

Julius-Maximilians-Universität Würzburg

Einführung in wissenschaftliches Arbeiten

Mitschrift der Vorlesung von Dipl.-Psych. Ingo Totzke
im SS 2008

Roland Pfister



Inhaltsverzeichnis

0. Vorwort	3
1. Was ist wissenschaftliches Arbeiten?	4
1.1. Richtlinien für die wissenschaftliche Praxis	4
1.2. Wissenschaftliches Arbeiten	6
2. Literaturarbeit	7
2.1. Internetressourcen	7
2.1.1. Online-Nachschlagewerke.....	7
2.1.2. Online-Tageszeitungen und -Magazine.....	7
2.1.3. Freie Internetressourcen für Psychologen.....	8
2.1.4. E-Books und elektronische Zeitschriften	8
2.2. Datenbankrecherche	8
2.2.1. Datenbanken und Kataloge	8
2.2.2. Datenbanken für Psychologen.....	8
2.2.3. Suche in Datenbanken	8
2.2.4. Spezielle Datenbanken (revisited).....	10
2.2.5. Übung im CIP-Pool.....	11
2.3. Informationsbeschaffung und -bewertung III	13
2.3.1. Informationen im Internet.....	14
2.3.2. Bewertung von Quellen	14
2.4. Informationsintegration	15
2.4.1. Visualisierung von Quelleninhalten.....	16
2.4.2. Kodierung	16
2.4.3. Anmerkung I: Primärstudien	17
2.4.4. Anmerkung II: Wikipedia.....	17
3. Berichtslegung	18
3.1. Schriftliche Darstellung wissenschaftlicher Arbeiten	18
3.1.1. Struktur von Berichten	18
3.1.2. Stil und Argumentation	19
3.1.3. Zitierung und Literaturverzeichnis.....	23
3.1.4. Positionierung von Veröffentlichungen	24
3.2. Visuelle Darstellung wissenschaftlicher Arbeiten	25
3.2.1. Eigene Visualisierungen	26
3.2.2. Tipps.....	26
3.3. Präsentation	27
3.3.1. Konzeption.....	27
3.3.2. Durchführung.....	27
3.3.3. Diskussion	28

3.3.4.	Nachbereitung	28
4.	Datenanalyse mit Excel und SPSS.....	29
4.1.	Allgemeine Anmerkungen.....	29
4.2.	Excel.....	29
4.2.1.	Statistische Zusatzfunktionen	29
4.2.2.	Zufallszahlen	29
4.2.3.	ANOVA mit Excel	30
4.3.	SPSS	30
5.	Quantitatives Vorgehen	31
5.1.	Deskriptive Analyse.....	31
5.1.1.	Numerische Kennwerte	31
5.1.2.	Graphische Darstellung: Präsentation	33
5.2.	Graphische Analyse.....	35
5.2.1.	Datenkontrolle	36
5.2.2.	Vergleich empirischer Verteilungen	38
5.2.3.	Kategorisierung von Informationen.....	39
5.2.4.	Prüfung von Voraussetzungen / Verteilungen I: Varianzanalyse	39
5.2.5.	Prüfung von Voraussetzungen / Verteilungen II: Regressionsanalyse	41
5.3.	Inferenzstatistische Analyse.....	42
5.3.1.	Grundlagen der Inferenzstatistik.....	42
5.3.2.	Entscheidungsfehler und Teststärke	43
5.3.3.	Bedeutung des Stichprobenumfangs.....	43
5.3.4.	Praktische Bedeutsamkeit	44
5.4.	„Reading Statistics“	44
5.4.1.	Interpretation deskriptiver Informationen	44
5.4.2.	Interpretation inferenzstatistischer Informationen I: Inferenzstatistische Grundprinzipien	46
5.4.3.	Interpretation inferenzstatistischer Informationen II: Unterschiedshypothesen.....	46
5.4.4.	Interpretation inferenzstatistischer Informationen III: Zusammenhangshypothesen	46
5.4.5.	Interpretation graphischer Informationen.....	47

0. Vorwort

Dozent: Dipl.-Psych. Ingo Totzke

Termin: Di, 12:00-14:00 Uhr

Klausur: - Termin: ???
- Mündliche Prüfung

Web: www.izvw.de bzw. www.psychologie.uni-wuerzburg.de/methoden

1. Was ist wissenschaftliches Arbeiten?

Wissenschaft zeichnet sich immer durch spezifisch methodisches Vorgehen aus. Nicht Inhalte einer Handlung, sondern das Wie des Handelns machen eine Wissenschaft auf – also der Umgang mit wissenschaftlichen Standards (Methoden). In der Grundstudiumsveranstaltung Forschungsmethoden der Psychologie wurde hierbei vor allem das Spiralenmodell von Sarris (1992) exemplarisch dargestellt.

