleich der beiden im Test verwendeten Formulartypen wurde gesagt, dass die Seiten des TA-Layout „einfach zu lang und damit anstrengend für das Auge“ waren und dass die „letzten Seiten [...] beinahe überfordernd“ waren (dabei sind Kategorien der Einkommens- & Vermögensfeststellung gemeint). Ergänzend wurde aber geäußert, dass die Teilnehmenden im OTPP-Formular „gerne noch eine Status-Leiste gehabt [hätten,] um schnell zu bestimmtem Punkt zurückzugehen [und um zu sehen was sie bereits] abgehakt habe[n].“. Denn negativ über das OTPP-Formular wurde geäußert, dass es schwer sei darin zu navigieren und man „umständlich zurück [müsste]“. Ebenfalls wurde geäußert, dass die vielen Einzelseiten „etwas nervig“ seien und eine Übersicht bei „Einkünften und Vermögen“ fehle, von all dem „was als nächstes kommt“. Am häufigsten wurde das „zu viele Weiterclick[en]“ kritisiert und dass dadurch bei manchen der Eindruck entstehe, dass das OTPP-Formular „länger dauerte“. Tatsächlich äußerten mehr Personen, dass das TA-Formular „schneller auszufüllen“ sei als beim OTPP-Formular. Auch hat mehreren Teilnehmenden am TA-Formular besser gefallen, dass man „nicht benötigte Fragen [...] ausblenden“ konnte. Andererseits wurde aber auch in Bezug auf das Aufklappen bzw. Ausblenden von gewissen Fragen gesagt, dass es angenehmer ist, wenn „sich die Seiten nicht [ändern], wenn man eine gewisse Option ausgewählt [hat]. Das wirkte so hektisch und hat mir das Gefühl gegeben, ich hätte das falsche gewählt, da ich mehr ausfüllen musste“. Einzeln wurde auch von Teilnehmenden genannt, dass „Mehr Infos pro Seite“ ihnen besser gefielen.

In jedem Fall sollen diese Ergebnisse aber im Kontext gesehen werden, über den sich die Teilnehmenden äußerten, dass „Ob ich es gut oder schlecht finde, spielt hierbei nicht wirklich eine Rolle. Es ist halt der Ausfüllprozess eines Formulars. Das geschieht emotionslos und weil ich es muss. Da kommen keine Gefühle hoch“ bzw., dass „so oder so ist das Ausfüllen von langen Formularen immer nervig“.

5 Diskussion

Dieses finale Kapitel hat das Ziel die Ergebnisse der Studie zusammenzufassen, zu interpretieren und in den bisherigen Forschungsstand einzuordnen. Daraufhin wird diskutiert welche Implikationen der Studie in verschiedene Bereiche des öffentlichen Lebens und damit der Praxis haben kann. Anschließen werden die Limitation erläutert, die im Rahmen des Studienprojekts auftraten, um beurteilen zu können, ob und inwieweit das Ergebnis sich als hilfreich erweisen können. Final werden Überlegungen und Anregungen zu weiterführende Forschungsarbeiten die Arbeit abschließen und in einem Fazit münden.

5.1 Zusammenfassung und Interpretation der Ergebnisse

Da die Ergebnisse der Vorstudie lange vor den Ergebnissen der Hauptstudie vorlagen und in gewisser Weise als Grundlage dienten, werden die Ergebnisse getrennt voneinander thematisiert. Diese Trennung dient ebenfalls dem Zweck herausarbeiten zu können, inwiefern die Ergebnisse sich unterscheiden und um Übertragungseffekte zu vermeiden. Zudem ist die Interpretation einer qualitativen Arbeit wie der Vorstudie eine andere als die Interpretation einer quantitativen Arbeit wie der Hauptstudie

5.1.1 Vorstudie

Die in der Vorstudie festgestellten Antworttendenzen zeigen ein hauptsächlich positives Bild in Bezug das OTPP-Layouts. Da es überschaubarer und kompakter wirkt und nicht zu viel auf einmal verlangt. Ebenso wie der Ausfüllprozess angenehmer und entspannter sei als bei langen Formularseiten, da die Testpersonen sich durch das Formular geleitet fühlen und dies teilweise auch motiviert. Ob bzw. inwiefern diese Antworttendenz aber von einem Intervieweffekt beeinflusst wurde, lässt sich in dieser Vorstudie nicht ermitteln. Negative

