Freie Universität Berlin

Fachbereich Mathematik und Informatik

Bachelorarbeit

im Studiengang Informatik

Thema:	Computer Literacy Skala für erfahrene Nutzer– Entwicklung und Itemanalyse
eingereicht von:	Maria Spiering (maria.spiering@fu-berlin.de, Matr-Nr: 4059687)
eingereicht am:	13. Juli 2010
Gutachter:	Prof. Dr. sc. nat. Hartmut Wandke Prof. Dr. Lutz Prechelt
Betreuer:	Dipl.-Psych. Michael Sengpiel

Eidesstattliche Erklärung

Hiermit versichere ich an Eides Statt, dass die vorliegende Bachelorarbeit von niemand anderem als meiner Person selbst verfasst wurde und dabei keine anderen als die angegebenen Quellen und Hilfsmittel verwendet wurden.
Die Arbeit wurde bisher in gleicher oder ähnlicher Form keiner anderen Prüfungskommission vorgelegt und auch nicht veröffentlicht.

Berlin, den 13. Juli 2010

Maria Spiering

Inhaltsverzeichnis

1 Einleitung **1**
1.1 Computer Literacy . . . 2
1.1.1 Bedeutung und Definition . . . 2
1.1.2 Arbeitsdefinition . . . 4

2 Methode **5**
2.1 Material . . . 5
2.1.1 Computer Literacy Skala . . . 5
2.1.2 Aufgaben- und Skalentypen . . . 10
2.2 Versuchspersonen . . . 11
2.3 Versuchsdurchführung . . . 11

3 Ergebnisse **13**
3.1 CLS-Punkte und Computererfahrung . . . 14
3.2 Itemanalyse . . . 15
3.2.1 Rohwerte . . . 15
3.2.2 Reliabilität . . . 17
3.2.3 Schwierigkeit . . . 18
3.2.4 Trennschärfe . . . 20

4 Diskussion **22**
4.1 Interpretation . . . 22
4.1.1 Rohwerte . . . 22
4.1.2 CLS-Punkte und Computererfahrung . . . 23
4.1.3 Reliabilität . . . 24
4.1.4 Schwierigkeit . . . 25
4.1.5 Trennschärfe . . . 25
4.2 Itemselektion . . . 26
4.2.1 Selektionskennwert . . . 26
4.2.2 Rationale Selektion . . . 27
4.3 Validität . . . 28
4.4 Probleme . . . 29
4.5 Ausblick . . . 30

5 Zusammenfassung **31**

6 Literaturverzeichnis **32**

7 Anhang **34**
32
7.1 Fragebogen 34
7.2 Items 38
7.3 Itemselektion 45
7.4 Anwerberschreiben 47
7.5 verwendete Plattformen 48
7.6 Grafiken 49
7.7 Rohdaten 54
7.8 R-Code 55

1 Einleitung

Seit Mitte der 80er Jahre hat die Bedeutung von Computern in unserem Leben immer mehr zugenommen (Turner, Sweany & Husman, 2000). Diese Entwicklung ist vor allem auf die nachlassenden Preise von Produkten wie Mikrocomputern und dem Personal Computer zurückzuführen (Lockheed, Hunter, Anderson, Beazly & Esty, 1983), wodurch erst größere Teile der Bevölkerung Zugang zu dieser Technologie bekamen. Immer mehr Menschen haben einen privaten Computer (Poynton, 2005) und auch die explosionsartige Verbreitung des Internets hat ihren Teil dazu beigetragen, dass Computer eine so wichtige Rolle in unserer heutigen Gesellschaft spielen und kaum noch wegzudenken sind (Naumann, Richter & Groeben, 2002; Poynton, 2005).
Computer erscheinen uns als unabkömmlich, weil sie längst nicht mehr nur Arbeits-, sondern auch Unterhaltungsmittel sind und sie gewinnen durch ihre Kommunikations- und Informationsfunktionen zunehmend an Einfluss auf fast allen Bereichen der Gesellschaft (Gupta, 2006; Idowu, Adagunodo & Popoola, 2003; Naumann et al., 2002; Turner et al., 2000).
Dieses enorme Potential von Computern und deren Bedeutung erkannte man vor allem im Bildungsbereich sehr früh und so sollten Lehrpläne umstrukturiert werden und fortan das Vermitteln von Computerwissen beinhalten (Cheng & Stevens, 1985; Lockheed et al., 1983). Um einen angemessenen Lehrplan entwickeln zu können, ist es unabdingbar zuvor das schon bestehende Wissen der Schüler und Studenten einzuschätzen und zu messen, weswegen zahlreiche Umfragen und Skalen mit unterschiedlichen Schwerpunkten entwickelt wurden (Hall, 2004; Lockheed et al., 1983; Pyrczak, 1990; Turner et al., 2000).

Wie anfangs schon angedeutet, spielt Computerwissen aber mittlerweile nicht nur im Bildungswesen eine Rolle - auch wenn es dort am effektivsten gelehrt werden kann - sondern beeinflusst alle Bereiche unseres Lebens und ist von ganzheitlichen gesellschaftlichem Interesse (Gupta, 2006). Des Weiteren ist keines der Messinstrumente zur Bestimmung von Computerwissen kurz und objektiv, weswegen unter anderem Sengpiel und Dittberner (2008) auf diesen Mangel reagierten und eine neue Computer Literacy Skala für ältere Menschen entwickelten, die sowohl interaktives Computerwissen - in Form von Symbolen und Begriffen - abfragt als auch eine Selbsteinschätzung beinhaltet. Die entwickelte Skala ist jedoch aufgrund des niedrigen Schwierigkeitsgrades für Menschen mit höherer Kompetenz nicht geeignet (Sengpiel & Dittberner, 2008).

Das Ziel dieser Arbeit ist es demzufolge eine Skala für erfahrene Nutzer, das heißt Menschen mit einem höherem Maß an Computer Literacy zu entwickeln, wobei neue und schwierigere Items benutzt und einer Analyse zu unterzogen werden. Dabei soll

zum Einstieg zunächst der Begriff Computer Literacy und der Test von Sengpiel und Dittberner (2008), der die Basis dieser Arbeit darstellt, genauer erläutert werden. Anschließend wird die Herangehensweise zur Findung der neuen Items als auch die Items selbst grob vorgestellt. Die entstandene Neufassung des Fragebogens (Testvorform), der diese Items beinhaltet, wird dann Versuchsteilnehmern vorgelegt und ausgewertet. Nach abgeschlossener Analyse der Items werden die Items vorgestellt, die weiter verwendet werden können bzw. welche es sich nicht mehr lohnt beizubehalten.

1.1 Computer Literacy

1.1.1 Bedeutung und Definition

Was also ist Computer Literacy?
„Computer Literacy“ ist seit der Einführung des Computers ein weit verbreitetes Konzept, welches besagt, dass man in der Lage ist dem Computer zu sagen was er machen soll, damit er auch das tut was man möchte.

Dies war einer der ersten Vorschläge von Arthur Luehrmann und nur einer von vielen für eine Definition dieses Konzepts (Lockheed et al., 1983). Denn bisher konnte man sich in der Literatur noch auf keine übereinstimmende Definition festlegen (Hall, 2004; Mason & Morrow, 2006; Turner et al., 2000; Vliet, Kletke & Chakraborty, 1994, und andere).

Ein erster naiver Ansatz wäre es den englischen Begriff Computer Literacy mit dem verwandten englischen Begriff für die Lese- und Schreibfähigkeit also die Alphabetisierung (engl. literacy) zu vergleichen. Anhand dessen wäre es dann vielleicht möglich Rückschlüsse für eine Definition zu ziehen. Aber genau dieser Vergleich ist aufgrund der zu unscharfen und uneinheitlichen Definition von Computer Literacy zum Scheitern verurteilt und lässt einen Vergleich der beiden Begriffe gar nicht erst zu. Möchte man diese beiden Begriffe dennoch gegenüberstellen, so zieht dies eine Einführung von neuen, vergleichbaren Kunstbegriffen nach sich (McMillan, 1996).

Dabei ist aber allgemein klar, dass ein gewisser Grad an Computer Literacy als auch an Alphabetisierung in der Gesellschaft von Nutzen ist. Denn durch dieses Wissen fällt es mir zum Beispiel leicht Straßenschilder zu lesen oder eine Fahrkarte am Automaten zu kaufen. Dabei ist es nicht notwendig ein guter Programmierer zu sein oder einen Abschluss in Informatik zu haben, um „Computer Literate“ zu sein und diese Tätigkeiten auszuführen (Gupta, 2006).

Aber wie soll ein Fragebogen und eine Skala für Computer Literacy entwickelt wer-

den, wenn eine Bestimmung des Begriffes nur in einem groben Rahmen möglich ist?

Eine mögliche Herangehensweise ist die bisherigen Definition zu betrachten. In einer Studie von Gupta (2006) ist zu lesen, dass dieser Computer Literacy als Fähigkeit, um bessere Arbeit zu leisten, ansieht. Das heißt Computer Literacy ist nicht losgelöst von seiner Umwelt, sondern steht in einer Beziehung zu einem Job oder einem anderen Umfeld und zu der Person, die dieses Wissen verwendet. Auf ein ähnliches Ergebniss sind auch Mason und Morrow (2006) in einer Studie gleichen Jahres gekommen. Bei einem Vergleich von Definitionen erkannten sie, dass Veränderungen in der Definition von Computer Literacy parallel zur Weiterentwicklung im Computerwesen verlaufen (Mason & Morrow, 2006).
Das Verständnis von Computer Literacy variiert also von Person zu Person, von Beruf zu Beruf und von Zeit zu Zeit (Lockheed et al., 1983). Wobei die Zeit der Faktor ist, der einer einheitlichen Definition wohl am meisten im Wege steht (Turner et al., 2000), denn die Geräte (Hardware) und grafischen Oberflächen (Software), die wir benutzen werden sich immer verändern und weiterentwickeln. Weswegen sich wohl auch eine Definition von Computer Literacy kontinuierlich ändern wird (Van Dyke, 1987).

Wie auch bei Cheng und Stevens (1985) könnte an dieser Stelle zu einem Hilfsmittel gegriffen und eine zeitgemäße „Arbeitsdefinition“ für eine bestimmte Zielgruppe nach der Bottom-Up-Vorgehensweise erstellt werden (Cheng & Stevens, 1985).
Bei dieser Vorgehensweise werden zuerst zentrale Themen identifiziert und anschließend nach Wichtigkeit bewertet. Die wichtigsten Themen bilden anschließend die Grundlage für eine vorläufige Definition. Aus dieser Arbeitsdefinition werden dann der Fragebogen und dessen Items entwickelt.
Cheng und Stevens (1985) haben in ihrer Studie ermittelt, dass die „Bedienung eines Computers“ als wichtigstes und das Wissen um die „Geschichte des Computers“ als unwichtigstes Thema herausstechen. Alle weiteren potentiell bedeutenden Themen sind nach der Bewertung sehr eng beieinander und liefern somit keine eindeutige Aussage (Cheng & Stevens, 1985).
Im Gegensatz dazu sind für Mason und Morrow (2006) „Bewusstsein“ und „Fähigkeiten“ zentrale Komponenten von Computer Literacy, wobei jedoch der Themenkomplex „Bewusstsein“ unter anderem die Geschichte des Computers beinhaltet und der Bereich zur Förderung der Fähigkeiten eher praktisch ausgerichtet ist und sich nicht für einen Fragebogen eignet.
Ebenso praktisch als auch Microsoft orientiert sind die Themen, die Gupta (2006) als fundamentales Wissen einstuft.

Da sich mit diesem Verfahren auf Grund der vielen aufkommenden Widersprüche, die

sich vermutlich auf das Fehlen einer einheitlichen Definition zurückführen lassen, keine grundlegenden Themen finden lassen, wird die Definition von Sengpiel und Dittberner (2008) als Leitfaden dienen.

1.1.2 Arbeitsdefinition

Die von Sengpiel und Dittberner (2008) formulierte Definition für die Computer Literacy Skala (CLS) betrachtet zwar nur einen kleinen aber dafür wesentlichen Aspekt des Computer Literacy Konzepts und benutzt diesen als Indikator für die tatsächlich vorhande Computer Literacy einer Person.

Die von ihnen aufgestellte Definition, soll auch als Grundlage dieser Studie dienen:
„Wenn Alphabetisierung als die Fähigkeit Symbole lesen und benutzen zu können angesehen wird, dann kann Computer Literacy als Fähigkeit computerbezogene Symbole, funktionale Elemente und interaktive Muster zu verstehen und zu benutzen, verstanden werden“ (Sengpiel & Dittberner, 2008, S. 2).

Nachteil dieser Definition ist, dass sie keine thematischen Vorgaben macht. Andererseits legt sie dafür fest, welche Arten von Items der Fragebogen beinhalten soll.

Dabei kann man beispielsweise das Fragezeichen als computerbezogenes Hilfe-Symbol, Knöpfe und Tasten als funktionale Elemente und Checkboxes als interaktive Muster verstehen.
Neben der durch die Definition festgelegten Vorgaben zu den Itemtypen (dazu mehr in Kap. 2.1.2), enthielt der Fragebogen von Sengpiel und Dittberner (2008) eine Struktur, die sich aus zwei Komponenten zusammensetzt:

- Einen Teil zur Selbsteinschätzung der eigenen Erfahrung und Nutzungshäufigkeit des Computers für unterschiedliche Aufgaben.
- Einen Teil zur Erhebung von theoretischem Computerwissen. In diesem sollten computerbezogene Begriffe wie zum Beispiel Tooltip den entsprechenden Erklärungen zugeordnet werden.

Diese Komponenten werden ebenfalls in den Fragebogen mit aufgenommen, obwohl sie in der Definition nicht explizit mit aufgelistet ist.

2 Methode

2.1 Material

2.1.1 Computer Literacy Skala

Aufbau und Entwicklung
Sengpiel und Dittberner (2008) haben in einer Studie bereits eine Definition als auch einen Fragebogen ausgearbeitet. Beides stellt die zentrale Grundlage für den neuen Fragebogen dar.
Das heißt, dass der Aufbau in zwei Teilen und auch der jeweilige Aufgabentypus beibehalten und in den neuen Fragebogen mit einfließen werden.
Verändert werden:

1. die Anzahl der abgefragten Items. Dies ist vor allem sinvoll, weil es sich bei dieser Arbeit um einen einzelnen Iterationsschritt in dem Entwicklungsprozess des Fragebogens handelt.
   Zunächst werden also eine Reihe von Items in den Fragebogen mit aufgenommen, um am Ende des Schrittes - also nach der Analyse aller Items - festzustellen, ob sich ein Item zur Messung von Computer Literacy eignet oder nicht. Es kann also in einem zukünftigen Iterationsschritt beibehalten, ersetzt oder entfernt werden. Es ist also von Vorteil mit mehr als unter Umständen nötigen Items zu beginnen.

2. die Items selbst. Der Grund dafür liegt darin, dass die Schwierigkeit des Fragebogens erhöht werden soll. Bricht man dieses Problem herunter, gilt dies im einzelnen dann auch für die Items.

Inhaltlich baut sich der Fragebogen wie folgt auf:
Der erste von zwei Teilen, Teil A (Erfahrung mit Computern), beinhaltet eine Selbsteinschätzung der Teilnehmer. Die zentralen Fragen sind hier: Wieviel Erfahrung habe ich bereits mit Computern und in welchen Bereichen habe ich diese Erfahrung?
Der umfangreichere Teil B beschäftigt sich mit dem eigentlichen computerbezogenem Wissen des Teilnehmers (Computer Literacy). Es müssen Begriffe, Symbole und ähnliches den entsprechenden Bedeutungen zugeordnet werden. Teil B ist also der eigentliche Computer Literacy Test.

Die Entscheidung für einen gleichbleibenden Aufbau ist vor allem der Überlegung zu besseren Vergleichbarkeit zuzuchreiben. So erhoffte ich mir durch eventuelle Unterschiede oder Gemeinsamkeiten bei den Ergebnissen Fehler im neuen Fragebogen aufzudecken oder Bestätigung für diesen zu finden.

Enwickelt wurde der Fragebogen zunächst als Papierversion (siehe Anhang 7.1). Diese

Form der Entwicklung ermöglichte es Änderungen schnell umzusetzen, erste Tests an dem Prototyp zeitnah vorzunehmen und aufgedeckte Probleme sofort zu beseitigen.
Der Prototyp selbst wurde dreimal überarbeitet bis die endgültige Online-Version fertig war.
Die Umsetzung der fertigen Papierversion in einen webbasierten Fragebogen wurde im Rahmen einer Diplomarbeit von Nico Zeißig vorgenommen (Zeißig, 2009). Der Fragebogen war vorerst unter der Adresse `http://macs2.psychologie.hu-berlin.de/ALISA/clsplus.seam` und später unter `http://141.20.68.27/ALISA/clsplus.seam` erreichbar.

Teil A: Erfahrung mit Computern
Im ersten Teil des Fragebogens hat sich der Teilnehmer zu drei unterschiedlichen Erfahrungsbereichen selbst einzuschätzen. Wobei sich die Erfahrung mit Computern in Dauer, Intensität und Vielfalt gliedert und wie folgt erfragt wird:

Dauer Seit wie vielen Jahren benutzen Sie einen Computer?