Im deutschsprachigen Raum sind nach einer Analyse von Hager (2005) vor allem gerichtete Zusammenhangsfragestellungen anzutreffen, die mit Varianz- und/oder Regressionsanalysen geprüft werden. Gerne werden hierbei Voraussetzungen und statistische Probleme wie multiples Testen übersehen. Auch Effektstärken werden selten berichtet.

Nicht jede Forschung ist jedoch als wissenschaftlich zu bezeichnen. Streng genommen umfasst der Begriff **wissenschaftliches Arbeiten** die folgenden 5 Kriterien:

- Systematisch (zielgerichtet, methodisch kontrolliert)
- Vollständig
- Objektiv
- Allgemeingültig (in definierten Grenzen)
- Überprüfbar

1.1. Richtlinien für die wissenschaftliche Praxis

Die Universität Würzburg nennt hierbei einige Richtlinien zur Sicherung guter Wissenschaftlicher Praxis (vgl. auch Richtlinien.pdf). Diese sollen im Folgenden kurz behandelt werden.

§ 1 Grundsätze

- (1) Wissenschaftlerinnen und Wissenschaftler, die an der Julius-Maximilians-Universität Würzburg tätig sind, sind verpflichtet,
- lege artis zu arbeiten
 - Resultate zu dokumentieren
 - alle Ergebnisse konsequent selbst anzuzweifeln
 - strikte Ehrlichkeit im Hinblick auf die Beiträge von Partnern, Konkurrenten und Vorgängern zu wahren,
 - wissenschaftliches Fehlverhalten zu vermeiden und ihm vorzubeugen.

§ 4 Leistungs- und Bewertungskriterien

- Originalität und Qualität haben als Leistungs- und Bewertungskriterien für Prüfungen, für die Verleihung akademischer Grade, für Einstellungen, Beförderungen, Berufungen und Mittelzuweisungen stets Vorrang vor Quantität.
- Leistungs- und Bewertungskriterien sind an diesem Grundsatz auszurichten.

§ 5 Sicherung und Aufbewahrung von Primärdaten

- Primärdaten als Grundlagen für Veröffentlichungen sind auf haltbaren und gesicherten Trägern in der Einrichtung, in der sie entstanden sind, für zehn Jahre aufzubewahren.

§ 6 Wissenschaftliche Veröffentlichungen

- (1) Die Bezeichnung als Originalarbeit kann nur der erstmaligen Mitteilung neuer Beobachtungen oder experimenteller Ergebnisse einschließlich der Schlussfolgerungen zukommen. Demzufolge ist die mehrfache Publikation derselben Ergebnisse, abgesehen von vorläufigen Kurzmitteilungen in aktuellen Fällen, nur unter Offenlegung der Veröffentlichung vertretbar.
- (2) Wissenschaftliche Untersuchungen müssen nachprüfbar sein. Demzufolge muss ihre Publizierung eine exakte Beschreibung der Methoden und der Ergebnisse enthalten.
- (3) Befunde, welche die Hypothese der Autorin bzw. des Autors stützen oder sie in Frage stellen, sind gleichermaßen mitzuteilen.
- (4) Befunde und Ideen anderer Forschender sind ebenso wie relevante Publikationen anderer Autorinnen und Autoren in gebotener Weise zu zitieren.
- (5) Die Fragmentierung von Untersuchungen mit dem Ziel, die Anzahl scheinbar eigenständiger Publikationen zu erhöhen, ist zu unterlassen.