Tendenzen bzgl. einer feingliedrigen Aufteilung waren in erster Linie, dass es die Teilnehmenden stört, so häufig die Seite zu wechseln. Dies wurde allerdings von der Teilnehmenden geäußert, die die meiste Vorerfahrung mit dem deutschen eGovernment und dessen Formularen aufwies und zudem merklich schneller las als die anderen Testpersonen. Dies kann die These unterstützt, dass die Vorteile eines OTPP-Formular/einer feingliedrigen Aufteilung weniger gewichtig sind, je mehr Vorerfahrungen die Benutzenden aufweisen. Allerdings basiert dies nur auf der Meinung einer einzelnen Testperson, sodass es diese These in weiterführenden Studien zu falsifizieren gilt. Diese Interpretation kann somit bestenfalls als Hinweis dienen, dass es einen solchen Effekt bei steigender Vorerfahrung geben kann. Über alle Testpersonen hinweg gibt es aber eindeutig Unterschiede in der Art und Weise wie das Formular ausgefüllt wird. Dies zeigt sich zum einen darin, ob die Ausfüllenden die Texte lediglich überfliegen oder präzise lesen. Was sich bspw. darin niederschlagen kann, ob die vorgegebenen Formate – beispielsweise für ein Datum mittels (MM/JJJJ) – beachtet werden oder nicht oder darin ob gewisse formularspezifischen Fachausdrücke die Ausfüllende zum Stocken bringen oder nicht. Unterschiede zeigen sich damit auch nicht nur in der benötigten Zeit zum Ausfüllen des Formulars, sondern auch darin, ob kleine Pausen entstehen, in denen über die Plausibilität der Angaben nachgedacht wird. Außerdem liegen große Unterschiede zwischen den Nutzer:innen darin, ob ihre Daten eher bereitwillig eingegeben werden oder sie sich eher sträuben und überzeugt werden müssen, Informationen preiszugeben, die sie selbst betreffen. Allerdings waren keine festen Muster oder Korrelationen zu erkennen, die aus der Bearbeitungsweise eine gewisse Präferenz in Bezug auf die Formulargestaltung erlaubt hätten abzuleiten. Ebenso wenig wie sich vorhersagen lässt, welche Nutzer:innen nun welche Eigenschaften im Ausfüllprozess zeigen werden. Klar zeichnet sich hingegen ab, dass es unterschiedliche, persönliche Präferenzen gibt in Bezug darauf, welche Fragen zusammen auf einer Seite stehen sollen oder nicht – gut zu erkennen am Beispiel der Unterkategorien der Einkommens- oder Vermögenssituations – und dass hierbei ein ideale Formularaufteilung

jeweils der Situation und den Eigenschaften des Nutzenden obliegt. Auch wenn Parallelen dazu zu erkennen sind, dass Informationen, welche auch in einem semantischen Netz eine bedeutende Nähe aufweisen, zusammenstehen sollen. Als Beispiel hierfür kann gut der eigene Name und Vorname gelten, die zusammen stehen sollen wohingegen das eigene Geschlecht nicht auf derselben Seite stehen sollte. Obwohl die Nennung zu welchem Geschlecht eine Person sich selbst zuordnet ebenfalls sehr einfach abrufbar und ebenso essenziell zur Charakterisierung seiner selbst ist. Durchaus schwerer zu prognostizieren, waren allerdings die Informationen, die scheinbar nicht unbedingt zusammenstehen müssen. Dies zeigt abermals, dass ein solcher formativer nutzer:innenzentrierter Test essenziell ist bei der Erstellung eines OTPP-Formulars. Denn das Ergebnisformular zeigt eindeutig auf, dass nicht alle Eingabefelder einer Formularkategorie auf einer Seite stehen müssen. Weiterführend und interessant ist es auch, sich die verschiedenen Eigenschaften der Kommentare der verschiedenen Iterationsrunden zu betrachten. Denn während in der ersten Runde noch weniger positive Kommentare fielen, da das Formular noch zu kleinteilig erschien, wurden dabei folgerichtig die meisten Kommentare genannt, die zum Zusammenziehen gewisser Fragen beitrugen. Nachdem die Eingabefelder gemäß dem Vorgaben der ersten Runde zusammengezogen wurden, gab es schon deutlich mehr positive Kommentare in Bezug auf das Formularlayout, dafür aber weniger Kommentare wie das Formular weiterhin zusammengezogen werden sollte. Allerdings zeichneten sich dabei erstmalig Kontroversen ab zwischen den verschiedenen Präferenzen von einzelnen Nutzer:innen. Besonders in Bezug auf die Kategorien des Einkommens und des Vermögens, wie bereits erwähnt. Nachdem die Kommentare der Iterationsrunde ebenfalls eingearbeitet wurden, spiegelte die Kommentare der dritten Iterationsrunde das bislang beste Meinungsbild der befragten Personen wider, was nicht zuletzt dazu führt, dass in der dritten Iterationsrunde der formative Designprozess beendet wurde, durch das Erreichen des Usabilitygoals. Die Aufteilung des OTPP-Formular, welches das Ergebnis dieser Vorstudie darstellt, kann im Anhang A2 und A3 eingesehen werden.