Intensität Wie viele Stunden pro Woche benutzen Sie normalerweise einen Computer?

Vielfalt Wie oft benutzen Sie einen Computer für folgende Tätigkeit?
(Anmerkung: Als Tätigkeiten sind unter anderem Textverarbeitung, Tabellenkalkulation, Computerspielen und Programmieren aufgeführt.)

Bis zu diesem Punkt gleicht der neue Fragebogen seinem Vorgänger. Dieser wiederum basiert auf dem „Inventar zur Computerbildung“ (Naumann, Richter & Groeben, 1999). Das heißt auch die metrischen Skalen für die Dauer und Intensität bleiben erhalten. Gleiches gilt für die Berechnung der Vielfalt: Die Häufigkeiten der einzelnen Tätigkeiten werden mit den Werten nie = 0, selten = 1, gelegentlich = 2 und oft = 3 aufsummiert.

Die metrischen Angaben für Dauer und Intensität waren in anderen Studien keine guten Prädiktoren für Computerwissen (Sengpiel & Dittberner, 2008; Arning & Ziefle, 2008), werden aber dennoch mit in den Fragebogen aufgenommen.
Bei der Studie von Sengpiel und Dittberner (2008) könnte der Grund dafür darin liegen, dass das Alter der Studienteilnehmer sehr unterschiedlich war. Dies führt dazu, dass jüngere Teilnehmer auf Grund ihres Alters grundsätzlich eine geringere Dauer an Computererfahrung (in Jahren) haben als ältere, obwohl sie unter Umständen mehr Computer Literacy besitzen. Bei Arning und Ziefle (2008) wurden ausschließlich ältere Menschen getestet, weswegen über die Nutzungsdauer und -intensität der jüngeren Generation keine Aussage getroffen werden kann.
Die Vielfältigkeit der Computernutzung war in beiden eben genannten Studien der beste Prädiktor für Computer Literacy (Sengpiel & Dittberner, 2008) und wird im neuen

Test um zwei Items erweitert: Zu den bisherigen Items kommen die zwei Tätigkeiten „Chat“ und „VoIP“.

Die Wahl auf diese Tätigkeiten fiel aus zwei Gründen: Erstens handelt es sich um internetbasierte Tätigkeiten, die häufig von jüngeren und von Menschen mit höherer Computer Literacy benutzt werden und zweitens waren sie bereits zuvor einmal Teil des Fragebogens (Sengpiel, 2007).

Teil B: Zuordnung von Symbolen und Begriffen

Für die Entwicklung diesen Teils standen an erster Stelle folgende Überlegungen:
Wenn der zuvor entwickelte Wissenstest für Menschen mit mehr Computer Literacy zu einfach war, dann muss die Zuordnung der Items zu einfach gewesen sein. Doch was macht ein Item zu einem schweren oder leichtem? Wann genau fällt es schwer ein Item seiner Bedeutung zuzuordnen?

Mit diesen Vorüberlegungen bin ich an die Auswahl von Items herangegangen und habe zunächst einige Kriterien für deren Schwierigkeitsgrad aufgestellt. Wobei mir durchaus bewusst ist, dass die folgenden Kriterien kein Garant für den Schwierigkeitsgrad sind, sondern dieser erst nach statistischer Auswertung festgestellt werden kann.
Nichtsdestotrotz habe ich mir überlegt, dass ein Element potentiell schwieriger ist, wenn es ...

> ... neu und deshalb noch nicht weit verbreitet/ bekannt ist.
>
> ... nicht der deutschen Sprache entspringt.
>
> ... anderen Symbolen bzw. Begriffen ähnlich ist und so eine Unterscheidung schwerer fällt.
>
> ... nur in bestimmten Umfeldern benutzt wird.
>
> ... optisch bekannt, aber terminologisch unbekannt ist (oder umgekehrt).

Betrachtet man diese - unter Umständen unvollständige - Auflistung, erscheint es als ein leichtes Items zu finden, die mindestens einem Kriterium entsprechen.
Nehmen wir also an, wir haben ein solches Element gefunden und viele Menschen haben Probleme dem Element seine Bedeutung zuzuordnen. Dann ist trotz der gewollten Schwierigkeit noch offen wie relevant dieses Element überhaupt im Alltag ist.

Zur genaueren Erläuterung der Problematik soll der Begriff „Datenbank“ als Beispiel verwendet werden.

Der Begriff scheint beim ersten Lesen nicht wirklich in eine der oben genannten Kategorien zu passen, da er weder neu ist noch aus einer anderen Sprache kommt und viele Leute das Gefühl haben zu wissen was sich dahinter verbirgt. Es kann aber recht schnell gezeigt werden, dass man ihn eigentlich sogar gleich in zwei Kategorien einteilen kann (c und d):
Würde man zusätzlich zu diesem Begriff noch nach einem weiteren Fragen, z.B. „Datenbanksystem", so bedingt die wortverwandtschaft die Einteilung in Kategorie c. Zusätzlich wird dieser Begriff eher in der IT verwendet und dort auch klar von den Begriffen „Datenbanksystem" und „Datenbankmanagementsystem" abgegrenzt.
Ohne die einzelnen Begriffe genauer zu erklären und deren Unterschiede hier darzulegen, ist hoffentlich deutlich geworden, dass es sich selbst bei einem „gefühlt leichtem" Begriff wie „Datenbank" um ein schwieriges Item handeln kann.
Außerdem sollte ohne zusätzliche Erläuterung klar sein, dass es für den Alltag nicht notwendig ist, genau zu wissen was eine Datenbank ist, geschweige denn klar sein muss, was der Unterschied zu einem „Datenbanksystem" und einem „Datenbankmanagementsystem" ist.
Mit diesem Beispiel sollte verdeutlicht werden, dass die Wahl eines schwierigen Items nicht trivial ist. Des Weiteren sollte bei der Wahl der Items im Vordergrund stehen, dass Computer Literacy bedeutet, genügend Wissen und Fertigkeiten zu besitzen, um mit der Technologie der Computergesellschaft zusammenzuleben (Major & Prueher, 1984).

Es reicht also nicht aus so komplizierte Items wie möglich zu erfragen. Es ist vielmehr so, dass die Wahrscheinlichkeit auf eines dieser Elemente im Alltag zu treffen sehr hoch sein sollte. Die Items also einen alltäglichen „gesellschaftlichen" Nutzen haben. Sei es um mit dem Wissen über dieses Item im Beruf effektiver zu arbeiten (Gupta, 2006), E-Learning Plattformen für sein Studium zu nutzen (Hall, 2004) oder um einfach in der Lage zu sein sich im Alltag zurechtzufinden.
Um diesen übergreifenden Anspruch an den Items gerecht zu werden, habe ich nach Elementen gesucht, die bei elektronischen Geräten oder Interfaces unterschiedlichster Art verwendet werden. An je mehr Geräten sich dieses Element befindet, desto höher ist die zuvor angesprochene Wahrscheinlichkeit auf dieses Element zu treffen.
Darauf aufbauend habe ich die Frage nach Symbolen aus dem Fragebogen von Sengpiel und Dittberner (2008) wieder aufgegriffen und thematisch unterteilt. So gab und gibt es in dem Fragebogen mit Symbolen versehene Tasten als auch softwarbasierte interaktive Muster, die häufig in Interfaces auftreten.
Im Folgenden sind die einzelnen Teile grob beschrieben. Ein Teil kann aus mehreren Bereichen bestehen, wobei jeder Bereich genau ein Distraktor-Item enthält.
Detailierte Informationen zu den insgesamt 58 Items sind im Anhang (Kap. 7.2) zu

finden.

Im ersten Teil geht es darum optisch bildliche Elemente in Form von Icons, Symbolen auf Tasten oder interaktive Muster wiederzuerkennen und benennen zu können. Dabei ist jeder Bereich dieses Teils durch eine Box abgegrenzt und beinhaltet wie eben beschrieben andere Typen von Elementen. Diese Arten von Items sind gegenüber dem Fragebogen von Sengpiel und Dittberner (2008) keine Neuheit. Einzige Neuheit ist thematisch der erste Bereich dieses Teils.
Bereich a) fokusiert thematisch Möglichkeiten der Vernetzung und Datenübertragung zwischen Geräten. Dabei fällt das Kensington-Schloss zum Schutz vor der Entwendung von portablen Geräten raus und dient als Distraktor ohne selbst das Distraktor-Item zu sein.
In Bereich b) sind ausschließlich Tasten von Geräten wie Computer-Tastaturen bis hin zu Knöpfen handelsüblicher Kameras abgebildet.
Der letzte Bereich c) des ersten Teils enthält interaktive Muster wie zum Beispiel das häufig verwendete Drop-Down-Menü oder auch das eher selten anzutreffende Pie-Menü.

Teil zwei besteht aus zwei neuen Bereichen, die es zuvor nicht gab.
Bereich a) greift eine Idee von Arning und Ziefle (2008) auf, die in ihrem Fragebogen zur Erfassung von Computerkenntnissen in Frage 16 danach gefragt haben, was passiert, „wenn [man] während der Arbeit an einem Dokument die Tastenkombination „Strg“ + „V“ „drückt (Arning & Ziefle, 2008). Verändert wurde der Ausdruck „Strg“, da diese Taste nur auf der Windows-Tastatur existiert und bei Apple als die Befehlstaste bekannt ist (*Mac OS X: Tastenkombinationen*, 2007). Zur zusätzlichen Erschwernis wurde der Kontext, in dem die Tastenkombination verwendet wird, weggelassen.
Im Bereich b) sind Dateiendungen den entsprechenden Bedeutungen zuzuordnen. Diese Form der Abfrage im Kontext von Computer Literacy habe ich in keiner Literatur gefunden. Nichtdestotrotz halte ich sie aus folgendem Grund für sinnvoll: Wenn ich in der Lage bin eine Abkürzung einem Programm zuzuordnen oder ich zumindest weiß, worum es sich bei diesem Datei-Typ handelt, kann sich ein potentielles Problem um ein vielfaches vereinfachen. Dabei ist dieses Wissen sowohl für die Benutzung am Computer, als auch mit dem Handy im Arbeitsumfeld oder im privaten Bereich einsetzbar.

Teil drei der Item-Gruppen ist thematisch ebenfalls aus dem Fragebogen von Sengpiel und Dittberner (2008) übernommen. In den drei zugehörigen Bereichen werden jeweils Begriffe der Computer-Terminologie abgefragt.

2.1.2 Aufgaben- und Skalentypen

Wie eingangs bereits erwähnt, bleiben die Aufgabentypen und dementsprechend auch die Skalentypen dieselben wie die bei der von mir verwendeten Basis für diesen Fragebogen (Sengpiel & Dittberner, 2008).

Teil A: Erfahrung mit Computern

Dauer metrische Werte

Intensität metrische Werte

Vielfalt vier Punkte Likert-Skala mit folgenden Angaben:
0 = nie, 1 = selten, 2 = gelegentlich, 3 = oft
mit einer maximalen Punktzahl von $13 \cdot 3 = 39$

Nur bei der Abfrage der Vielfalt könnte man von einem Aufgabentypus reden, den man als Stufen-Antwort-Aufgabe bezeichnet. Dieser ist Teil der gebundenen Aufgabentypen und bei diesem „wird jede Antwort gewertet und gewichtet“, zusätzlich wird bei der Auswertung „für jede Antwort eine entsprechende Punktzahl [..] für den Gesamttestpunktwert verrechnet“ (Lienert & Raatz, 1994, S. 19).
Eine Diskussion darüber wie sinnvoll es ist die Nominal- zu einer Ordinalskala mit Ausprägung (Likert-Skala) zu machen und die damit verbundenen - und vor allem für jede Tätigkeit gleichgewichteten - Werte zu addieren, soll an dieser Stelle mit der Begründung von Sengpiel und Dittberner (2008) abgekürzt werden: „Diese [Werte] werden als Vergleich zum objektiven Wissentest in Teil B heranzgezogen“ (Sengpiel & Dittberner, 2008). Sie sind also nicht Bestandteil des eigentlichen Tests, sondern dienen vielmehr als Prädiktor für die Ergebnisse im Wissentest.

Teil B: Zuordnung von Symbolen und Begriffen
Das in Teil B abgefragte Wissen zur Bestimmung eines Aspekts von Computer Literacy wird über die Summe der richtigen Antworten berechnet. Für jede richtige der insgesamt 58 Zuordnungen gibt es einen Punkt, was zu maximal 58 Punkten führt. Es handelt sich hierbei um eine Verhältnisskala und die hier berechneten Punkte werden im weiteren Verlauf als CLS-Punkte bezeichnet.

Bei allen acht - durch umschließende Rahmen getrennten - Teilaufgaben handelt es sich um gebundene Aufgaben vom Typ Zuordnungs-Aufgabe (ZO-Aufgabe). „Hier müssen die beiden Elemente einer Aufgabe - Problem und Lösung, Frage und Antwort - zusammengefügt, einander zugeordnet werden“ (Lienert & Raatz, 1994, S. 19). Dabei gibt es je Teilaufgabe keine vollständige Zuordnung, da sonst die Lösung der letzten

Zuordnung eines Items durch die Lösung aller vorhergehenden Zuordnungen notwendigerweise eindeutig bestimmbar wäre. Aus diesem Grund gibt es in jeder Teilaufgabe ein Element, welches keinen Zuordnungspartner hat (Distraktor).

Vorteile dieses Aufgabentyps sind vor allem, dass a) ein Test „schnell und objektiv über einen ausgedehnten Bereich elementaren Tatsachenwissens“ informieren kann und außerdem b) eine „große Anzahl von Aufgaben [...] auf einer relativ kleinen Fläche untergebracht werden“ kann, sofern es sich um verbale Aufgaben handelt (Begriffs-/ Bedeutungszuordnung) (Lienert & Raatz, 1994, S. 27).

2.2 Versuchspersonen

Den erstellten Fragebogen haben insgesamt 81 Versuchpersonen vollständig ausgefüllt und abgeschlossen. Zur Analyse der Items wurde jedoch nur ein Teil der Datensätze verwendet.
Damit sich der Fragebogen - wie zuvor gefordert - an erfahrenere Nutzer richtet und schwerer wird, müssen auch die Daten die Versuchspersonen dieser Zielgruppe entsprechen. Aus diesem Grund sind nur die Daten der Versuchspersonen, die mindestens ein Abitur als Abschluss haben, in die Auswertung mit eingegangen.
Ausgeschlossen wurden dadurch Teilnehmer mit einem „MSA“ (Mittlerer Schulabschluss), einer „Fachhochschulreife“, einem Abschluss an einer „Fachoberschule“ oder vergleichbarem. Genommen wurden Teilnehmer mit „Abitur“, „Bachelor“, „Master“, „Diplom“, „Magister“, „Vordiplom“ oder einem „Doktor“ jeglicher Art (*Akademischer Grad*, o. J.).

Durch diese Selektion reduzierte sich die Zahl der Versuchspersonen auf 70 (M = 28, SD = 9), die die 58 Items bearbeitet haben. Das Alter bewegt sich zwischen 19 und 67 Jahren mit einem Durchschnittsalter von 28 Jahren. Teilgenommen haben insgesamt 41 weibliche und 29 männliche Teinehmer. (Für eine vollständige Altersverteilung aufgeschlüsselt nach Geschlecht siehe Anhang 7.6 Abbildung 4.)
Von den selektierten Versuchspersonen haben 52 der 70 Versuchspersonen das Abitur als höchsten Schulabschluss. Einen Diplomabschluss haben 9, einen Bachelor haben 3, einen Master oder ein Vordiplom und einen abgeschlossenen Doktor oder Magisterstudiengang haben jeweils zwei bzw. eine Versuchsperson (siehe Anhang 7.6 Abbildung 5).

2.3 Versuchsdurchführung

Die Versuchspersonen wurden ausschließlich über Plattformen im Internet gewonnen, da der Fragebogen ebenfalls webbasiert ist. Demzufolge erfolgte die Bearbeitung des

Fragebogens auch ohne Ein- und Anweisung, sondern völlig frei.
Ein verfasstes Anwerberschreiben bzw. eine minimal abgeänderte Kurzfassung davon wurde auf folgenden Plattformen veröffentlicht:

Foren , die von Studenten genutzt werden (*Über 100 Social Networks aus Deutschland*, 2008)

studiVZ Gruppen

PESA Probandenserver des Instituts für Psychologie der Humboldt Universität Berlin

Eine vollständige Liste der verwendeten Foren, studiVZ-Gruppen und das Anwerberschreiben kann im Anhang 7.5 und 7.4 gefunden werden.

3 Ergebnisse

Für die Analyse der erfassten Daten wurde die für Statistik und stochastische Simulation entwickelte Programmiersprache und das gleichnamige Software-System R verwendet (Sawitzki, 2008). Da es hier aber hauptsächlich um die erfassten und ausgewerteten Daten geht und nicht um den in R verfassten Programmcode, wird dieser nicht im Fließtext zu finden sein. Um die Berechnungen kontrollieren und nachvollziehen zu können, ist der vollständige Programmcode auf der beiliegenden CD zu finden. Die in R programmierten Formel-Berechnungen der einzelnen Itemkennwerte sind im Anhang 7.8 vorgestellt.