§ 7 Umgang mit wissenschaftlichem Fehlverhalten

Wissenschaftliches Fehlverhalten liegt vor, wenn in einem wissenschaftserheblichen Zusammenhang bewusst oder grob fahrlässig Falschangaben gemacht werden...

- a) Erfinden von Daten,
- b) Verfälschung von Daten und Quellen oder das Unterdrücken relevanter Quellen,
- c) unrichtige Angaben in einem Bewerbungsschreiben oder einem Förderantrag (einschließlich Falschangaben zum Publikationsorgan und zu in Druck befindlichen Veröffentlichungen),
- d) unrichtige Angaben zur wissenschaftlichen Leistung von Bewerberinnen und Bewerbern in Auswahlkommissionen;

... geistiges Eigentum anderer verletzt wird...

- a) unbefugte Verwertung oder Anmaßung der Autorenschaft (Plagiat),
- b) Ausbeutung von insbesondere noch nicht veröffentlichten Forschungsansätzen und Ideen, insbesondere als Gutachter (Ideendiebstahl)
- c) Anmaßung oder unbegründete Annahme wissenschaftlicher Autor- oder Mitautorenschaft,
- d) Verfälschung des Inhalts,
- e) unbefugte Veröffentlichung oder unbefugtes Zugänglichmachen gegenüber Dritten, solange das Werk, die Erkenntnis, die Hypothese, der Lehrinhalt oder der Forschungsansatz noch nicht veröffentlicht ist,
- f) Inanspruchnahme der (Mit-)Autorenschaft einer anderen Person ohne deren Einverständnis.

... oder die Forschungstätigkeit Anderer beeinträchtigt wird

- a) Sabotage von Forschungsvorhaben anderer

- b) Beseitigung von Primärdaten, soweit damit gegen gesetzliche Bestimmungen oder fachspezifisch anerkannte Grundsätze wissenschaftlicher Arbeit verstoßen wird.

Wissenschaftliches Fehlverhalten besteht auch in einem Verhalten, aus dem sich eine Mitverantwortung für das Fehlverhalten anderer ergibt, insbesondere durch:

- (1) Aktive Beteiligung am Fehlverhalten anderer,
- (2) Mitwissen um Fälschungen durch andere,
- (3) Mitautorenschaft an fälschungsbehafteten Veröffentlichungen, oder
- (4) Grobe Vernachlässigung der Aufsichtspflicht.

1.2. Wissenschaftliches Arbeiten

Wie bereits genannt, ist wissenschaftliches Arbeiten grundsätzlich systematisch, vollständig, objektiv, allgemeingültig und überprüfbar. Dabei sind die Kernbestandteile der Tätigkeit des wissenschaftlichen Arbeitens jedoch immer **Informationsverarbeitung** im Sinne von Informationsbeschaffung (Literaturrecherche), -aufnahme und -auswertung sowie die **Gestaltung schriftlicher Arbeiten** hinsichtlich Gliederung, Aufbau, korrektes Zitieren und Bibliographieren.

Als Erweiterung muss jedoch auch die **Informationsvermittlung** durch Vorträge, Präsentationen etc. in den Kanon des wissenschaftlichen Arbeitens mit aufgenommen werden.

Die folgenden Kapitel gliedern sich daher in die Themenblöcke der Informationsverarbeitung (Literaturarbeit, quantitatives Vorgehen) und der Gestaltung schriftlicher Arbeiten (Berichtslegung).

2. Literaturarbeit

Zunächst werden Informationsbeschaffung und Informationsbewertung (2.1, 2.2, 2.3) und anschließend die Informationsintegration (2.4) betrachtet

2.1. Internetressourcen

Im Folgenden werden behandelt:

- Online-Nachschlagewerke
- Online-Tageszeitungen und -Magazine
- Freie Internetangebote für Psychologen
- Elektronische Bücher und Zeitschriften



Viele Lizenzpflichtige Nachschlagewerke können über das Uni-Netzwerk genutzt werden. Hierfür wird der Cisco Systems VPN-Client oder der Docweb-Proxy-Server des Rechenzentrums benötigt (<http://docweb.rz.uni-wuerzburg.de/>).