5.1.2 Hauptstudie

Aus den Ergebnissen lässt sich ableiten, dass die Effektivität des OTPP-Formulars zwar geringfügig, aber nicht signifikant besser ist als die des TA-Formulars. Dies gilt für die Messung anhand der Fehleranzahl, ebenso wie für die Abbruchrate. In Bezug auf die zweitgenannte Messung ist zu sagen, dass im Verhältnis zu den 47 beendeten Experimenten und 23 Abbrüchen an anderen Stellen ist der Unterschied zwischen den 3 Abbrüchen während des OTPP-Formular und den 4 Abbrüchen während des TA-Formulars marginal. Ein sehr ähnliches Bild ergibt sich bei den Messungen der Effizienz, denn auch hier sind die Messwerte des OTPP-Formulars zwar geringfügig, aber nicht signifikant besser als die des TA-Formulars. Dies gilt für die Gesamtzeit, ebenso wie für die Betrachtung der Zeiten der Einzelkategorien. Für die beiden Fälle der leistungsbasierten Messgrößen der Usability kann resümierend gesagt werden, dass das OTPP- ebenso gut ist wie TA-Layout und dadurch eine echte Alternative darstellt für die Formulargestaltung. Wobei die Stärke der OTPP-Formulargestaltung im Test in der deutlich höheren Zufriedenstellung lag und damit bei der nutzer:innenorientierten, meinungsbasierten Messgröße der Usability. Denn im Experiment hat sich das OTPP-Formular als weitaus zufriedenstellender erwiesen als das TA-Formular. Diesen Eindruck unterstreicht eine Zufriedenheitsdifferenzen in Abbildung 6 von bis zu 4,5 Punkten bei einer Maximaldifferenz von 6, welche im Kontrast zu nur sehr wenig Gegenstimmen steht. Vor dem Hintergrund einer sonst recht hohen Varianz in Bezug auf das OTPP-Formular und des wie am Beispiel der Effizienz dargestellten ausgeglichenem Verhältnis (das beinahe an eine Standardnormalverteilung erinnert) ist dieses Ergebnis noch relevanter. Die höhere Zufriedenheit spiegelt sich auch in den Kommentaren der Teilnehmenden wider, die aussagen, dass durch die einfache Struktur und das minimalistische Design des OTPP-Antrags die Teilnehmenden größere Erfolgserlebnissen erfuhren und zudem weniger Stress und Frustration entstand. Dennoch kann bezüglich der gestellten Forschungsfragen, welche ist, ob das BAföG-Formular durch Unterschiede des Seitenlayouts in der Usability und in der vom Nutzenden

wahrgenommene Beanspruchung verändert werden, im Vergleich zu dem bestehenden Formulartyp keine eindeutige Antwort gegeben werden. Denn ob Unterschiede in der wahrgenommenen Beanspruchung vorliegen, konnte leider in dem Experiment nicht gezeigt werden, aufgrund des bereits thematisierten Datenfehlers. Mit Sicherheit kann jedoch gesagt werden, dass keine signifikante Veränderung in der Effektivität oder der Effizienz bewiesen werden konnte, dafür aber ein signifikant verändertes Ergebnis in der Messgröße der Zufriedenstellung. Somit ist zu resümieren, dass die Usability auch nur bedingt verändert werden konnte, durch Unterschiede des Seitenlayouts.