Was wurde also analysiert?
Zur Beantwortung dieser Frage, möchte ich einen kurzen Exkurs geben, wie ein angestrebter idealer Test möglichst aussehen soll. Dieser Exkurs soll dazu dienen zu verstehen, welche Teile des Tests warum analysiert wurden und warum diese wichtig sind.

Ziel ist es mit der Analyse ungeeignete Aufgaben bzw. Items zu filtern, weil diese zum Beispiel nicht für die Messung von Computer Literacy geeignet sind. Nach der Selektion ist es wünschenswert einen Test zu erhalten, der in „möglichst hohem Grade objektiv, reliabel und valide ist. Diese Qualitäten kann ein Test nur dann besitzen, wenn auch die einzelnen Aufgaben [diese Kriterien erfüllen]“ (Lienert & Raatz, 1994, S. 29).
Es ist dementsprechend wichtig zu wissen, was genau die Kriterien sind und in welcher Beziehung die einzelnen Aufgabenkriterien zueinander stehen:

Objektivität	→	Reliabilität	→	Validität
Schwierigkeit	→	Trennschärfe	→	

Um die Qualität der einzelnen Aufgaben und damit auch des ganzen Tests zu überprüfen, werden im Anschluss sowohl Reliabilität als auch Schwierigkeit und Trennschärfe statistisch analysiert (siehe Kap. 3.2). Das Kriterium der Objektivität wird nicht speziell überprüft werden. Zur Erläuterung warum und wie dies zu vertreten ist, ohne dabei die Kriterien-Kette *Objektivität* → *Reliabilität* → *Validität* zu verletzen, folgt die Definition der Objektivität:
„Eine Aufgabe ist dann *objektiv*, wenn ihre Beantwortung von verschiedenen Beurteilern übereinstimmend als richtig oder falsch, als kennzeichnend für das Vorhandensein oder Fehlen bzw. einen stärkeren oder schwächeren Ausprägungsgrad des untersuchten Persönlichkeitsmerkmals gewertet wird.“ (Lienert & Raatz, 1994, S. 29; Hervorhebung im Orginal)
Anhand dieser Definition greife ich darauf zurück, dass die Auswahl der Items von den zwei Beurteilern Prof. Dr. sc. nat. Hartmut Wandke und Dipl.-Psych. Michael Sengpiel

für in Ordnung befunden wurde.

Nichtsdestotrotz soll uns auch interessieren, wieviel Erfahrung die Versuchspersonen im Umgang mit Computer haben. Hierbei spielen die Dauer, Intensität und Vielfalt der Computernutzung eine Rolle (siehe Kapitel 3.1). Im gleichen Kapitel ist außerdem auch eine Zusammenfassung der Ergebnisse von „Teil B: Zuordnung von Symbolen und Begriffen“ des Fragebogens zur Erfassung der CLS-Punkte zu finden:

3.1 CLS-Punkte und Computererfahrung

Der Versuchspersonen haben im Bereich Computer Literacy - Teil B des Fragebogens - im Durchschnitt ca. 42 der maximal 58 möglichen Punkte erreicht (Median = 43, Min = 14, Max = 58, SD = 10). (Für eine grafische Verteilung der Punkte siehe Anhang 7.6 Abbildung 6.)
Bei Sengpiel und Dittberner (2008) erzielte die jüngere Gruppe, für die der Test zu einfach war, im Mittel 92% der möglichen Punkte (M = 23.9 von maximal 26 möglichen Punkten), wobei hier die Versuchsperonen im Durchschnitt nur 72% der Punkte erreichten.

Im Bereich Erfahrung mit Computern ergab sich für die Versuchspersonen eine Nutzungsdauer von im Mittel 13 Jahren (Min = 5, Max = 25). Weiterhin ergab sich eine Intensität von durchschnittlichen 31 Stunden Computernutzung pro Woche (Min = 4 Stunden, Max = 100 Stunden).
Im Vergleich zu Sengpiel und Dittberners Ergebnissen (2008) in denen die jüngeren Probanden im Schnitt 10 Jahre (Min = 3, Max = 17) - also 3 Jahre weniger - den Computer genutzt haben und im Mittel 27 Stunden pro Woche (Min = 6, Max = 50) den Computer verwenden.

Zusätzlich wurde mit einer 4-Punkt-Likert-Skala für 13 Items erfasst, wie vielfältig die Versuchspersonen den Computer nutzen. Die Vielfalt der Nutzung in Punkten reichte von minimal 12 bis maximal 34 in einem Wertebereich von 0 bis 39 Punkten (M = 24, Median = 25, SD = 5). Die interne Konsistenz der Items liegt zwischen .62 und .69 (M = Median = .65). Die Popularität (Schwierigkeit) der Vielfalt-Items ist im Durchschnitt 1.88 (Min = .66, Max = 2.97), wobei der Computer am Meisten zum E-Mails schreiben (99%), zur Informationssuche und zum „Surfen“ im Internet (97%), und am wenigsten zum Programmieren (22%) verwendet wird (siehe Abbildung 8 im Anhang 7.6).
Im Vergleich zu den Ergebnissen von Sengpiel und Dittberner (2008) nutzen die Versuchspersonen dieses Tests den Computer zwar ähnlich (Textverarbeitung = 82%; Sur-

fen und E-Mail = 78%; Programmierung = 9%) , aber er wird, sofern man lediglich einen prozentualen Vergleich der Vielfalt erhebt, wesentlich mehr genutzt. Dieser Vergleich bezieht bei Sengpiel und Dittberner (2008) sowohl deren Subgruppe der Älteren (50 Jahre und älter) als auch die jüngere Untergruppe (unter 50 Jahre) mit ein. Vergleicht man also die durchschnittlich erzielten Vielfalt-Punkte der jüngeren Gruppe (61%) mit den Punkten der Versuchspersonen dieses Tests (62%), stellt sich heraus, dass beide Probanden-Gruppen den Computer ähnlich vielfältig nutzen.
Dieser Vergleich der beiden Gruppen ist durchaus sinvoll, da sich diese beiden sowohl vom Alter als auch von der Nutzung her sehr ähnlich sind.

3.2 Itemanalyse

3.2.1 Rohwerte

Rohwertermittlung
Die im Kapitel 3.1 dargelegten Punkte mussten im Rahmen der Rohwertermittlung erst berechnet werden. Dabei gibt es verschiedene Möglichkeiten den Rohwert (Punkte) für die Versuchspersonen zu ermitteln.

Teil A des entwickelten Fragebogens enthält 13 Items mit einer 4-Punkt-Likert-Skala zur Erfassung der vielfältigen Nutzung von Computern und Teil B beinhaltet 58 Items. In Teil A gab es je nach Antwort von 0 bis 3 Punkte (nie = 0, selten = 1, gelegentlich = 2, oft = 3).
Im Teil B gab es für die richtige Beantwortung eines Items einen und für eine falsche Beantwortung keinen Punkt. Für jede richtige Antwort wurde als ein Punkt vergeben, das heißt der Rohwert X_i einer Versuchsperson i ist demnach wie folgt definiert (Lienert & Raatz, 1994):

$$\begin{aligned} X_i =& R_i \\ &\text{mit} \\ &\quad X_i = \text{Rohwert der Versuchsperson i} \\ &\quad R_i = \text{Anzahl der richtigen Antworten} \end{aligned} \tag{1}$$

Für den entsprechenden R-Code siehe Anhang 7.8 Listing 1.

Punkteverteilung
Neben den zu untersuchenden Qualitätskriterien für Items, gibt es noch einen weiteren Punkt, der bei einer vollständigen Itemanalyse beachtet werden muss:

Die Bedeutung der Testpunktverteilung.
Denn nach Lienert und Raatz (1994) soll die Durchführung der Itemanalyse auch darüber Auskunft geben, „ob der Test eine hinreichende Streuung der Punktwerte [...] besitzt und ob die Testpunktwerte annähernd normal verteilt sind.“ (Lienert & Raatz, 1994, S. 58). Auch wenn die „Hinreichende Streuung, Symmetrie und Eingipfligkeit der Rohwertverteilung (wie bei einer Normalverteilung) [...] keine notwendigen Bedingungen eines guten Tests [sind]”(Lienert & Raatz, 1994, S.58), so können diese Angaben bei der Wahl der Items und zur Verbesserung des Tests sehr hilfreich sein.

Was ist also die Testpunktverteilung und warum ist sie für die Analyse wichtig?
Die Testpunktverteilung beschreibt die Verteilung der erzielten Punkte aller Versuchspersonen (siehe Kap. 7.6 Abbildung 6). Dabei können sich die Punkte ganz unterschiedlich verteilen: Es kann sich um eine schiefe, breit-, schmal- oder auch zweigipfelige Punktwertverteilung handeln. Wünschenswert ist jedoch die angestrebte Normalverteilung.

Um diese Verteilung statistisch auf Normalität zu überprüfen, soll zum Einen die Schiefe S und zum Anderen der Exzess E mit einem jeweiligen z-Test berechnet werden (Lienert & Raatz, 1994):

$$
\begin{aligned}
S &= \frac{\sum_i x_i^3}{N \cdot s_x^3} \\
s_S &= \sqrt{\frac{6}{N}} \\
z_S &= \frac{S}{s_S}
\end{aligned}
\tag{2}
$$

mit

S = Schiefe der Verteilung

(die symmetrische Normalverteilung hat eine Schiefe von S = 0)

N = Stichprobemumfang (Anzahl der Probanden)

s_x = Standardabweichung der gewonnen Rohwerte

$\sum_i x_i^3$ = Kubensumme der Abweichungswerte $(X_i - \bar{X})$

s_S = Standardfehler der Schiefe S

$$
\begin{aligned}
E =& \frac{\sum_i x_i^4}{N \cdot s_x^4} \\
s_E =& \sqrt{\frac{6}{N}} \\
z_E =& \frac{E-3}{s_E} \\
&\text{mit} \\
&\quad E = \text{Exzess der Verteilung} \\
&\quad \text{(bei der Normalverteilung ist der Exzess E = +3.0)} \\
&\quad \sum_i x_i^4 = \text{Summe der 4.Potenz von } (X_i - \bar{X}) \\
&\quad s_E = \text{Standardfehler des Exzesses E}
\end{aligned}
\tag{3}
$$

Wie in den Formeln angegeben (R-Code: Listing 2), ist der Wert für die Schiefe bei einer Normalverteilung S = 0 und der normale Exzess liegt bei E = +3.0.
Die errechneten Schiefe- und Exzess-Werte für die Punktwertverteilung des Tests liegen bei S = -.45 (z_S = -1.54) und E = 2.47 (z_E = -1.81).
Laut Lienert und Raatz (1994) muss bei einer geringeren Anzahl von Probanden (unter 400) schon ab einem Signifikanzniveau von 1% von der Annahme der Normalverteilung abgewichen werden. Dies trifft sowohl auf die Schiefe mit einem statistisch signifikantem 1.5%-Niveau als auch auf den Exzess mit einem statistischem 1.8%-Signifikanzniveau zu.
Die Testpunktverteilung entspricht also keiner Normalverteilung, sondern ist im Gegensatz zu dieser eher rechts- und schmalgipfelig.

3.2.2 Reliabilität

„Unter der *Reliabilität* oder Zuverlässigkeit eines Tests versteht man den Grad der Genauigkeit, mit dem er ein bestimmtes Persönlichkeits- oder Verhaltensmerkmal mißt“ (Lienert & Raatz, 1994, S. 9; Hervorhebung im Orginal) ohne die Frage zu beantworten „ob [der Test] auch das mißt, was er messen soll“ (Lienert & Raatz, 1994, S. 9).

Eine von vier Varianten, um den Grad der Reliabilität zu bestimmen, ist die Methode der Konsistenzanalyse. Diese Methode kann verwendet werden, wenn der Test bei einer Stichprobe einmal durchgeführt wird, wodurch man aus den Aufgabenkennwerten - mittels spezieller Rechenformeln - dann einen Schätzwert für die Reliabilität erhält (Lienert & Raatz, 1994). Man spricht von der Analyse der „inneren Konsistenz“ (Lienert & Raatz, 1994) bzw. der „internen Konsistenz“ eines Tests (Bortz, 1999).

Ein großer Vorteil dieser Methode ist es, dass der Test nicht mehrmals wiederholt werden muss, um einen Schätzwert für die Reliabilität zu erhalten.
Bei dieser Art der Reliabilitätsbestimmung wird der Test in zwei (oder mehr) äquivalente Teile geteilt, das Testergebnis für die Teile ermittelt und anschließend werden die Ergebnisse der Testteile miteinander korreliert. So erhält man - unter der Voraussetzung, dass alle Teile dasselbe Merkmal messen - einen Wert für die innere Konsistenz (Lienert & Raatz, 1994).
Wie schon angedeutet hängt die Berechnung der inneren Konsistenz von der verwendeten Rechenformel ab.
In dem Test (Fragebogen Teil B) sind alle Testteile (Items) gleich lang, weswegen es möglich ist die Berechnung mit dem α-Koeffizienten von Cronbach durchzuführen.
Die Besonderheit bei der Verwendung des α-Koeffizienten ist, dass „der Test in soviele Teile aufgeteilt wird, wie er Aufgaben enthält, die je mit 0 oder 1 Punkten bewertet werden“ (Lienert & Raatz, 1994, S. 192). Anders gesagt, werden bei dieser Methode „alle Testitems als eigenständige Tests für ein- und dasselbe Merkmal angesehen [und] die Reliabilität des Summenscores (α) ergibt sich als durchschnittliche Paralleltestreliabilität für alle möglichen Paare von Testitems“ (Bortz, 1999, S. 517).
Die verwendete Formel zur Berechnung der inneren Konsistenz lautet:

$$\alpha = r_{tt} = \frac{c}{c-1} \cdot \left[1 - \frac{\sum_j s_j^2}{s_x^2} \right] \tag{4}$$

mit

c = Anzahl der äquivalenten Testteile

s_x= Standardabweichung des gesamten Tests

s_j= Standardabweichung des Tests ohne Item j

Unter Verwendung der Formel (4) (R-Code: Listing 3) kann für jedes Item errechnet werden, wie sich die innere Konsistenz des Gesamttest verhält, wenn das Item aus dem Test genommen wird.
Die errechnete interne Konsistenz für die jeweiligen Items ist hoch (α = .91 bis .92) und liegt in einem Wertebereich von 0 bis 1 (siehe Abbildung 1).

3.2.3 Schwierigkeit

„**Definition:** Der Schwierigkeitsindex einer Aufgabe ist gleich dem prozentualen Anteil P der auf diese Aufgaben entfallenden richigen Antworten in einer Analysestichprobe von der Größe N; der Schwierigkeitsindex liegt also bei schwierigen Aufgaben niedrig,

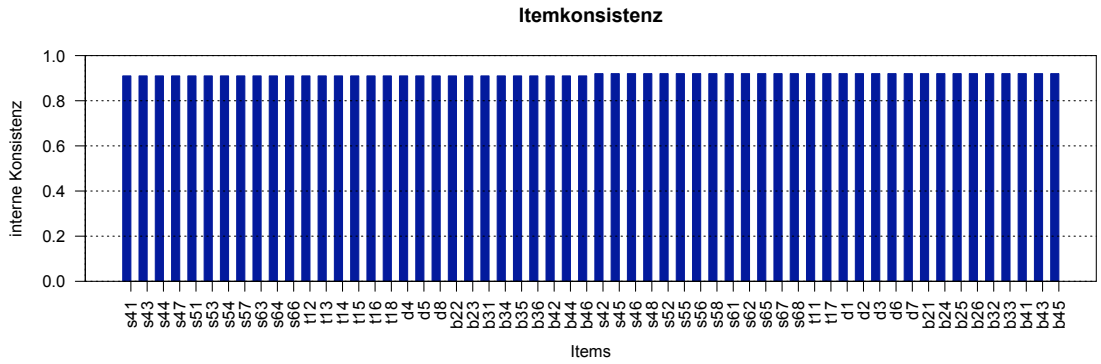


Abbildung 1: Itemkennwert - Innere Konsistenz

bei leichten hoch“ (Lienert & Raatz, 1994, S.73; Hervorhebung im Orginal).
Das heißt, je mehr Versuchspersonen in der Lage sind eine Aufgabe zu lösen, desto leichter ist diese Aufgabe. Und je leichter eine Aufgabe ist und desto größer wird ihr Schwierigkeitsindex.