Generell: Zuständig für die Internetressourcen für Psychologie und Pädagogik ist Viola Hämmner (0931 / 888-5909 bzw. viola.haemmer@bibliothek.uni-wuerzburg.de); auch für Probleme, Wünsche und Anregungen zu E-Books.

Für die Mitschriften aus den Übungssitzungen, siehe 2.2.5.

2.1.1. Online-Nachschlagewerke

- **Brockhaus Enzyklopädie:**
http://www.bibliothek.uni-regensburg.de/dbinfo/einzeln.phtml?bib_id=ub_wue&colors=63&ocolors=40&titel_id=5687
- **Duden:**
http://www.bibliothek.uni-regensburg.de/dbinfo/einzeln.phtml?bib_id=ub_wue&colors=63&ocolors=40&titel_id=423
- **Wörterbuch Psychologie (W. D. Fröhlich):**
http://www.bibliothek.uni-regensburg.de/dbinfo/einzeln.phtml?bib_id=ub_wue&colors=63&ocolors=40&titel_id=1605
- **Oxford Dictionary of Psychology (Colman A. M.):**
http://www.oxfordreference.com/views/BOOK_SEARCH.html?book=t87&authstatuscode=202.
- **Langenscheidts Fachwörterbuch Medizin (inkl. Übersetzung von Fachbegriffen):**
http://www.bibliothek.uni-regensburg.de/dbinfo/einzeln.phtml?bib_id=ub_wue&colors=63&ocolors=40&titel_id=893
- **Psyhyrembel Klinisches Wörterbuch (inkl. englischer Übersetzung der Begriffe im Glossar):**
http://www.bibliothek.uni-regensburg.de/dbinfo/einzeln.phtml?bib_id=ub_wue&colors=63&ocolors=40&titel_id=84
- **Rote Liste:** Informationen zu Arzneimitteln, z.B. Psychopharmaka inklusive Substanzklassen, Einzelstoffen, Handelsnamen, Nebenwirkungen etc. Wichtig: Die Rote Liste Online ist mit reduzierten Einträgen frei im Web verfügbar, Rote Liste Software bietet eine vollständige Datenbank, auf die nur im Hochschulnetz zugegriffen werden kann.
http://www.bibliothek.uni-regensburg.de/dbinfo/einzeln.phtml?bib_id=ub_wue&colors=63&ocolors=40&titel_id=90

2.1.2. Online-Tageszeitungen und -Magazine

Mehrere 1000 Zeitungen und Zeitschriften sind gleichzeitig und im Volltext über die Datenbank **LexisNexis** durchsuchbar:

http://www.bibliothek.uni-regensburg.de/dbinfo/einzeln.phtml?bib_id=ub_wue&colors=63&ocolors=40&titel_id=1670

2.1.3. Freie Internetressourcen für Psychologen

- Linkkatalog Psychologie ([PsychLinker](#)):
<http://www.zpid.de/redact/category.php?cat=1>
- Suchmaschine Psychologie ([PsychSpider](#)):
http://www.zpid.de/PsychSpider_info.php

2.1.4. E-Books und elektronische Zeitschriften

Derzeit findet sich eine relativ große Auswahl an E-Books von Springer und verschiedenen englischsprachigen Verlagen, nicht jedoch von Hogrefe, Beltz oder Kohlhammer.

Eine Übersicht findet sich auf der [E-Book-Seite der UB](#):

<http://www.bibliothek.uni-wuerzburg.de/digitalebibliothek/>

Zu elektronischen Zeitschriften existiert ebenfalls ein Zugang über die [Elektronische Zeitschriftenbibliothek \(EZB\)](#):

<http://rzb1x1.uni-regensburg.de/ezeit/search.phtml?bibid=UBW&colors=7&lang=de>

2.2. Datenbankrecherche

2.2.1. Datenbanken und Kataloge

Grundsätzlich muss zwischen Datenbanken und Katalogen unterschieden werden. [Kataloge](#) (z. B. [OPAC](#)) enthalten nur Bücher und keine Aufsätze. Sie sollten also dann herangezogen werden, wenn man schon weiß, welches Buch bzw. welche Zeitschrift man benötigt.