Ein über die Forschungsfrage und die Hypothesen hinausgehender, interessante Befund war jedoch die vielfache Nennung, dass eines der Formulare kürzer bzw. schneller war als ein anderes, obwohl das im Durchschnitt nicht stimmt und auch nicht unbedingt in den jeweiligen intrapersonellen Daten so widergespiegelt wurde. Dies kann einerseits damit erklärt werden, dass gemäß Sauro und Lewis (2009) Ease of Use assoziiert wird mit Performance. Sodass die Aussage dessen umgewandelt werden kann dazu, dass ein gewisser Personenkreis das OTPP-Formular als einfacher zu benutzen empfand und ein anderer, vermeintlich kleinerer, Personenkreis das TA-Formular. Dies wiederum verdeutlicht den Eindruck, wie er auch schon in der Vorstudie angeklungen ist, dass es große Unterschiede an Formularausfüllenden gibt. Manche kommen sehr gut mit dem neuen Layout des OTPP-Prinzip zurecht, andere nicht. Dies könnte auch die hohe Varianz in den Messungen des OTPP-Designs begründen, wie sie bspw. bei der Gesamtzeit aufgetreten ist. Dies kann entweder daran liegen, dass Leute mehr oder weniger gut mit Neuem zurechtkommen oder aber, dass das OTPP-Prinzip nur für eine bestimmte Gruppe von Formularausfüllenden geeignet ist. Dieser Eindruck wird ebenfalls von den Kommentaren der Hauptstudie und auch der Vorstudie bestätigt, da sich Teilnehmende darin teilweise sogar widersprachen.

5.2 Einordnung der Ergebnisse in den Forschungsstand

Zum einen kann die Theorie von Adam Silver (2018, S. 370), dass das OTPP-Prinzip für Noviz:innen sehr hilfreich sein wird, wohingegen es für Expert:innen weniger unterstützend wirkt, anhand der Ergebnisse der Hauptstudie nicht unterstützt werden. Denn im Experiment konnte sogar ein positiver Zusammenhang von Erfahrung und der Zufriedenstellung im OTPP-Formular festgestellt werden.

Ebenso wenig kann die These von Steve Krug (2014, S. 54) unterstützt werden, dass es nicht darauf ankommt, wie oft jemand klicken muss, solange jeder einzelne Klick eine unbekümmerte und unmissverständliche Auswahl darstellt. Denn es wurde sowohl in der Vor- als auch der Hauptstudie beanstandet, dass man so häufig klicken muss, um zu den nächsten Eingabefeldern zu gelangen.

Die These von Shneiderman und Leavitt (2006, S. 66), dass oftmals die untere Kante des Bildschirms auch als Boden der Seite angesehen wird und daher Seiten bestenfalls auf die Länge eines Bildschirms reduziert sein sollten, konnte ebenfalls nicht bestätigt werden durch die Ergebnisse der Studie. Denn auch längere Formularseiten wurden von den Teilnehmenden in Gänze ausgefüllt. Somit wird eher die These von Paul (2015) von den Ergebnissen getragen, dass die Nutzer:innen es inzwischen gewohnt sind zu scrollen.

Auch kann die These von Shneiderman und Leavitt (2006, S. 75) nicht bestätigt werden, die besagt, dass Nutzer:innen effizienter sind auf gut organisierten, kurzen Seiten. Denn im Experiment waren die Teilnehmenden in etwa gleich effizient auf kurzen wie auch langen Seiten, auf denen gescrollt werden musste.

Die vertretene Auffassung von Adam Silver (2017), dass der gesamte Designprozess des OTPP-Formulars weniger aufwendig ist, denn er ist stets minimalistisch und kann sehr einfach konsistent gehalten werden, kann nicht von Masteranden bestätigt werden. Denn entgegen der Aussage war der Designprozess mit Nutzer:innentest als Vorstudie, weitaus

aufwändiger als eine reine Kategorisierung der zu stellenden Fragen, wie für das TA-Formular benötigt.

Nur bedingt unterstützt werden kann die Aussage von Jarrett und Gaffney (2009, S. 102), dass auch sehr kurze Seiten in Ordnung sind, denn es existiere beim digitalen Formular keine Mindestlänge. Selbst wenn die Teilnehmenden auch die kürzesten Seiten ausgefüllt haben, wurden die Seiten mit nur einer Frage darauf doch in den Kommentaren der Studien des Öfteren angekreidet.

Hingegen stehen die Ergebnisse des Experiments im Einklang mit den Erkenntnissen von Wroblewski (2008, S. 85), dass es bei der Formulargestaltung von großer Bedeutung ist einen klaren Pfad vorzugeben, um die Nutzer:innen komfortabel von Anfang bis zum Ende zu geleiten. Denn die gute Strukturierung des OTPP-Formulars wurde mehrfach in den Kommentaren gelobt und schlägt sich auch in den Ergebnissen der Zufriedenstellung nieder.

Ebenfalls konform mit den Aussagen der Kommentarspalten geht die Aussage von Silver (2018, S. 375), dass die Testpersonen das OTPP-Prinzip als motivierend empfinden können, beim Abschließen der kurzen Formularseiten und dies zusätzlich ein Gefühl der Progression und Momentum kreiert. Auch dies kann ein Faktor der höheren Zufriedenstellung im Experiment gewesen sein.