Die Formel zur Berechnung des Schwierigkeitsindex hängt von der Art des Tests als auch der Art der Aufgabe ab.
Bei der Art des entwickelten Tests handelt es sich um einen „Niveautest“ bei dem es keine Vorschrift bezüglich der Testzeit gab. Außerdem wurden Aufgaben vom Typ Zuordnungs-Aufgabe mit mehr als vier Wahlmöglichkeiten verwendet, weswegen der Zufallseinfluss bei der Beantwortung einer Aufgaben vernachlässigt werden kann.
Resultierend aus diesen zwei Testeigenschaften, ergibt sich die Verwendung der Formel (5) zur Bestimmung des Schierigkeitsindex je Item (Lienert & Raatz, 1994):

$$P = 100 \cdot \frac{N_R}{N} \tag{5}$$

mit

N_R= Anzahl der Pbn, die die Aufgabe richtig beantwortet haben

N = Gesamtzahl der Probanden (Pbn)

Die mit Formel (5) errechneten Schwierigkeitindizes (R-Code: Listing 4) bewegen sich eigentlich in einem Wertebereich von 0 bis 100. Der Wertebereich wurde aber zur besseren grafischen Vergleichbarkeit mit den anderen Itemkennwerten durch 100 geteilt, so dass der neue Wertebereich wie der der anderen Itemkennwerte zwischen 0 und 1 liegt.
Im Mittel liegt die Schwierigkeit der Items bei P = .72 (Median = .76, Min = .23 ,

Max = .99) (siehe Abbildung 2).

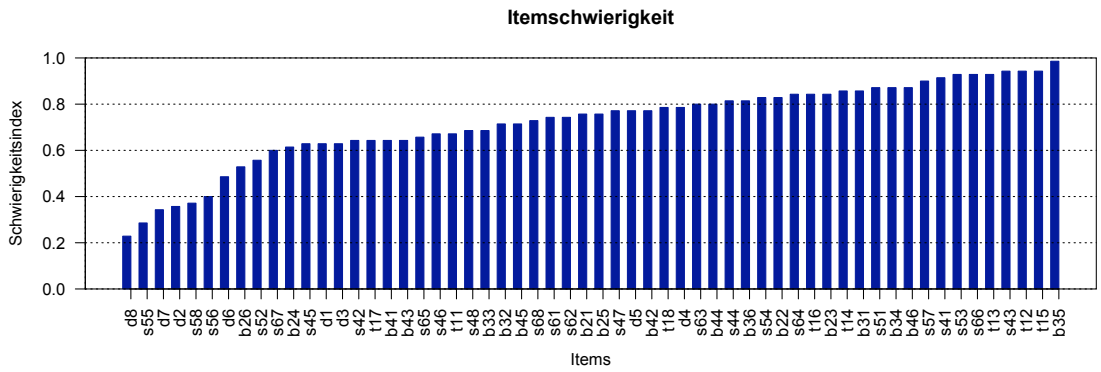


Abbildung 2: Itemkennwert - Schwierigkeitsindex

3.2.4 Trennschärfe

„**Definition:** Der Trennschärfekoeffizient einer Aufgabe j ist gleich der Korrelation zwischen dem Aufgabenpunktwert - meist 0 oder 1 - und dem Rohwert eines jeden von N [Probanden] Pbn der Analysenstichprobe“ (Lienert & Raatz, 1994, S. 78; Hervorhebung im Orginal).
Die Trennschärfe besagt also wie gut eine Aufgabe „bessere“ von „schlechteren“ Versuchspersonen trennt. Das heißt bessere Probanden beantworten diese Aufgabe meist richtig und schlechtere Probanden meist falsch (Lienert & Raatz, 1994).

Auch bei rechnerischen Ermittlung des Trennschärfekoeffizienten gibt es wiederum verschiedene Methoden und Formeln, die verwendet werden können.
Handelt es sich um eine vollständige Aufgabendarbietung und dient der Rohwert (Anzahl der Punkte) als quantitatives Analysekriterium, wie im entwickelten Fragebogen, dann verwendet man die punktbiserale Korrelation zur Berechnung des Zusammenhangs zwischen dem Item und dem Gesamtpunktwert (Lienert & Raatz, 1994):

$$r_{pbis} = \left[\frac{\sum_i X_{Ri}}{N_R} - \frac{\sum_i X_i}{N} \right] \sqrt{\frac{N_R}{N - N_R}} \cdot \frac{N}{\sqrt{N \sum_i X_i^2 - (\sum_i X_i)^2}} \tag{6}$$

mit

N = Anzahl aller Probanden (Pbn)

N_R = Anzahl der Pbn, die die Aufgabe richtig beantwortet haben

X_{Ri} = Testrohwert des Pb_i, der die Aufgabe richtig beantwortet hat

Die mit der Formel (6) errechneten Trennschärfekoeffizienten (R-Code: Listing 5) liegen im Mittel bei $r_{pbis} = .41$ und ordnen sich über alle Items zwischen $r_{pbis} = -.02$ und .68 in einem Wertebereich von -1 bis 1 ein (siehe Abbildung 3).

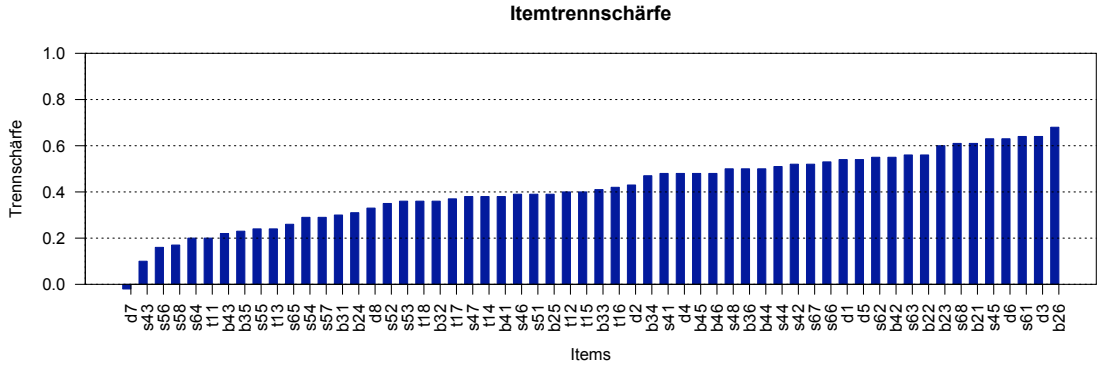


Abbildung 3: Itemkennwert - Trennschärfe

4 Diskussion

Im Folgenden sollen die zuvor lediglich beschriebenen Ergebnisse diskutiert und interpretiert werden. Dabei ist es das Ziel jedes der enthaltenen Items anhand der Ergebnisse, bezüglich der berechneten Itemkennwerte, zu bewerten. Aus dieser Bewertung heraus soll in Kapitel 4.2 für jedes Item entschieden werden, ob dieses ausgewählt (selegiert) oder ausgeschlossen (eliminiert) wird. Das heißt das jeweilige Item wird auch zukünftig Teil des Fragebogens sein oder nicht.
Außerdem soll die Einflussfaktoren Validität des Tests kritisch betrachtet und ein Ausblick bzw. Vorschläge für weiterführende iterative Entwicklungsschritte gegeben werden.

4.1 Interpretation

4.1.1 Rohwerte

Die errechneten Werten für die Schiefe und den Exzess haben ergeben, dass es sich um eine rechts- und schmalgipfelige Verteilung handelt. Die Versuchspersonen erzielen also im Schnitt mehr Punkte und diese sind nicht so weit wie gewünscht über den gesamten Punktwertebereich gestreut.
Ursachen für diese Umstände können folgender Art sein (Lienert & Raatz, 1994):

(a) Die Analysestichprobe kann heterogen sein.

(b) Der Test kann mangelhaft konstruiert sein.

(c) Das untersuchte Persönlichkeitsmerkmal kann anomal verteilt sein.

Dabei sind ursächliche Fehler bei (a) und (b) meinerseits zu verantworten.

Stichprobenheterogenität
Angenommen es gibt so eine Heterogenität innerhalb der Versuchspersonen, so müsste in einem weiteren Schritt mittels einer Regressionsanalyse untersucht werden, welches Merkmal (Alter, Abschluss, Geschlecht usw.) Ursache dieser Heterogenität ist.
Da so eine aufwendige Analyse jedoch den Umfang dieser Bachelorarbeit sprengen würde und sowohl das Anschreiben und als auch die nachträgliche Selektion der Versuchspersonen dazu geführt haben sollte, dass sich nur noch Studenten (52 Vpn mit Abschluss = „Abitur“) und Alumni (18 Vpn) in der Analysestichprobe befinden, werde ich an dieser Stelle diese Ursache zunächst mit Vorsicht ausschließen und auf Ursache (b) als Grund für eine anomale Verteilung verweisen.

Mangelhafte Testkonstruktion
Nach Lienert und Raatz (1994) kann eine nicht vorhandene Normalverteilung auch darin begründet sein, dass eine unzweckmäßige Schwierigkeitsgraduierung vorliegt. Das heißt „die mittlere Aufgabenschwierigkeit (= arithmetisches Mittel der Schwierigkeitsindizes) [ist] sehr gering, also $\bar{P}$ sehr hoch, d.h. [...] der Test als ganzer [ist] zu *leicht* gewesen, [und] so kann eine rechtsgipfelige RW-Verteilung erwartet werden“ (Lienert & Raatz, 1994, S. 155; Hervorhebung im Orginal).
Genau dieser Fall tritt hier ein, denn wie in Kapitel 3.2 (unter Schwierigkeit) vorgestellt, liegt die mittlere Aufgabenschwierigkeit weit über dem Mittel (M = .72). Wodurch es zu einer rechtsgipfeligen Verteilung kommt.
Der Grund für den Mangel an Exzess - also der Breite der Glockenkurve - kann ebenfalls auf diese asymmetrische Form der Kurve zurückgeführt werden. Die erhaltene Verteilungskurve der Punkte wird nach rechts abgebrochen und so kommt es am rechten Ende der Rohwertskala (hohe Punkte) zu einer *Testwerte-Stutzung* (Lienert & Raatz, 1994).

Auch wenn der Ausschluss der Ursache (a) an dieser Stelle fälschlicherweise stattgefunden haben sollte, so hat diese Entscheidung nur geringfügige Auswirkungen auf die Itemselektion. Der Grund dafür liegt darin, dass 74% der Versuchspersonen derselben gewünschten Studenten-Gruppe angehört.

Nichtsdestotrotz ist die Erkenntnis über die zu geringe Aufgabenschwierigkeit für die Itemselektion sehr wichtig und mit zu beachten.

4.1.2 CLS-Punkte und Computererfahrung

Für die weitere Diskussion der Ergebnisse stellt sich nun die Frage, ob es sich bei den Nutzern wirklich um *erfahrene* Nutzer handelt und ob der Test wie gewünscht schwieriger ausfällt.

Ein Punkt der - unabhängig von den Analysen - für die nötige Erfahrung der Nutzer spricht, ist der Umstand, dass die Versuchspersonen bereits in der Lage sein mussten einen Computer inklusive eines Webbrowser und einer von den zuvor erwähnten Plattformen (siehe Kap. 2.3) zu benutzen, um an der Studie überhaupt teilnehmen zu können. Das heißt, die Ergebnisse fallen vermutlich besser aus als wäre die Studie mit der Papierversion - auf der Straße - durchgeführt worden.

Des Weiteren wurden in Kapitel 3.1 zur Beantwortung der obigen Frage die analysierten Ergebnisse bereits vorgestellt und mit den Werten von Sengpiels und Dittberners Arbeit (2008) verglichen.

Dabei zeigte sich, dass die jüngere und erfahrene Gruppe (unter 50 Jahre) der soeben genannten Autoren den Computer im Schnitt 3 Jahre (Dauer) und 3 Stunden pro Woche (Intensität) weniger verwenden als die Versuchspersonen in diesem Test. Beim Vergleich der Vielfalt der Computernutzung erzielten beide Gruppen jedoch sehr ähnliche Ergebnisse (61% bzw. 62%). Dieser Vielfalts-Wert ist auch eher von Interesse, da er im Gegensatz zur Dauer und zur Intensität der Nutzung auch der beste Prädiktor für die CLS-Punkte ist (Sengpiel & Dittberner, 2008). Benutzt also jemand den Computer durchaus vielfältig, so schneidet er auch im Computer Literacy Wissenstest „Zuordnung von Symbolen und Begriffen“ besser als andere ab.
Da sich die beiden Gruppen (jüngere von Sengpiel und Dittberner (2008) und Versuchspersonen hier) auf einem ähnlichen Vielfalt-Niveau befinden, müssten die Gruppen - sofern die beiden Tests gleich schwer sind - auch eine ähnliche Anzahl von CLS-Punkten haben.
Dies ist aber nicht der Fall. Die jüngere Gruppe erreichte wesentlich mehr CLS-Punkte (M = 92%, SD = 9%) als die Versuchspersonen in diesem Test (M = 72%, SD = 17%). Der diesem zugrunde liegende Test von Sengpiel und Dittberner (2008) ist also leichter, was im Umkehrschluss heißt, dass der neu entwickelte Test wie gewünscht schwieriger ausfällt.

4.1.3 Reliabilität

Wie bereits in Kapitel 3.2 dargestellt, liegt die interne Konsistenz der Items bei der einen Hälfte bei $\alpha = r_{tt} = .91$ und bei der anderen bei $r_{tt} = .92$ in einem Wertebereich von 0 bis 1.
Auch wenn die errechneten Werte sehr hoch sind, ist damit aber noch nicht klar, wie diese einzuschätzen sind. Ist das Ergebnis sehr gut oder nur bedingt zufriedenstellend?

Laut Lienert und Raatz (1994) werden gegenwärtig oft übertrieben hohe Anforderungen an den Wert der Reliabilität gestellt, so dass die interne Konsistenz ungefähr bei $r_{tt} = .95$ liegen sollte.
Dieser Wert wurde verfehlt. Nichtsdestotrotz ist der errechnete Reliabilitätswert immer noch sehr hoch. Wie Lienert und Raatz (1994) schrieben, ist der Anspruch übertrieben, da es auch mehr als ausreichend ist, wenn ein Test eine interne Konsistenz von über $r_{tt} \geq .80$ besitzt (Carmines & Zeller, 1979). Dies wird auch durch Cronbach selbst bestärkt, der in seiner Arbeit über den α-Reliabilitätswert schrieb, dass zwar ein hoher Wert angestrebt wird, dieser aber nicht perfekt sein muss (Cronbach, 1951).

Der entwickelte Test und alle beinhaltenden Testitems besitzen also eine hohe und ausreichend gute Genauigkeit zur Bestimmung eines Merkmalsaspektes von Computer Literacy.

Dies hat zur Folge, dass bei der Auswahl der Items dieser Kennwert kein Ausschlusskriterium für ein Item ist. Bei der Itemselektion muss dieses Kriterium also nicht betrachtet werden. Die betrachteten Selektionskriterien sind also alleinig die Schwierigkeit und die Trennschärfe des jeweiligen Items.

4.1.4 Schwierigkeit

Die in Kapitel 3.2 vorgestellten Itemschwierigkeiten der Items liegen in einem Bereich von $P = .23$ bis .99 (M = .72). Dieser Schwierigkeitsindex ist je höher desto leichter ein Item bzw. eine Aufgabe zu lösen ist. Bei einer mittleren Aufgabenschwierigkeit von $P = .72$ sind die Items tendenziell zu leicht, da dieser Mittelwert für eine gute Differenzierung der Versuchspersonen möglichst nah an $P = .50$ liegen sollte (Lienert & Raatz, 1994).
Des Weiteren ist bei Niveautests gewünscht, dass sich „die Schwierigkeitsindizes [...] über den ganzen Bereich des Persönlichkeitsmerkmals [Computer Literacy] streuen, wenigstens jedoch von $P = 20$ bis $P = 80$ [(hier .2 bis .8)]“ (Lienert & Raatz, 1994, S. 115). Ein solche breite Schwierigkeits-Verteilung der Items ist schon annähernd vorhanden - auch wenn einige zusätzliche sehr schwierige Aufgaben durchaus wünschenswert wären - und sollte mit der Itemselektion nicht verloren gehen.

Für die Itemselektion gilt also eine möglichst breite Streuung der Schwierigkeit beizubehalten und etwas mehr leichte (hoher P-Wert) als schwere Items zu eliminieren.

4.1.5 Trennschärfe

Der errechnete Wert der Trennschärfe gibt an, wie gut ein Item „bessere“ von „schlechteren“ Versuchspersonen unterscheiden kann. Je besser die Unterscheidungskraft des Items ist desto höher ist dessen errechnete Trennschärfe r_{pbis}. Die Werte der Trennschärfen je Item bewegen sich in einem Bereich von $r_{pbis} = -.02$ und .68.

Da es das Ziel ist, die Versuchspersonen gut zu unterscheiden, sollte die Trennschärfe so hoch wie möglich sein. Items mit fehlender oder sogar negativer Trennschärfe sind nicht wünschenswert und sollten im Zuge der Itemselektion eliminiert werden (Lienert & Raatz, 1994).
Dabei kann es auf Grund der ebenfalls zu beachtenden Item-Schwierigkeit und dem parabolischen Zusammenhang der Trennschärfe mit dieser durchaus zu Ausnahmen kommen. Diese sollen aber erst im Kapitel 4.2 Itemselektion angeführt und diskutiert werden.