Ist dies nicht der Fall oder will man sich vertieft mit einem Thema beschäftigen, so sollte man auf [Datenbanken](#) zurückgreifen, die schwerpunktmäßig meist Aufsätze enthalten. Ein Beispiel ist das Datenbankinformationssystem [DBIS](#) der UB Würzburg:

<http://www.bibliothek.uni-wuerzburg.de/>

2.2.2. Datenbanken für Psychologen

- [PsyndexPlus](#): Enthält Monographien und Aufsätze der deutschsprachigen Literatur sowie psychologische Tests.
- [PsycINFO](#): Enthält überwiegend englischsprachige Literatur (ebenfalls Monographien und Aufsätze).
- [Science Citation Index](#): Fachübergreifende Datenbank mit speziellen Suchmöglichkeiten.

2.2.3. Suche in Datenbanken

2.2.3.1. Genauigkeit vs. Vollständigkeit

Dieser Themenkomplex wird im Kontext der Metaanalyse (Methoden der Analyse qualitativer Daten) nochmals in Form der Recall-Precision-Curves aufgegriffen.

Insgesamt liefert eine sehr genaue Suche nach dem spezifischen Thema (z.B. Verhaltenstherapie und Nikotinabhängigkeit) wenige, dafür aber relevante Treffer. Eine breite Suche (etwa Therapie und Sucht) liefert viele Treffer, die jedoch nicht zwingend relevant sein müssen.

2.2.3.2. Auswahl und Kombination von Suchbegriffen

Bei einer Suche erscheinen (**Titel**)**Stichworte** genau so im Dokument, es wird also zwischen „Klientenzentrierter Therapie“ und „Gesprächspsychotherapie“ unterschieden, was gerade bei thematischen Suchen problematisch ist. Dagegen sind Schlagworte normierte Begriffe und umfassen Synonyme, andere Sprache usw. Hier gibt es nur das eine **Schlagwort (Subject Heading)** „Client Centered Therapy“.

Wichtige Verknüpfungen zur Suche nach mehreren Wörtern sind und-, oder- und nicht-Verknüpfung. Bei thematischen Suchen, werden quasi immer Schlagworte verwendet.

Bei jeder Suche sollten nicht nur die tragenden Begriffe einer Fragestellung verwendet werden. Vielmehr sollte auch nach Synonymen, verwandten Begriffen, Übersetzungen, Schreibvarianten, spezifischeren und allgemeineren Begriffen gesucht werden.

Tragende Begriffe	Synonyme, verwandte Begriffe	Übersetzungen, Schreibvarianten	Spezifischere Begriffe	Allgemeinere Begriffe
Informationsverarbeitung	Kognitive Prozesse	information processing, cognitive processes	Urteilsbildung, Wahrnehmung	Psychische Prozesse
Emotionen	Gefühle	emotion(s), feelings	Freude, Traurigkeit, ...	Psychische Prozesse

2.2.3.3. Eingabe von Suchbegriffen

Wenn man mit Stichworten und nicht mit Schlagworten sucht, gibt es einige hilfreiche Techniken, vor allem Trunkierung und Maskierung.

Unter **Trunkierung** versteht man die Verwendung eines Platzhalters, der beliebig viele Zeichen ersetzt. Das häufigste Trunkierungszeichen ist dabei der *. Phob* findet also bspw. Phobie, phobia, Phobien, phobisch...

Maskierung hingegen ist die Verwendung eines Platzhalters, der genau ein Zeichen ersetzt. Andere Bezeichnungen sind **Wildcard** oder **Joker**. Das häufigste Maskierungszeichen ist das ?. Der Suchbegriff sui?id? findet also bspw. suicide, Suizid, Suizide...

2.2.3.4. Reduktion der Trefferzahl

Wenn zu viele Treffer erzielt wurden, gibt es zwei Möglichkeiten: Suche thematisch spezifizieren oder Limits setzen.

Die **Spezifikation** kann sehr einfach über weitere und/oder spezifischere Schlagwörter erfolgen. Reizvoll ist die Nutzung von **Classification Categories**. Bspw. bietet sich für eine Suche nach Interventionsmöglichkeiten im betrieblichen Kontext neben den Schlagwörtern „Alcoholism“ und „Psychotherapy“ die Verwendung des Classification Code 36 (= Industrial and Organizational Psychology) an.