4.2 Itemselektion

Für Auswahl der Aufgaben stellen sich zwei grundlegende Fragen, die es zu beantworten gilt:

(a) „Nach welchen Gesichtspunkten sollen Aufgaben aus der Testvorform ausgeschieden werden?“ (Lienert & Raatz, 1994, S.114) und

(b) „Welche und wieviele Aufgaben sollen aus der Testvorform ausgeschieden werden?“ (Lienert & Raatz, 1994, S.114)

Einige Hinweise zu den Gesichtspunkten aus Frage (a), nach denen entschieden werden soll, wurden schon bei der Diskussion der einzelenen Itemkennwerte gegeben. Warum diese Gesichtspunkte gewählt wurden, hängt von der Art des Tests ab.

Bei dem entwickelten Fragebogen handelt es sich um einen so genannten „Leistungstest“. Genauer gesagt sogar um ein „Niveautest“, da es für die Bearbeitung der Aufgaben keine zeitliche Beschränkung gab (Lienert & Raatz, 1994). Außerdem handelt es sich hier um einen Test homogener Art (siehe Kap. 3.2).
Aus dieser Bestimmung des Testtyps ergibt sich das primäre Kriterium der maximalen Trennschärfe. Zweitwichtigstes Kriterium ist die Schwierigkeit. Dies führt dazu, dass Items mit geringer Trennschärfe nach ihrer Schwierigkeit ausgewählt (selegiert) oder ausgeschlossen (eliminiert) werden. Dieses Vorgehen wird verfolgt, um am Ende „genügend Aufgaben von geringer und sehr hoher Schwierigkeit in der Testendform zu behalten“ (Lienert & Raatz, 1994, S. 115). Damit ist die Vorgehensweise der Itemselection und die eingangs gestellte Frage (a) beantwortet.

Frage (b) besteht aus zwei Teilen, bei der als erstes zu klären ist, wieviele Items am Ende noch übrig bleiben sollen, um als zweites eine explizite Eliminierung der Items durchführen zu können.

Zur Beantwortung der Frage nach der Anzahl der Items, die eliminiert bzw. selegiert werden sollen, verweise ich ein weiteres Mal auf den diesem zugrunde liegenden Test von Sengpiel und Dittberner (2008) bei dem von anfänglich 60 untersuchten Items die Hälfte eliminiert wurden, so dass noch 30 Items übrig blieben. So sollen auch hier von den anfänglich 58 Items 30 am Ende beibehalten werden.

4.2.1 Selektionskennwert

Die soeben beschriebene Anforderung, bei der Selektion von homogenen Niveautest gleichermaßen Trennschärfe und Schwierigkeit zu beachten, hat sich auch die Methode der Itemselektion entwickelt durch Mittenecker, Ebel und Lienert zum Ziel gesetzt

(Lienert & Raatz, 1994). Bei dieser Methode werden diese beiden Itemkennwerte zu einem Selektionskennwert S zusammengefasst, anhand dessen eine Auswahl vollzogen wird. Die in R umgesetzte Formel (Anhang 7.8 Listing 6) zur Berechnung dieses Kennwertes lautet wie folgt:

$$
\begin{aligned}
S_j =& \frac{r_j}{2 \cdot \sqrt{p_j \cdot q_j}} \\
&\text{mit} \\
&\quad S_j = \text{Selektionskennwert des Items j} \\
&\quad r_j = \text{Trennschärfe j} \\
&\quad p_j = \text{Schwierigkeit des Items j P / 100} \\
&\quad q_j = 1 - p_j
\end{aligned}
\tag{7}
$$

„Dieser Kennwert erlaubt eine Selektion nach dem Prinzip, daß Aufgaben mit niedrigem S-Wert ausgeschieden werden können, ohne daß man zu befürchten braucht, es würden auf diese Weise zu viele Aufgaben mit extremer Schwierigkeit verlorengehen“ (Lienert & Raatz, 1994, S. 118).

Nach Berechnung des Selektionskennwertes für jedes Item werden die Items mit den 30 höchsten S-Werten selegiert (siehe Anhang 7.3 Wahl 1).

4.2.2 Rationale Selektion

Betrachtet man die getroffene Auswahl der Items mittels des Selektionskennwertes, so kann man feststellen, dass kaum schwierige Items (niedriger P-Wert) ausgewählt wurden (siehe Anhang 7.3 Wahl 1). Dies hat zur Folge, dass nur leichte Aufgaben in dem neuen Test existieren würden. Dies ist aber nicht wünschenswert, da „der Test eine im Bezug auf die Population mittlere Schwierigkeit besitzen [sollte]“ (Lienert & Raatz, 1994, S. 147).
Aus diesem Grund soll an dieser Stelle zusätzlich eine rationale Selektion erfolgen, wobei Lienert und Raatz (1994) davor warnen „bei der Aufgabenselektion [sich] jedes Perfektionismus [zu] enthalten und ein Fingerspitzengefühl für diese Technik und die hierzu erforderliche Kompromißbereitschaft [zu] erwerben“ (Lienert & Raatz, 1994, S. 116).

Im Zuge dessen habe ich mich dazu entschieden nur die Items (unter den ersten 30 höchsten S-Kennwerten) zu behalten, die eine Trennschärfe von $r_{pbis} \geq .50$ besitzen. Durch diese zusätzliche Anforderung bleiben nichtsdestotrotz 20 der vorherigen

30 Items unverändert selegiert.
Die verbleibenden 10 Item-Plätze habe ich durch Items mit absteigendem S-Wert je nach Schwierigkeit und Trennschärfe ersetzt. Bei der Ersetzung lag die höchste Prioriät auf der Selegierung von Aufgaben mit geringem P-Wert (schwierige Items).
Die dadurch enstehende Auswahl von Items ist im Anhang unter 7.3 *Itemselektion* (Wahl 2) zu finden.

Definiert man den zuvor entwickelten und analysierten Test als so genannte Testvorform, so handelt es sich nach abgeschlossener Analyse und Aufgabenselektion (ohne eine Aufgabenrevision) um die Testendform.

4.3 Validität

Die eingangs erläuterte Abhängigkeit der Validität (Gültigkeit) von den Gütekriterien des Tests, wurde nach Lienert und Raatz (1994) wie folgt dargestellt:

Objektivität	→	Reliabilität	→	Validität
Schwierigkeit	→	Trennschärfe	→	

Aus diesem Zusammenhang lässt sich ableiten, dass es vier Voraussetzungen gibt, um einen validen Test zu konstruieren. Das heißt damit ein Test valide sein kann, muss er reliabel sein. Aber nur weil der Test reliabel ist, ist er nicht gleich valide.

Nach Analyse und Auswertung der vorausgesetzten Komponenten, fehlt zu einer vollständigen Begutachtung des Test nur noch die Validität.
Die Untersuchung der Validität des Tests (Validierung) ist im Rahmen dieser Arbeit nicht erfolgt und auch nicht vorgesehen gewesen, da eine Validierung an einem Außenkriterium in dem zeitlich vorgegebenen Rahmen nicht dürchführbar gewesen wäre. Um dieses wichtige Thema aber nicht unangerissen zu lassen, wird dazu im Hinblick auf weitere Untersuchungen noch ein kleiner Ausblick gegeben (siehe Kap. 4.5).
Der entwickelte Test zur Erfassung von Computer Literacy weist also - wenn überhaupt - nur eine psychologische Validität bzw. Augenscheingültigkeit (face validity) auf. Diese wird vor allem an den Versuchspersonen festgemacht, die einen Zusammenhang zwischen dem Inhalt des Tests (Aufgaben) und dem zu untersuchenden Merkmal (Computer Literacy) herstellen können (Lienert & Raatz, 1994).

Auch wenn die eigentliche Validierung fehlt, sollen hier einige Faktoren aufgeführt werden, die die externe Gültigkeit unter Umständen verringert haben könnte:
Der anfänglich auf Papier entwickelte Fragebogen zur Erfassung von Computer Literacy wurde für die eigentliche Umfrage in einen webbasierten Fragebogen überführt. Das

heißt die Versuchspersonen mussten bereits in der Lage sein einen Computer zu benutzen, um an der Studie teilnehmen zu können. Auch wenn dies gewünscht war, um erfahrenere Versuchspersonen zu erreichen, so könnte dieses Vorgehen dennoch einen unerwünschten Einfluss auf die Validität haben. Denn nichtsdestotrotz wurde nur ein kleiner Ausschnitt von Personen damit erreicht.
Außerdem ist die Anzahl der Versuchspersonen und damit auch der analysierten Datensätze relativ klein. Auch wenn im Rahmen dieser Studie ca. 80 Versuchspersonen gewünscht waren, so ist im Fall der Versuchspersonen mehr auch immer besser. In Anbetracht der Anzahl von Versuchspersonen könnte die Gültigkeit der Ergebnisse durch diesen Faktor durchaus vermindert sein.
Des Weiteren wurde nicht die Ursache dafür gefunden, dass die untersuchten Rohwerte (Punkte) keiner Normalverteilung unterliegen. Die Ursache dafür kann sowohl bei der heterogenen Stichprobe als auch an der mangelhaften Konstruktion des Tests liegen. Beides kann dazu führen, dass die Ergebnisse unter anderen Testbedingungen anders ausfallen und damit die externe Validität negativ beeinflussen.

Eine weitere Einschränkung in der Anwendbarkeit des Test zur Erfassung von Computer Literacy ergibt sich durch Veränderung des Begriffs Computer Literacy und die schnelle Entwicklung der Informationsgesellschaft (Naumann, Richter & Groeben, 2001). Denn diese Entwicklung bedeutet auch, dass sich die verwendeten Symbole und Begriffe verändern oder sogar veralten können. Das bedeutet, dass der Test inklusive der Items nicht immer der gleiche bleibt, sondern kontinuierlich überarbeitet und angepasst werden muss. Andernfalls kann es dazu kommen, dass trotz tempärer hoher Validität, der Test keine Gültigkeit mehr besitzt.

4.4 Probleme

Abschließend sollen auch die Probleme, die im Zuge der Erstellung dieser Arbeit auftraten, geschildert und nicht unterschlagen werden. Vorab sei gesagt, dass die Probleme vor allem zeitverzögernder Natur waren.

Zum Einen erforderte, nach Abschluss der fertigen Papierversion des Fragebogens, die Umsetzung des webbasierten Fragebogens viel Zeit. Diese Umsetzung wurde im Zuge einer Diplomarbeit von Nico Zeißig (Zeißig, 2009) vorgenommen.
Allen voran stellte dabei die Erreichbarkeit des Online-Fragebogens ein großes Problem dar. Der Zugriff auf den Fragebogen war über weite Strecken nur sporadisch möglich. Dies behinderte sowohl die Umsetzung, als auch die Fehlersuche und die Durchführung des Pilottests.
Zum Anderen musste ein erster Versuch der Datenerhebung (Ausfüllen des Fragebo-

gens durch Versuchsteilnehmer) abgebrochen werden, da noch Änderungen am Online-Fragebogen durch Nico Zeißig vorgenommen werden mussten. Dabei änderte sich vor der erfolgreichen Datenerhebung auch die Adresse (URL) unter der der Fragebogen erreichbar war.

4.5 Ausblick

Ein Vorschlag für weiterführende Untersuchungen wäre zum Einen die Testendform (Test nach Itemselektion) der Computer Literacy Skala von erfahrenen Nutzer zu überprüfen. Dabei müsste dieser erneut mit Versuchspersonen getestet und auswertet werden. Nur so kann festgestellt werden, ob die Itemselektion auch den gewünschten Erfolg hat und eine normale Rohwertverteilung nach sich zieht.
Dabei sollte die Reihenfolge der Items so verändert werden, dass der Schwierigkeitsgrad kontinuierlich ansteigt (Lienert & Raatz, 1994). Dies kann bei dem neuen Testaufbau beachtet werden, da nun Werte für die Schwierigkeit (Schwierigkeitsindex) vorliegen.
In Bezug auf die Schwierigkeit wäre es auch erstrebenswert noch zusätzliche schwere Items in den Test mit aufzunehmen. Ein Teil der Items ist immer noch zu leicht, weswegen diese durch schwerere ersetzt werden könnten.

Die wohl wichtigste Weiterführung dieser Studie betrifft die Validierung des Tests, die zum Beispiel anhand von Aufgaben an einem Fahrkartenautomat durchgeführt werden könnte. Denn die entscheidene Frage, ob der Test auch das misst, was er vorgibt zu messen - nämlich Computer Literacy - , ist und kann ohne die Durchführung einer Validierung nicht beanwortet werden.
Mit anderen Worten ist ohne eine Validierung des Tests an einem externen Kriterium, dieser auch nicht gültig und damit hinfällig.

5 Zusammenfassung

Um computergestützte Systemene erfolgreich benutzen zu können, ist es notwendig computerbezogene Kompetenz in Form von Symbolen, Begriffen und interaktiven Mustern zu besitzen.
Ein Prädiktor für dieses computerbezogene Wissen, welches einen Aspekt von Computer Literacy beschreibt, bietet eine Skala zur Erfassung von Computer Literacy. Die bisher entwickelten Skalen gelten jedoch nur für Menschen mit einem geringem Maß an Computer Literacy, da die verwendeten Items zur Erfassung des Wissens zu einfach für erfahrene Nutzer waren.

Für den neu entwickelten Test wurden zunächst ausgewählt Items zu einem Fragebogen zur Erfassung von Computer Literacy (und Computererfahrung) thematisch zusammengefasst. Anschließend wurde dieser von möglichst erfahrenen Nutzern ausgefüllt. Die verwendeten Items dann wurden im Rahmen einer durchgeführten Analyse für die weitere Verwendung bewertet und abschließend eliminiert oder selegiert.

Aus dieser vorgenommenen Selektion von Items erschließt sich der neue Fragebogen zur Erfassung von Computer Literacy für erfahrene Nutzer.

6 Literaturverzeichnis

Akademischer Grad. (o. J.). (Online verfügbar unter `http://de.wikipedia.org/wiki/Akademischer_Grad`, zuletzt besucht am 22. Mai 2010.)

Arning, K. & Ziefle, M. (2008). Development and validation of a computer expertise questionnaire for older adults. *Behaviour & Information Technology*, *27* (1), 89–93.

Bortz, J. (1999). *Statistik für Sozialwissenschaftler.* Springer.

Carmines, E. G. & Zeller, R. A. (1979). *Reliability and validity assessment.* Sage Publications, Inc.

Cheng, T. T. & Stevens, D. J. (1985). Prioritizing Computer Literacy Topics. *Comput. Educ.*, *9* (1), 9–13.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16* (3), 297–334.

Gupta, G. K. (2006). Computer Literacy: Essential in Today's Computer-Centric World. *SIGCSE Bull.*, *38* (2), 115–119.

Hall, B. (2004). *E-learning: IT Competencies, computer literacy and student attitudes to e-learning.*

Idowu, P. A., Adagunodo, E. R. & Popoola, B. I. (2003). Computer Literacy Level and Gender Differences among Nigerian University Staff. *An On-line African Educational Research Journal*, *3* (3).

Lienert, G. A. & Raatz, U. (1994). *Testaufbau und Testanalyse.* Beltz, Psychologie Verlags-Union.

Lockheed, M. E., Hunter, B., Anderson, R. E., Beazly, R. M. & Esty, E. T. (1983). *Computer Literacy: Definition and Survey Items for Assessment in Schools.*

Mac OS X: Tastenkombinationen. (2007, Juni). (Online verfügbar unter `http://docs.info.apple.com/article.html?artnum=75459-de`, zuletzt besucht am 20. Februar 2010.)

Major, J. H. & Prueher, R. F., Jr. (1984). Functional Computer Literacy. In *Siguccs '84: Proceedings of the 12th annual acm siguccs conference on user services* (S. 135–138). New York, NY, USA: ACM.

Mason, J. & Morrow, R. M. (2006). YACLD: yet another computer literacy definition. *J. Comput. Small Coll.*, *21* (5), 94–100.

McMillan, S. (1996). Literacy and Computer Literacy: definitions and comparisons. *Comput. Educ.*, *27* (3-4), 161–170.

Naumann, J., Richter, T. & Groeben, N. (1999). *Inventar zur Computerbildung.*

Naumann, J., Richter, T. & Groeben, N. (2001). Das Inventar zur Computerbildung INCOBI: Ein Instrument zur Erfassung von Computer Literacy und computer-

vezogenen Einstellungen bei Studierenden der Geistes- und Sozialwissenschaften. *Psychologie in Erziehung und Unterricht*, *48*, 1–13.

Naumann, J., Richter, T. & Groeben, N. (2002). Validierung des INCOBI anhand eines Vergleichs von Anwendungsexperten und Anwendungsnovizen. *Zeitschrift für Pädagogische Psychologie*, *15*, 219–232.

Poynton, T. A. (2005). Computer literacy across the lifespan: a review with implications for educators. *Computers in Human Behavior*, *21*, 861–872.

Pyrczak, F. (1990). Development of Diagnostic Tests for Computer Literacy. *Comput. Educ.*, *14* (3), 213–216.