Das Setzen von **Limits** kann sich auf verschiedene Aspekte der Suche beziehen, etwa Erscheinungsjahr, Sprache, Population, Altersgruppe, Methode.

2.2.3.5. Treffer exportieren

Unter den erzielten Treffern kann eine Auswahl getroffen werden. Diese Auswahl kann auf Festplatte gespeichert, per Mail gesendet oder gedruckt werden.

Wichtig: Bei Verwendung der Standardeinstellung wird das Abstract nicht mit exportiert. Das muss extra ausgewählt werden.

2.2.3.6. Dokumente öffnen

Über den [SFX-Button](#) kommt man direkt zum entsprechenden Dokument, sofern es verfügbar ist. Der Button stellt eine Verlinkung in die EZB dar, die Zugriff auf den Volltext vieler Artikel bietet. Bücher werden direkt im Katalog der UB (OPAC) angezeigt.

Die erste Alternative ist dabei immer die beste.

Eine andere Alternative ist der nur bei manchen Treffern angezeigte [Digital Object Identifiziert \(DOI\)](#), der ebenfalls ein Link zum Dokument ist.

2.2.3.7. Auf dem Laufenden bleiben

Um nicht ständig alle interessanten Gebiete nach neuen Artikeln durchsuchen zu müssen, gibt es einen [Alerting Service](#) ([PsycINFO: selective dissemination of information SDI](#)), der sich über die Search History einrichten lässt.

Einmal eingerichtet informiert der Alerting Service per Mail über neue Treffer unter den entsprechenden Schlagworten.

2.2.4. Spezielle Datenbanken (revisited)

Die Datenbank [PsycINFO](#) ist die größte psychologische Datenbank weltweit mit ca. 2.5 Mio. Einträgen im April 2008. Sie enthält Literatur aus Psychologie und Nachbardisziplinen, wobei etwa 2150 Zeitschriften ausgewertet werden.

Vorwiegend finden sich englischsprachige Einträge, aber auch die wichtigsten deutschsprachigen Zeitschriften sind enthalten. Die Datenbank wird dabei wöchentlich aktualisiert.

Die Datenbank [PsyndexPlus](#) ist die größte Datenbank zur deutschsprachigen psychologischen Literatur und besteht aus 2 Teilen: Literatur mit über 185.000 Literaturnachweisen, die monatlich aktualisiert werden und Tests mit über 5.000 Tests mit teilw. detaillierter Beschreibung (halbjährliche Aktualisierung).

Der [Thesaurus](#) beinhaltet die Schlagwörter ([PsycINFO: Index Terms](#)), die hierarchisch und thematisch miteinander verknüpft sind. Der Befehl [Search Marked](#) sucht genau das ausgewählte Schlagwort, [Explode Marked](#) sucht das gewählte Schlagwort und alle verknüpften Unterbegriffe. Der Thesaurus eignet sich auch zur Generierung von Synonymen (bei Stichwortsuchen).

Der [Science Citation Index](#) ist eine fächerübergreifende Datenbank. Dabei ist auch eine Suche über Zitierungen möglich ([Find Related Records](#): Findet Artikel mit ähnlichen Literaturverzeichnissen, [Times Cited](#): Findet neuere Artikel, die die gefundenen Artikel zitieren).

Der Science Citation Index ist dabei mit dem Journal Citation Index verknüpft und listet die [Impact Faktoren](#) der jeweiligen Zeitschrift auf.

2.2.5. Übung im CIP-Pool

2.2.5.1. Technisches

- Bücher bestellen geht auf der UB-Seite unter Quicklinks: „Frag die UB“ → Bücher bestellen oder per Mail an Viola Hämmer
- Zugang zur UB:
 - [Cisco Systems VPN-Client](#)
 - [DocWeb-Server des Rechenzentrums](#). Zugang erhält man auf der Unibib-Seite unter Infos von A-Z → E → Externer Zugang. Dort meldet man sich mit seiner Novell-Kennung an.
Zwar sind hier nicht alle UB-Dienste freigeschaltet, aber elektronische Zeitschriften, eBooks und Datenbanken über das Datenbank-Informationssystem (DBIS) sind weitgehend verfügbar. Alle Dienste der UB, die auf dem zentralen bayerischen CD-Rom-Server hinterlegt sind (etwa das Handwörterbuch der Psychologie oder das Wörterbuch der Psychologie) sind allerdings nur über den VPN-Client zugänglich.
- Quicklinks der EZB: eZeitschriften lesen