Sawitzki, G. (2008). *Statistical Computing: Einführung in R.*

Sengpiel, M. (2007). *Computer literacy scale (preview).*

Sengpiel, M. & Dittberner, D. (2008). *The computer literacy scale (CLS) for older adults - development and validation.*

Turner, G. M., Sweany, N. W. & Husman, J. (2000). Development of the Computer Interface Literacy Measure. *Journal of Educational Computing Research*, *22* (1), 37–54.

Über 100 Social networks aus Deutschland. (2008, Mai). (Online verfügbar unter `http://netzwertig.com/2007/10/16/zn-uber-100-social-networks-aus-deutschland/`, zuletzt besucht am 22. Mai 2010.)

Van Dyke, C. (1987, May). Taking "computer literacy"literally. *Communications of the ACM*, *30*, 366–374.

Vliet, P. J. A. van, Kletke, M. G. & Chakraborty, G. (1994). The measurement of computer literacy: a comparison of self-appraisal and objective tests. *Int. J. Hum.-Comput. Stud.*, *40* (5), 835–857.

Zeißig, N. (2009). *Entwurf und Umsetzung einer webbasierten Diagnoseplattform zur Erhebung von deklarativem und prozeduralem Interaktionswissen.* Unveröffentlichte Diplomarbeit, Humboldt Universität zu Berlin.

7 Anhang

7.1 Fragebogen

Fragebogen
zur Bachelorarbeit

Freie Universität Berlin

Im Folgenden interessiert uns, welche Erfahrungen Sie mit Computern haben. Im ersten Teil dieses Fragebogens bitten wir Sie um einige Angaben zu Ihrer Nutzung von Computern. Im zweiten Teil besteht Ihre Aufgabe darin, unter anderem Begriffe verschiedenen Symbolen und Bedeutungen zuzuordnen. Die Bearbeitung wird etwa 15 Minuten dauern. Bitte lesen Sie die Erklärung zur Bearbeitung des Fragebogens genau durch. Es ist wichtig, dass Sie alle Fragen beantworten. Vielen Dank für Ihre Unterstützung!

Teil A: Erfahrung mit Computern

1.) Seit wie vielen Jahren benutzen Sie einen Computer? _____

2.) Wie viele Stunden pro Woche benutzen Sie normalerweise einen Computer? _____

3.)

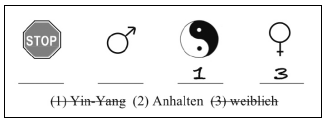
Wie oft benutzen Sie einen Computer für folgende Tätigkeiten?	**nie**	**selten**	**gelegent -lich**	**oft**
Textverarbeitung	❑	❑	❑	❑
Tabellenkalkulation	❑	❑	❑	❑
Präsentationserstellung	❑	❑	❑	❑
Bildbearbeitung	❑	❑	❑	❑
Programmieren	❑	❑	❑	❑
Computerspielen	❑	❑	❑	❑
E-Mail	❑	❑	❑	❑
Chat	❑	❑	❑	❑
VoIP	❑	❑	❑	❑
Internet-Surfen	❑	❑	❑	❑
Gezielte Informationssuche	❑	❑	❑	❑
Online-Einkäufe	❑	❑	❑	❑
Online-Banking	❑	❑	❑	❑

CLS+ Seite 1 von 4 Beitrag zum Projekt ALISA

Teil B: Zuordnung von Symbolen und Begriffen

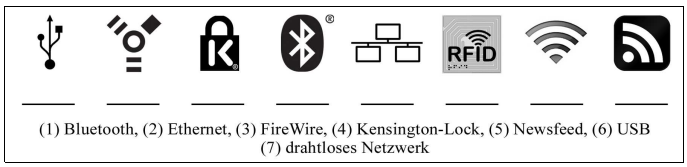
Auf dieser Seite sehen Sie verschiedene Symbole, die für die Bedienung von technischen Geräten oder Computern relevant sind. Bitte ordnen Sie diese ihren jeweiligen Bedeutungen zu. Zur Veranschaulichung dient dieses Beispiel mit bekannten Alltagssymbolen:

Die Bedeutungen verfügen jeweils über eine Nummer. Diese tragen Sie bitte unter das passende Symbol ein. Bitte beachten Sie dabei, dass es nicht zu jedem Symbol eine passende Bedeutung gibt, sondern immer ein Symbol übrig bleibt.

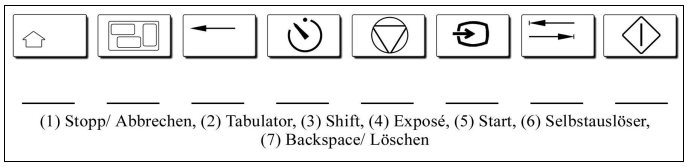


Hinweis: Die Zuordnung wird einfacher, wenn Sie die bereits zugeordneten Bedeutungen – wie im Beispiel dargestellt – einfach durchstreichen.

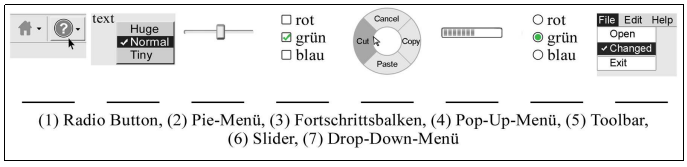
Es ist nur natürlich, dass Sie nicht alle Antworten kennen. Dennoch bitten wir Sie, allen Bedeutungen möglichst zügig ein Symbol zuzuordnen. Wenn Sie nicht sicher sind, raten Sie einfach, welche Bedeutung am Besten passen könnte.
Viel Erfolg!



(1) Bluetooth, (2) Ethernet, (3) FireWire, (4) Kensington-Lock, (5) Newsfeed, (6) USB
(7) drahtloses Netzwerk



(1) Stopp/ Abbrechen, (2) Tabulator, (3) Shift, (4) Exposé, (5) Start, (6) Selbstauslöser,
(7) Backspace/ Löschen



(1) Radio Button, (2) Pie-Menü, (3) Fortschrittsbalken, (4) Pop-Up-Menü, (5) Toolbar,
(6) Slider, (7) Drop-Down-Menü

Bitte ordnen Sie nun in gleicher Weise die folgenden Tastenkombinationen ihren jeweiligen Bedeutungen zu.

Drücken der Befehlstaste:

+ A	+ C	+ O	+ P	+ S	+ V	+ X	+ Z
_____	_____	_____	_____	_____	_____	_____	_____

(1) Ausschneiden, (2) Drucken, (3) Einfügen, (4) Kopieren, (5) Öffnen, (6) Sichern/ Speichern (7) Widerrufen/ Rückgängig

Im Folgenden ordnen Sie bitte die Beschreibungen den entsprechenden Dateiendungen zu.

.app	.csv	.iso	.rar	.sh	.swf	.tiff	.ttf
_____	_____	_____	_____	_____	_____	_____	_____

(1) Shell-Skript
(2) komprimiertes Dateiarchiv
(3) ausführbare Datei
(4) Flash-Animation
(5) CD-Image
(6) Vektorgrafisches Format für Schriftschnitte
(7) Textdatei, die durch Komma getrennte Werte beinhaltet

Diesmal sind folgende Begriffe ihren jeweiligen Erklärungen zuzuordnen.

Thumbnail	Tooltip	Cache	ASCII	URL	Shell
_____	_____	_____	_____	_____	_____

(1) Betriebssysteme haben grafische Benutzeroberflächen und Kommandozeileninterpreter, wobei mit diesem Begriff meist letztere gemeint sind.
(2) Eine verkleinerte Grafik zur Vorschau von Bildern.
(3) Identifiziert und lokalisiert eine Ressource über das verwendete Netzwerkprotokoll und den Ort der Ressource im Computernetzwerk.
(4) Schneller Speicher mit geringer Kapazität, der als Puffer eingesetzt wird.
(5) Kurzer Text, der eingeblendet wird, wenn der Mauszeiger längere Zeit auf einem Icon verweilt.

Stream	Desktop	Taskleiste	HTTP-Cookie	IP-Adresse	Captcha
_____	_____	_____	_____	_____	_____

(1) Automatischer Test um Menschen von Computern zu unterscheiden.
(2) Information, die ein Webserver an einen Browser schickt und clientseitig gespeichert ist.
(3) Teil einer grafischen Oberfläche, der z.B. ein zentrales Programmauswahlmenü bereitstellt, derzeit laufende Programme und aktuelles Datum und Uhrzeit anzeigt.
(4) Nummmer zur Identifizierung eines Computers oder Gerätes in einem Netzwerk.
(5) Arbeitsoberfläche und unterste Fensterebene eines grafischen Betriebssystems.

Formatierung	Konfiguration	Validierung	Defragmentierung	Indexierung	Partitionierung
_____	_____	_____	_____	_____	_____

(1) Beschreibt den Vorgang und das Ergebnis einer Neuordnung von Datenblöcken auf einem Speichermedium durch ein spezielles Programm.
(2) Verfahren der Kennzeichnung, bei dem einem Dokument bestimmte vorher festgelegte beschreibende Elemente zugeordnet werden.
(3) Aufteilung einer Festplatte in mehrere logische Bereiche.
(4) Im Bereich der Datenspeicherung ist dies ein Prozess durch den ein Speichermedium zur Aufnahme von Daten vorbereitet wird.
(5) Eine bestimmte Einstellung von Programmen oder Hardwarebestandteilen.

Vielen Dank für das Ausfüllen des Fragebogens!

Bitte notieren Sie abschließend Ihr Alter, Geschlecht und sofern zutreffend den Studiengang oder Studienabschluss. Danke!

Alter

_____

Geschlecht

❑ männlich ❑ weiblich

Studiengang/-abschluss

_____

7.2 Items

Item (Bezeichnung)	Quelle / Rechtliches	Name / Begründung
Symbole		
(s41)	• http://de.wikipedia.org/w/index.php?title=Datei:Usb-svg.svg • unter Creative Commons-Lizenz (Attribution ShareAlike 2.5) Autor: DocMiller	• USB • weit verbreiteter Standard • Einsatzgebiete: Handys, Computer, Drucker usw.
(s42)	• Apple Inc. • http://developer.apple.com/softwarelicensing/agreements/firewire.html	• FireWire • IEEE 1394 Standard • Einsatzgebiete: PlayStation 2, Computer, Autos
(s43)	• Kensington Computer Products Group • http://uk.kensingtonimages.com • Nutzungsrecht per E-Mail erhalten von Magdalena Kijek European PR and Campaign Executive	• Kensington-Lock • Diebstahlsicherung • Einsatzgebiete: portable Geräte wie Beamer, Notebooks usw.
(s44)	• Bluetooth SIG, Inc. • Nutzungsrechts und Symbol per E-Mail erhalten von Nathan Funkhouser Membership Sales Manager	• Bluetooth • Standard zur kabellosen Datenübertragung • Einsatzgebiete: Handys, Computer usw.
(s45)	• erfunden von David W. Hill (unter IBM) • kein standardisiertes Symbol • Grafik selbst nachgestellt	• Ethernet • gebräuchlichster LAN Standard (IEEE 802.3) • Einsatzgebiete: Computer und andere Netzwerkgeräte
(s46)	• Andreas Weigand - Gewinner des Logo-Wettbewerbs vom Informationsforum RFID • (noch) kein standardisiertes Symbol	• RFID • Barcode-Nachfolger; Speicherung von Produktdaten • Einsatzgebiete: hauptsächlich i.d. Industrie, Bibliotheken, WM2006
(s47)	• kein standardisiertes Symbol • Grafk selbst nachgestellt	• Wireless • Datenübertragung via Funk • Einsatzgebiete: Handys, Computer

Item (Bezeichnung)	Quelle / Rechtliches	Name / Begründung
(s48)	• Mozilla Foundation • http://www.feedicons.com	• News-Feed • Markierung von Feeds im RSS- oder Atom-Format • Einsatzgebiete: Webseiten
Tasten		
(s51)	• Grafik selbst erstellt	• Shift • Einsatzgebiete: jede Tastatur
(s52)	• Tastatur: Mac-Produktreihe • Grafik selbst erstellt	• Exposé • Einsatzgebiete: Apple-Tastatur
(s53)	• Grafik selbst erstellt	• Backspace/ Löschen • Einsatzgebiete: jede Tastatur
(s54)	• Grafik selbst erstellt	• Selbstauslöser • Einsatzgebiete: Foto-Kameras
(s55)	• Scanner: HP Scanjet 5590 • Kopierer: Canon PC-D450 • Multifunktions- und Faxgeräte: Epson Photo RX700, Canon Pixma MP950 • Grafik selbst erstellt	• Stopp/ Abbrechen • Einsatzgebiete: Scanner, Kopierer, Multifunktions- und Faxgeräte
(s56)	• Displays: Dell 2707WFP • Fernbedienungen: Marke unbekannt (http://www.faqs.org/photo-dict/photofiles/list/824/1251remote_control.jpg) • Grafik selbst nachgestellt	• Input • Einsatzgebiete: Displays, Fernbedienungen
(s57)	• Grafik selbst erstellt	• Tabulator • Einsatzgebiete: jede Tastatur

Item (Bezeichnung)	Quelle / Rechtliches	Name / Begründung
(s58)	• Waschmaschinen: Clatronic • Drucker: Canon i-Sensys Fax L-3000 • Kopierer: Kyocera FS-Produktreihe • Multifunktions- und Faxgeräte: Epson Stylus DX9400F, Canon Pixma MP950 • Grafik selbst nachgestellt	• Start • Einsatzgebiete: Waschmaschinen, Drucker, Kopierer, Multifunktions- und Faxgeräte
graphische Elemente		
☐ rot ☑ grün ☐ blau (s61)	• Grafik selbst nachgestellt	• Checkbox • Einsatzgebiete: Formulare in Programmen, Betriebssystemen etc.
File Edit Help Open ✓ Changed Exit (s62)	• Grafik selbst nachgestellt	• Drop-Down-Menü • Einsatzgebiete: Programme jeglicher Art, Betriebssysteme
Cancel Cut Copy Paste (s63)	• Grafik selbst nachgestellt	• Pie-Menü • Einsatzgebiete: Spielen (Die Sims, Second Life) Mozilla Firefox
text Huge ✓ Normal Tiny (s64)	• Grafik selbst nachgestellt	• Pop-Up-Menü • Einsatzgebiete: Programme, Betriebssysteme
(s65)	• Grafik selbst nachgestellt	• Fortschrittsbalken • Einsatzgebiete: Programme, Betriebssysteme
○ rot ◉ grün ○ blau (s66)	• Grafik selbst nachgestellt	• Radio Button • Einsatzgebiete: Formulare, Programme, Betriebssysteme
(s67)	• Grafik selbst nachgestellt	• Slider • Einsatzgebiete: Lautstärkeregelung (Betriebsystem, Youtube)

Item (Bezeichnung)	Quelle / Rechtliches	Name / Begründung
(s68)	• Grafik selbst nachgestellt	• Toolbar • Einsatzgebiete: Betriebssysteme, installierbare Browser-Toolbars, Office-Produkte
Tastenkombination		

Befehl + A (t11)	• Grafik selbst nachgestellt
	• Quellen für die Tastenkombinationen:
	  • http://docs.info.apple.com/article.html?artnum=75459-de
	  • http://support.microsoft.com/kb/301583/de
	  • http://windows.microsoft.com/de-DE/windows-vista/Keyboard-shortcuts
		• Alles Markieren
		• Einsatzgebiete (gilt für alle Tastenkombinationen): alle Betriebssysteme, hardware Tastatur

Befehl + C (t12)	• Grafik selbst nachgestellt	• Kopieren
Befehl + O (t13)	• Grafik selbst nachgestellt	• Öffnen
Befehl + P (t14)	• Grafik selbst nachgestellt	• Drucken
Befehl + S (t15)	• Grafik selbst nachgestellt	• Sichern/ Speichern
Befehl + V (t16)	• Grafik selbst nachgestellt	• Einfügen
Befehl + X (t17)	• Grafik selbst nachgestellt	• Ausschneiden
Befehl + Z (t18)	• Grafik selbst nachgestellt	• Widerrufen/ Rückgängig
Dateiendungen		
.app (d1)		• ausführbare Datei • Einsatzgebiete: Apple-Produkte
.csv (d2)		• Datei mit durch Komma getrennten Werten • Einsatzgebiete: Office-Produkte, Datenbanken

Item (Bezeichnung)	Quelle / Rechtliches	Name / Begründung
.iso (d3)		• CD-Image • Einsatzgebiete: CD, DVD
.rar (d4)		• Dateiarchiv • unbekannter als zip
.sh (d5)		• Shell-Skript • Einsatzgebiete: Programmieren
.swf (d6)		• Flash-Animation • Einsatzgebiete: Web
.tiff (d7)		• Bild • Einsatzgebiete: Bildbearbeitung
.ttf (d8)		• Schriftschnitte
Begriffe I		
Thumbnail (b21)		• verkleinerte Grafik (Vorschau) • Einsatzgebiete:
Tooltip (b22)		• kurzer Hilfe-Text • Einsatzgebiete:
Cache (b23)		• schneller Speicher • Einsatzgebiete:
ASCII (b24)		• kurzer Hilfe-Text • Einsatzgebiete:
URL (b25)		• Lokalisierung einer Ressource • Einsatzgebiete:
Shell (b26)		• ”Kommandozeile” • Einsatzgebiete:

Item (Bezeichnung)	Quelle / Rechtliches	Name / Begründung
Begriffe II		
Stream (b31)		• kontinuierliche Abfolge von Daten • Einsatzgebiete: Internet
Desktop (b32)		• Arbeitsoberfläche • Einsatzgebiete: Betriebssystem
Taskleiste (b33)		• "zentrale Programmauswahl" • Einsatzgebiete: Betriebssysteme
Http-Cookie (b34)		• "Webseiten-Information" • Einsatzgebiete: Webseiten
IP-Adresse (b35)		• Identifikationsnummer im Netzwerk • Einsatzgebiete: Internet
Captcha (b36)		• Test zur Unterscheidung von Menschen und Maschinen • Einsatzgebiete: E-Mail, Online-Umfragen, Blogs, sonstige Webseiten
Begriffe III		
Formatierung (b41)		• Vorbereitung zur Daten-Aufnahme • Einsatzgebiete: Festplatten, USB-Sticks, MP3-Player
Konfiguration (b42)		• bestimmte Einstellung • Einsatzgebiete: Programme
Validierung (b43)		• Gültigkeitsbeweis • Einsatzgebiete: W3C, Programmierung, IT-Sicherheit

Item (Bezeichnung)	Quelle / Rechtliches	Name / Begründung
Defragementierung (b44)		• ”Daten-Umordnung” • Einsatzgebiete: Festplatten, sonstige Speichermedien
Indexierung (b45)		• Zuordnung von Deskriptoren • Einsatzgebiete: Such-Maschinen - online oder im Betriebssystem
Partitionierung (b46)		• Festplatteneinteilung • Einsatzgebiete: Festplatten, sonstige Speichermedien

7.3 Itemselektion

Itemselektion

Wahl 1	Wahl 2	i	Item	Bezeichnung	Trennschärfe	Schwierigkeit	Selektions-kennwert
☑	☑	22	Radio Button	s66	0,53	92,86	1,03
☑	☐	51	IP-Adresse	b35	0,23	98,57	0,97
☑	☐	1	USB	s41	0,48	91,43	0,86
☑	☐	26	Befehl + C	t12	0,40	94,29	0,86
☑	☐	29	Befehl + S	t15	0,40	94,29	0,86
☑	☑	43	Cache	b23	0,60	84,29	0,82
☑	☑	42	Tooltip	b22	0,56	82,86	0,74
☑	☑	17	Checkbox	s61	0,64	74,29	0,73
☑	☐	58	Partitionierung	b46	0,48	87,14	0,72
☑	☑	41	Thumbnail	b21	0,61	75,71	0,71
☑	☐	11	Backspace	s53	0,36	92,86	0,70
☑	☑	19	Pie-Menü	s63	0,56	80,00	0,70
☑	☐	50	Http-Cookie	b34	0,47	87,14	0,70
☑	☑	24	Toolbar	s68	0,61	72,86	0,69
☑	☑	46	Shell	b26	0,68	52,86	0,68
☑	☑	4	Bluetooth	s44	0,51	81,43	0,66
☑	☑	35	.iso	d3	0,64	62,86	0,66
☑	☑	5	Ethernet	s45	0,63	62,86	0,65
☑	☑	54	Konfiguration	b42	0,55	77,14	0,65
☑	☑	37	.sh	d5	0,54	77,14	0,64
☑	☑	52	Captcha	b36	0,50	81,43	0,64
☑	☑	18	Drop-Down-Menü	s62	0,55	74,29	0,63
☑	☑	38	.swf	d6	0,63	48,57	0,63
☑	☑	56	Defragementierung	b44	0,50	80,00	0,62
☑	☐	9	Shift	s51	0,39	87,14	0,58
☑	☐	30	Befehl + V	t16	0,42	84,29	0,58
☑	☐	36	.rar	d4	0,48	78,57	0,58
☑	☑	33	.app	d1	0,54	62,86	0,56
☑	☑	2	FirWire	s42	0,52	64,29	0,54
☑	☑	8	News-Feed	s48	0,50	68,57	0,54
☐	☐	28	Befehl + P	t14	0,38	85,71	0,54
☐	☑	23	Slider	s67	0,52	60,00	0,53
☐	☐	57	Indexierung	b45	0,48	71,43	0,53
☐	☐	15	Tabulator	s57	0,29	90,00	0,48
☐	☐	27	Befehl + O	t13	0,24	92,86	0,47

Wahl 1	Wahl 2	i	Item	Bezeichnung	Trennschärfe	Schwierigkeit	Selektions-kennwert
☐	☐	7	Wireless	s47	0,38	77,14	0,45
☐	☑	34	.csv	d2	0,43	35,71	0,45
☐	☐	45	URL	b25	0,39	75,71	0,45
☐	☐	32	Befehl + Z	t18	0,36	78,57	0,44
☐	☑	49	Taskleiste	b33	0,41	68,57	0,44
☐	☐	47	Stream	b31	0,30	85,71	0,43
☐	☑	6	RFID	s46	0,39	67,14	0,42
☐	☐	48	Desktop	b32	0,36	71,43	0,40
☐	☑	53	Formatierung	b41	0,38	64,29	0,40
☐	☑	31	Befehl + X	t17	0,37	64,29	0,39
☐	☑	40	.ttf	d8	0,33	22,86	0,39
☐	☐	12	Selbstauslöser	s54	0,29	82,86	0,38
☐	☑	10	Exposé	s52	0,35	55,71	0,35
☐	☑	44	ASCII	b24	0,31	61,43	0,32
☐	☑	13	Stopp	s55	0,24	28,57	0,27
☐	☐	20	Pop-Up-Menü	s64	0,20	84,29	0,27
☐	☐	21	Fortschrittsbalken	s65	0,26	65,71	0,27
☐	☐	55	Validierung	b43	0,22	64,29	0,23
☐	☐	3	Kensington	s43	0,10	94,29	0,22
☐	☐	25	Befehl + A	t11	0,20	67,14	0,21
☐	☐	16	Start	s58	0,17	37,14	0,18
☐	☐	14	Input	s56	0,16	40,00	0,16
☐	☐	39	.tiff	d7	-0,02	34,29	-0,02

Anmerkung
Wahl 1 - durch Selektionskennwert S
Wahl 2 - durch S-Wert und rationale Selektion

7.4 Anwerberschreiben

Betreff:
Studie über interaktives Computerwissen

Anschreiben:
Lieber Student, liebe Studentin,

hast Du dich schon einmal gefragt, wieviel Wissen Du eigentlich schon im Umgang mit Computern erworben hast?

Nein?!
Dabei benutzt Du große Teile dieses Wissens unterbewusst beim Surfen im Internet, beim täglichen Gebrauch Deines Handys und anderen elektronischen Geräten.

Dieses interaktive Wissen soll in dieser Studie abgefragt werden. Neugierig?

http://141.20.68.27/ALISA/clsplus.seam

Die Bearbeitungszeit des Fragebogens beträgt ca. 15 min.
Die Angaben werden selbstverständlich anonym und ausschließlich für wissenschaftliche Zwecke verwendet.

Im Voraus vielen Dank für die Unterstützung dieser Studie!

ACHTUNG: Der Fragebogen funktioniert derzeit nur mit Mozilla Firefox!

7.5 verwendete Plattformen

Studenten-Foren

	Forum	URL	Alexa.com (DE)	Seitwert.de - Ranking
☑	studiVZ	www.studivz.net	14	21
☐	Jappy	www.jappy.de	57	336
☑	gulli	www.gulli.com	79	892
☑	Unicum Online	www.unicum.de	1.920	6.291
☑	Studentum	www.studentum.de	9.409	5.742
☑	Study-Board	www.study-board.de	10.515	8.268
☑	nur!studenten	www.nurstudenten.de	24.533	28.358
☑	Studenten Forum	http://forum.student.de	14.789	

studiVZ-Gruppen

	Gruppe	Mitglieder
☑	Alles kostenlos - wo man als kluger Student sparen kann!	26.131
☑	Ich bin Student, ich hab' keine Zeit!	3.126
☑	Student sein ist geil, nur studieren ist scheisse...	969
☑	Ich bin Student - es liegt in meiner Natur...	830
☑	Ich bin wirklich Student und brauch mich hier nicht reinmogeln!!	691
☑	Vorlesung um 8? Ich bin Student und kein Bäcker...	369
☑	79er-Jahrgang und immer noch Student	159
☑	Hart, Härter, STUDENT	114
☑	Mein Name ist Hase. Ich bin Student.	79
☑	Es ist toll Student zu sein	110
☑	Ich bin Student und kann mir trotzdem die Playstation3 leisten	52

7.6 Grafiken

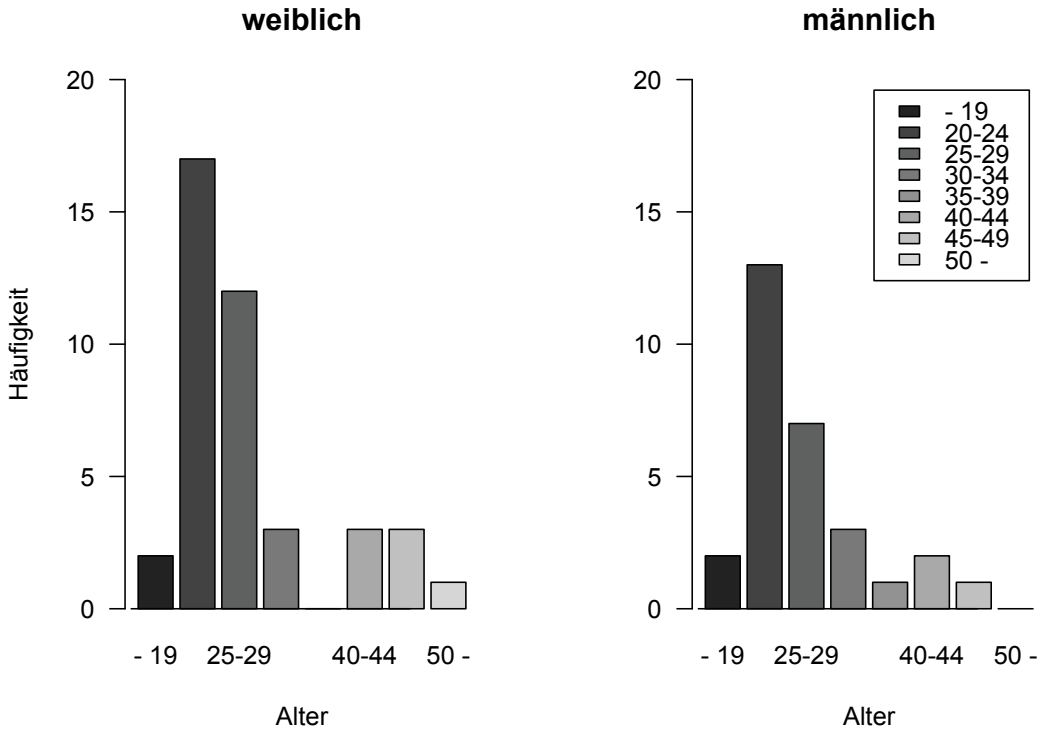


Abbildung 4: Altersverteilung der Versuchspersonen nach Geschlecht

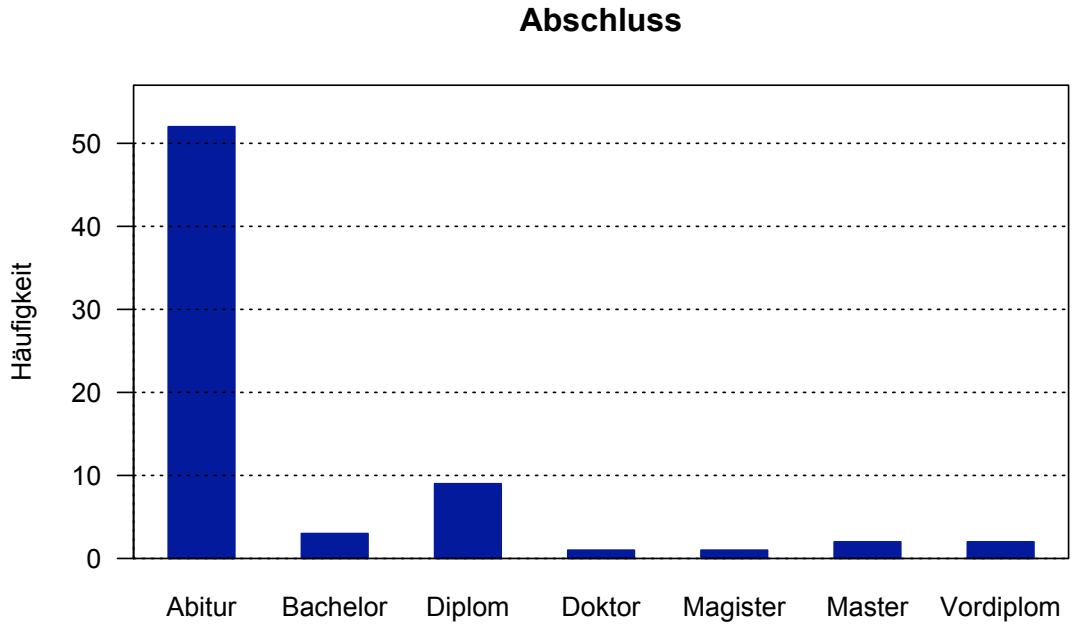


Abbildung 5: Abschlussverteilung der Versuchspersonen

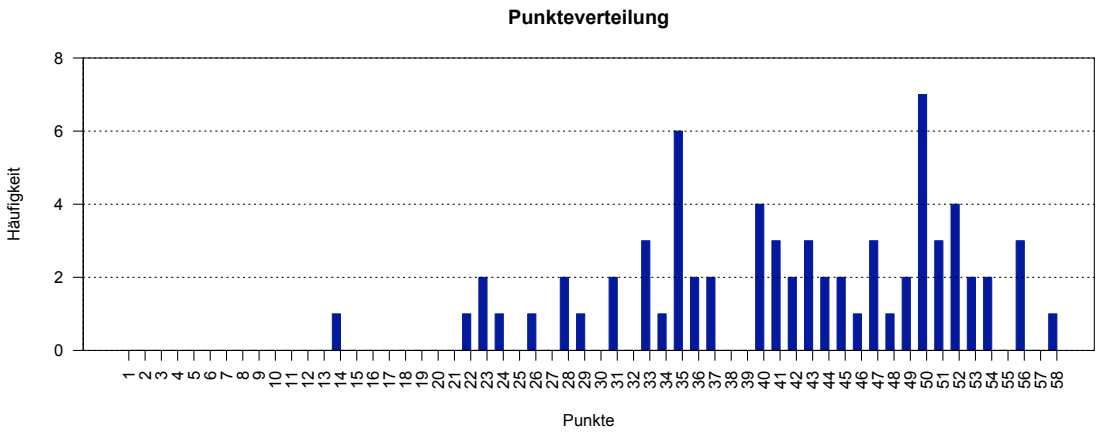


Abbildung 6: Punkteverteilung der Versuchspersonen

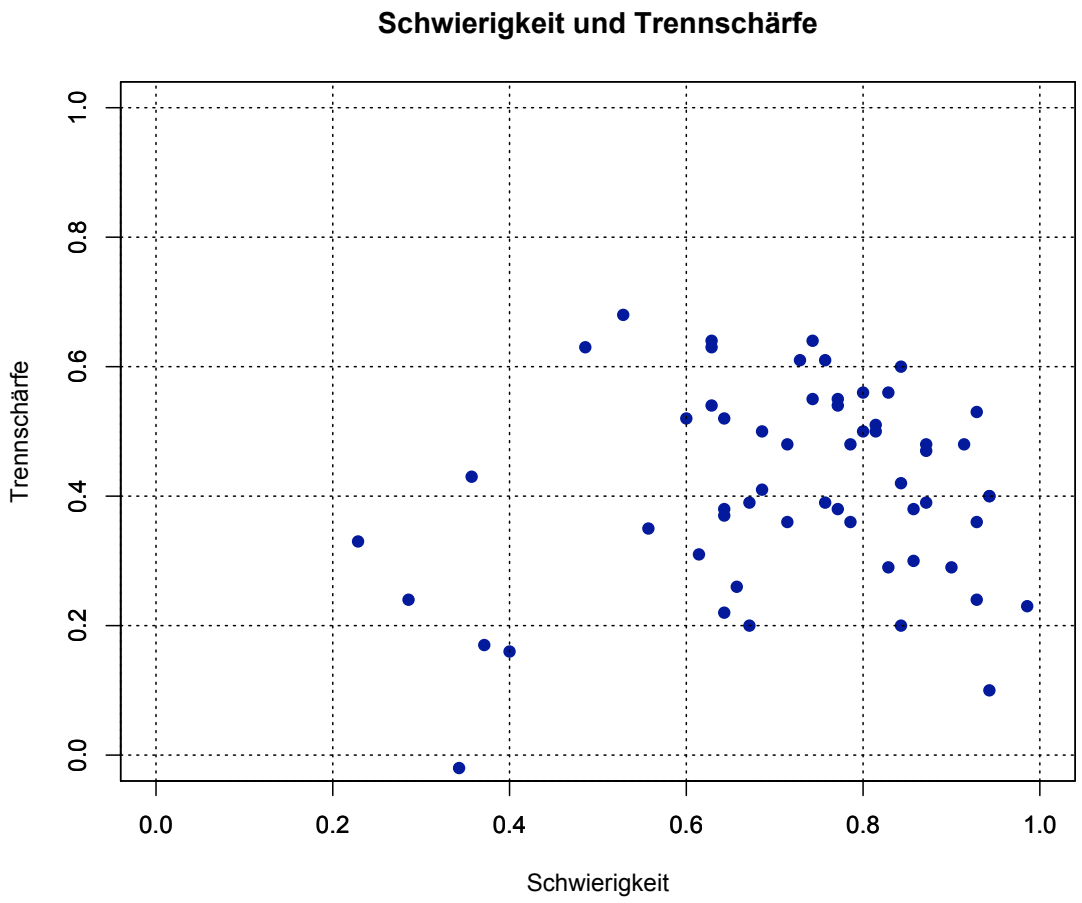


Abbildung 7: Beziehung zwischen Trennschärfe und Schwierigkeit (parabolische Abhängigkeit)