2.2.5.2. Angebote

- [Brockhaus online](#): Momentan können nur zwei Personen aus dem Hochschulnetz gleichzeitig zugreifen. Dort hat man allerdings auch Zugang auf die dpa-Bilddatenbank.
Starten kann man den Brockhaus über DBIS>Allgemeines>Enz & Nachschl>Brockhaus oder über den UB-Katalog (Suche nach Brockhaus und Dokumenttyp g: Online-Dokument). Wenn man im Hochschulnetz ist, wird man auf www.brockhaus-enzyklopaedie.de automatisch eingeloggt.
- [Britannica](#): Unbegrenzte Zugänge unter www.search.eb.com. Toll ist die Zitatsuche (Quotations) zu einem Suchwort.
- [Oxford Reference](#) (www.oxfordreference.com): Wörterbücher, Zitate, Nachschlagewerke, etc. Dabei können alle angebotenen Datenbanken auch zusammen durchsucht werden. Das kann auch interessant sein, um zu sehen, wie Begriffe in verschiedenen Disziplinen verwendet werden (Depression in der BWL oder Meteorologie). Bei den einzelnen Treffern wird angezeigt, in welchem Bereich ein Begriff gefunden wurde.
Zugang über: DBIS, eBookseite der UB>Digitale Bibliothek>eBooks >Fachübersicht. Psychologie findet sich hier unter Rechts-, Sozial- und Geisteswissenschaften. An die gleiche Stelle gelangt man mit dem DocWeb>eBooks.
- [Oxford Reference: Wörterbücher Psychologie](#). Psychologie gehört hier zu den Social and Political Sciences; Alternativ gelangt man hier über DBIS>Psychologie und Katalog>Suche nach Online-Dokument hierher. Die wichtigsten Wörterbücher sind Dictionary for Psychology und der Companion to the Mind (Kognitionswissenschaft). Tolles Werkzeug: Widen Your Search sucht nach Schreibvarianten etc.
- [Duden](#): Hier braucht man den VPN, da er nur als CD-Rom-Version verfügbar ist. Man kann ihn auch über DBIS und den Katalog erreichen.

- **Elektronische Zeitungen:** Erreichbar über Digitale Bibliothek > Elektronische Zeitungen. LexisNexis enthält dabei mehrere tausend nationale und internationale Zeitungen.
- **PsychSpider:** Psychologie-spezifische Suchmaschine. Hauptsächlich sind die Seiten psychologischer Institute indexiert.
- **Zentrum für psychologische Information und Dokumentation (www.zpid.de):** Psycholinker bietet eine psychologische Linksammlung.
- **Datenbank Diplomarbeiten im Fach Psychologie:** Eine Komplettsammlung von Diplomarbeiten seit 1998.

2.2.5.3. Handwerkszeug I: Allgemein

- Schlagwortsuchen (Subject Headings) sind Stichwortsuchen vorzuziehen
- Bei Stichwortsuchen ist die Trunkierung mächtig: * oder \$
- Aufpassen beim Nicht-Operator: Wenn im Titel steht „Es geht nicht um bipolare Störungen“, dann wird bei der Stichwortsuche „Depression und nicht bipolar“ dieser Artikel natürlich nicht gefunden.
- Fernleihe: Dauert 7-10 Tage und kostet 1,50 €.