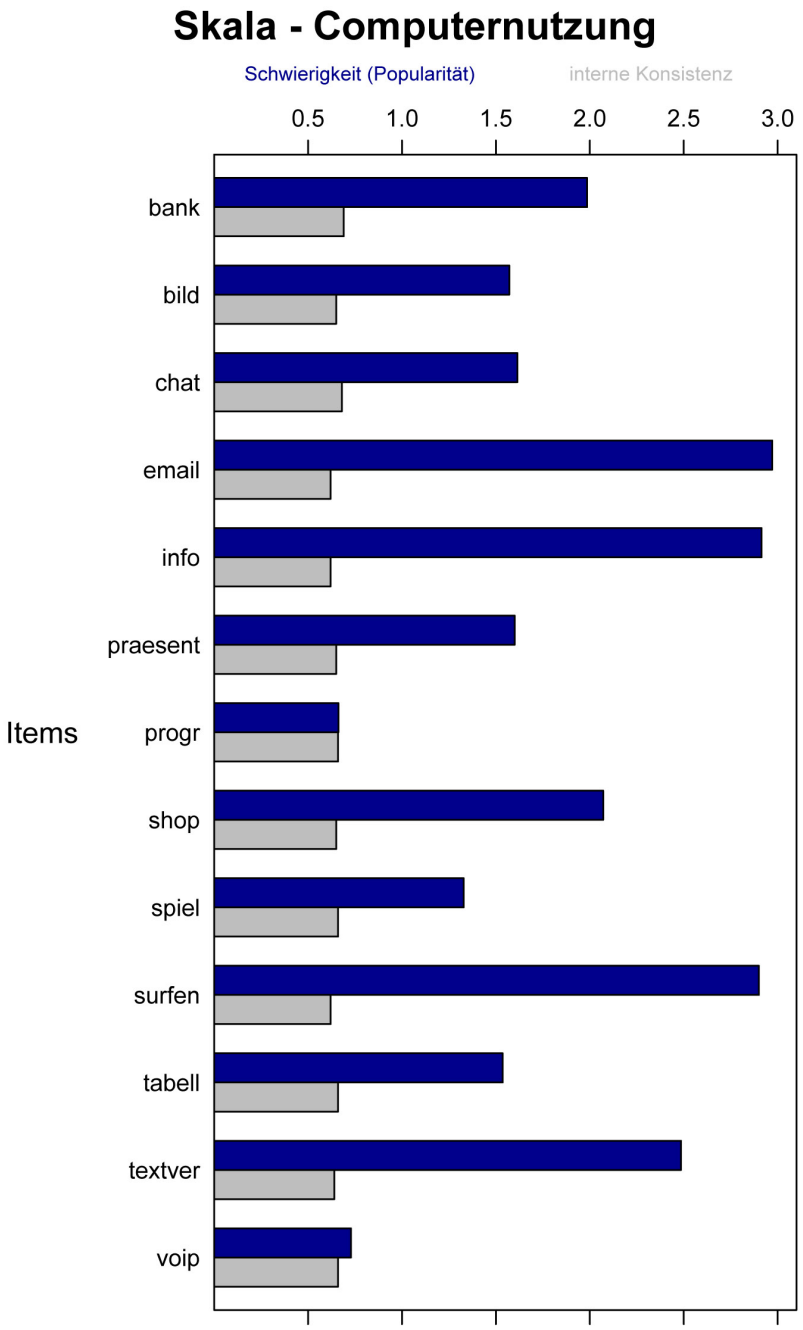


Abbildung 8: Computererfahrung - Popularität und interne Konsistenz je Item

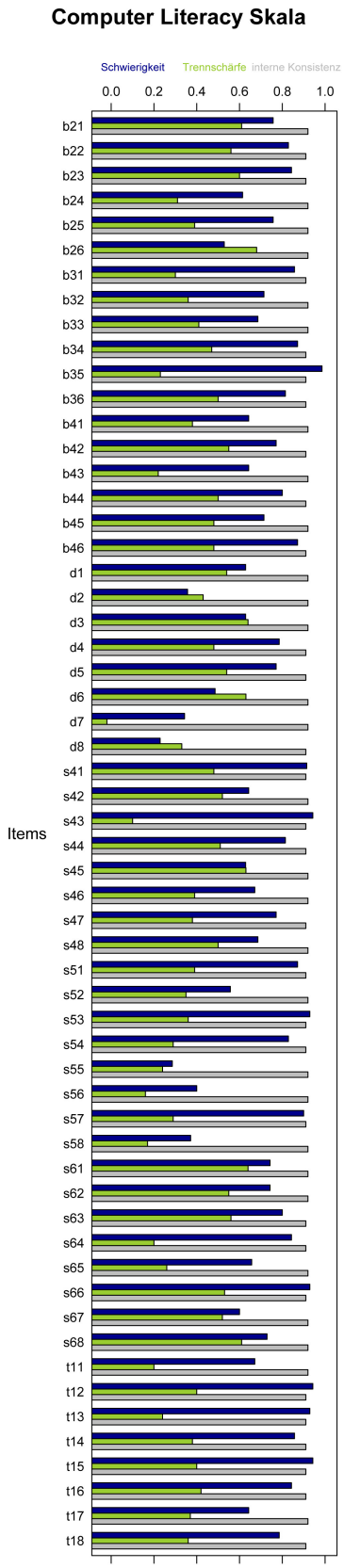


Abbildung 9: CLS - Zusammenfassung aller Itemkennwerte je Item

7.7 Rohdaten

UserID	Dauer	s41	s42	Alter	Geschlecht	Abschluss
56	11	6	3	22	0	Abitur
59	10	6	3	99	99	Professur in Informatik
60	12	6	3	20	1	Abitur
62	14	6	3	22	1	Abitur
63	15	6	2	22	1	Abitur
64	6	6	3	22	0	Abitur
65	20	6	3	49	0	Promotion
66	15	6	2	49	1	Diplom
69	15	6	3	23	0	Abitur
70	14	6	3	24	0	Hochschulreife

Tabelle 1: Auszug aus der csv-Datei

Tabelle 1 zeigt die Rohdaten von ausgewählten Spalten der ersten 10 Datensätze.

Die vollständigen Rohdaten sind an dieser Stelle aufgrund des Umfangs nicht dargestellt. Für einen vollständigen Überblick sei auf die beiliegende CD verwiesen.

7.8 R-Code

Der vollständige R-Code ist an dieser Stelle aufgrund seines Umfangs nicht dargestellt. Für einen vollständigen Überblick sei auf die beiliegende CD verwiesen.

Nichtsdestotrotz sollen die programmierten Umsetzungen der zuvor vorgestellten Formeln hier ihren Platz finden.

Nützliche Informationen:
Die in dem Fragebogen gemachten Angaben der Teilnehmer befinden sich in einem so genannten `data.frame` namens `data`.
Alle berechneten Itemkennwerte (Reliabilität, Trennschärfe usw.) der Analyse befinden sich - sofern nicht anders angegeben - im `data.frame analysis`.

Formel 1: Rohwert-/CLS-Punkte-Berechnung
Funktion: `points2data`
Zur Berechnung der CLS-Punkte müssen zunächst die gegebenen Antworten der Teilnehmer mit der Lösung verglichen werden (Funktion `comp.w.solution`). Aus diesem Abgleich entsteht ein boolescher `data.frame`, der mit booleschen Einträgen gefüllt ist, wobei eine *True* für eine richtige Antwort und eine *False* für eine falsche Antwort steht. Dieser entstandene `data.frame answers.bool` wird anschließend der Funktion `points2data` als `answers` übergeben. In dieser Funktion wird für jeden Teilnehmer die Summe der CLS-Punkte (`points`) berechnet, wobei in R *True = 1* und *False = 0* entspricht und somit genau wie in der Formel 1 die Punkte nur durch die Anzahl der richtig gegebenen Antworten bestimmt wird.
Zusätzlich werden in der Funktion `points2data` die errechneten Punkte noch in den Haupt-data.frame `data` überführt.

Listing 1: in Funktion „points2data“

```
points = 0
for(j in 1:n.items) {
        if(answers[row,j] != NO_ANSWER) {points = points + answers[row,j]}
}
```

Formel 2 und 3: Test der Punkte auf Normalverteilung
Funktion: `calc.point.distribution`
Die Ergebnisse der Untersuchung auf eine Normalverteilung sind im `data.frame analysis.points` zu finden.
Beim Vergleich der erhaltenen Testpunktverteilung mit einer Normalverteilung, wer-

den bei einer geringen Anzahl von Versuchspersonen (N < 200) die Punkte zunächst zu Gruppen zusammengefasst. Das heißt, dass bei einer Gruppengröße von *h=3*, die Punkte 1,2,3 und 4,5,6 jeweils in einer Gruppe sind. Dabei hängt die Gruppengröße von der Anzahl der insgesamt möglichen Punkte ab (hier: 58 Punkte), denn am Ende soll nicht eine Gruppe übrig bleiben, bei der zu wenig Punkte zusammengefasst werden, also zum Beispiel nur die Punkte 57 und 58. Die Gruppeneinteilung soll also genau aufgehen.
In der Tabelle `tab.81` werden anschließend die Werte nach Lienert und Raatz (1994) zur Bestimmung der Schiefe `S` und des Exzess `E` berechnet.
Der z-Test wird außerhalb der Funktion „calc.point.distribution“ vorgenommen (siehe Zeile 845).

Listing 2: in Funktion „calc.point.distribution“

```
# declaration variables
N        = nr.participants
X        = mean(data$punkte)
h        = -1
sx       = sd(data$punkte)
sx3      = sx^3
sx4      = sx^4

# range of points participants got
p.range = diff(range(data$punkte))+1

# for a small number of participants (N<200):
# why N < 200? see Lienert[p. 151]
# create classes of points in groups of size h
# find suitable group-size
if(N < 200) {
        for(h in 2:5) { if(p.range %% h == 0) break;}
}

nr.groups = p.range/h
tab.81 = data.frame(cbind(class=1:nr.groups, Xj=0, xj=0, f=0, fx3=0, fx4=0))

# fill tab.81 with values for each group
start = min(data$punkte)
for(i in 1:nr.groups) {
        end = start + h-1
        tab.81$class[i] = paste(start, "-", end)
        tab.81$Xj[i]    = mean(c(start,end))
        tab.81$xj[i]    = tab.81$Xj[i] - X
        tab.81$f[i]     = sum(p.distr[start:end,])
        tab.81$fx3[i]   = tab.81$f[i] * (tab.81$xj[i])^3
        tab.81$fx4[i]   = tab.81$f[i] * (tab.81$xj[i])^4

        start = end + 1
}

S = sum(tab.81$fx3) / (N * sx3)
E = sum(tab.81$fx4) / (N * sx4)
...
```

Formel 4: Reliabilität

Funktion: `calc.internal.consistency`

Zur Bestimmung der Reliabilität wurde die interne Konsistenz - genauer gesagt der α-Wert - des Tests bestimmt.

Zur Bestimmung eines α-Wertes für ein Items i, werden alle Kombinationen dieses Items mit den anderen Items durchgegangen, für die $i \neq j$ gilt. Aus der Summe der gegeben Antworten zum Quadrat und der Quadrat-Summe der Antworten ergibt sich der Wert `sj2`.

Weiterhin benötigt zur Bestimmung der internen Konsistenz, den über alle Items gleichen Wert `sx2`, bei dem alle erreichten Punkte über alle Teilnehmer (`sum(data[[points.in.data]])`) mit einfließen.

Listing 3: in Funktion „calc.internal.consistency“

```
N = nr.participants

# go through all items i
for(i in 1:n.items) {

        # calculate sj^2 for each item j
        # without current item i
        sj2 = 0
        not.answered = 0
        for(j in 1:n.items) {
                if(i != j) {
                        ...
                        # original: sj2 = sj2 + 1/(N-1) * (sum(answers.only
                            ^2) - ((sum(answers.only)^2) / N))
                        # -not.answered for "listenweisen ausschluss"
                        sj2 = sj2 + 1/(N-1-not.answered) * (sum(answers.only
                            ^2) - ((sum(answers.only)^2) / (N-not.answered)))

                } else {}       # do nothing, cause i == j
        }

        sx2 = 1/(N-1) * (sum(data[[points.in.data]]^2) - ((sum(data[[points.
            in.data]])^2) / N))

        # calc alpha
        alpha   = n.items/(n.items - 1) * (1 - (sj2 / sx2))
        ...
}
...
```

Formel 5: Schwierigkeitsindex

Funktion: `calc.diff.index`

Der Schwierigkeitsindex eines Items (`index`) bestimmt sich über das Verhältnis der richtigen Antworten `n.right` zu den insgesamt gegebenen Antworten `n`. Die Anzahl der maximal gegebenen Antworten entspricht der Anzahl der Versuchspersonen `n`.

Listing 4: in Funktion „calc.diff.index“

```
...
n = nr.participants

n.right = 0
not.answered = 0
for(j in 1:length(answers[,col])) { # add items only if answered (!=99)
        if(answers[j,col] != NO_ANSWER) {
                n.right = n.right + answers[j,col]
        } else {
                not.answered = not.answered+1
        }
}

# calc index
index   = 100 * n.right / (n - not.answered) # -not.answered == "listenweiser
     Ausschluss"
...
```

Formel 6: Trennschärfe

Funktion: `calc.selectivity` und `help.selectivity`

Die Trennschärfe eines Items `rpbis` errechnet sich über die Anzahl `n` der Versuchspersonen die einen jeweiligen Rohwert X (Punkte) erreicht haben (dieser Wert `n` bleibt über alle Items konstant) und über die Anzahl der Teilnehmer, die mit X Punkten das aktuell zu berechnende Item richtig beantwortet haben (`n.right`).

Die Funktion `help.selectivity` berechnet die Werte für die richtigen Antworten zum Item, welche sich je Item unterscheiden.

Listing 5: in Funktion „calc.selectivity“

```
...
select.frame = data.frame(cbind(X=1:(n.items+1)), row.names=c(as.character(1:
    n.items),"result"))
...
# calc constant columns 2 to 4
# ("n", "n.X", "n.X2")
for(row in 1:(nrow(select.frame)-1)) {
        points = select.frame$X[row]

        select.frame$n[row]             = sum(sapply(data[[points.col.name]],
             function(x) {x == points}))
        select.frame$n.X[row]   = select.frame$n[row] * points
        select.frame$n.X2[row]  = select.frame$n[row] * points^2
}
...
# for all items
for(i in 1:(length(answers)-1)) { # - 1 for UserID
        # create new columns
        # via help function "help.selectivity"
        select.frame = help.selectivity(answers[,i], answers$UserID, points.
            col.name, select.frame)
```

```
        # result row = last row
        result.row = nrow(select.frame)
        # calc result row (sum of a column)
        for(col in 1:ncol(select.frame)) {
                select.frame[result.row, col] = sum(select.frame[1:(result.
                    row-1),col])
        }
        ...

        # assign variables last value in each column from select.frame
        ...

        # calc rpbis
        rpbis   = (n.right.X/ n.right - n.X/ n) * sqrt(n.right/ (n-n.right))
            * n/ sqrt(n*n.X2 - n.X^2)
        ...
}
```

Formel 7: Selektionskennwert

Funktion: `calc.select.value`

Bei der Berechnung des Selektionskennwertes `S` eines Items fließen zum Einen der Schwierigkeitsindex (hier `p`) und zum Anderen die Trennschärfe (hier `r`) des jeweiligen Items mit ein.

Listing 6: in Funktion „calc.select.value“

```
...
# declare variables
p = result$diff.index[j] / 100
q = 1 - p
r = result$selectivity[j]

# calculate value for item selection
S = r / (2 * sqrt(p * q))
...
```