2.2.5.4. Handwerkszeug II: PsycINFO

- Auf der UB-Seite ist die Journal Coverage List angegeben. Ausgewertet werden über 2000 Zeitschriften seit 1887.
- Wöchentliches Update: Attraktiv ist die Automatisierung einer Suche mit Benachrichtigung über RSS-Feed oder eMail.
- Dissertationen sind sauschwer zu bekommen, v.a. weil sie häufig nicht einmal veröffentlicht sind.
- Auch sind wenige Bücher enthalten – und fast gar keine deutschsprachigen. Hier sucht man besser über den Katalog der UB.
- Spezielle Möglichkeiten:
 - **Basic Search:** Soll die Google-Suche imitieren und kann einen ersten Einstieg in die Suche darstellen. Das Thema sollte möglichst natürlichsprachlich eingegeben werden und Autoren dürfen nicht genannt werden.
Die Treffer werden nach Relevanz geordnet (was auch immer das genau ist) und man kann ähnliche Schreibvarianten über das Feld Related Terms einbeziehen.
 - **Find Citation:** Wenn man schon genau weiß was man will, lässt sich hier ein Literaturnachweis verifizieren.
 - **Search Tools:** Schlagwortverzeichnis. Hier muss man jeden Begriff einer neuen Suche einzeln überprüfen. Über das Kommando Explode werden Unterbegriffe mit in die Suche einbezogen. Was Focus bedeutet weiß keiner (keine Unterbegriffe?).
 - **Suchergebnisse:** Verknüpfung der Ergebnisse über die Search History möglich; auch empfiehlt es sich die angegebenen Schlagworte (Subject Headings) bei einem guten Treffer nochmals als Suchkriterium zu verwenden. Ähnlich funktioniert wohl Find Similar, aber keiner weiß, was das genau macht.
Der Button SFX bietet Zugang zum Volltext und stellt eine Verlinkung zum Bestand der UB dar. Oben sollte eigentlich immer die beste Alternative stehen.

- **Advanced Ovid Search:** Hier lässt sich die Suche eingrenzen. Praktisch sind die Additional Limits. Hier beziehen sich die Classification Codes XX_00_ auf Fächer und aller anderen Zahlen (nicht _00) auf Unterfächer.

2.2.5.5. Handwerkszeug III: Science Citation Index

- Synonym: Web of Knowledge
- Interdisziplinär, tolle spezielle Suchmöglichkeiten. Aber: Nur sinnvoll, wenn man schon etwas hat und ähnliches finden möchte.
- Weniger benutzerfreundlich als PsycINFO und findet viele irrelevante Artikel (Nachbardisziplinen etc.).
- Zugang zu Journal Citation Index und ImpactFactor.
- Related Records: Findet Artikel mit ähnlichen Literaturverzeichnissen (also eine ganz andere Art der Suche als Stich- und Schlagworte).
- Man kann der Datenbank sagen, dass man alle Nachweise haben möchte, die einen bestimmten Artikel zitieren.

2.3. Informationsbeschaffung und -bewertung III

Cooper (1984) unterscheidet verschiedene Informationskanäle:

- Primäre Kanäle: Bibliotheken, Analyse relevanter Arbeiten
- Sekundäre Kanäle: Bibliographien, Informationsstellen
- Informelle Kanäle: Eigene Forschung, Kontakt zu Wissenschaftlern, Kongresse
- Spezielle Informationssysteme bzw. Datenbanken: Dissertations Abstracts International, Social Sciences Citation Index, Psychological Abstracts oder PsycINFO, Psynex Plus, Medline, WISO
- Internetrecherche

Dabei lässt sich eine generelle Empfehlung für Suchstrategien aussprechen: Es sollte mit Enzyklopädien und Handbüchern begonnen werden¹. Danach werden Überblicks- bzw. Reviewartikel herangezogen (Psychological Bulletin, Psychological Review, Psychologische Rundschau) und erst dann werden Metaanalysen (Sekundärstudien) und Originalbeiträge (Primärstudien) herangezogen. Metaanalysen ermöglichen dabei Aussagen über Effektstärken und eventuelle Moderatorvariablen.

Enzyklopädien² → Handbücher → Überblicksartikel / Reviews →
→ Metaanalysen und Originalbeiträge

Unter einem **vertikalen Vorgehen** beim Suchen von Artikeln versteht man dabei nichts anderes, als das Lesen von Artikeln und das Durchforsten deren Literaturverzeichnisse, um evtl. weitere Artikel zu finden oder neue Suchen zu generieren – **ausgehend von einem zentralen Text**.

Im Folgenden werden nun die verschiedenen Informationskanäle kurz umrissen.

¹ „Es gibt hier kein richtig oder falsch; nur ein optimal oder nicht optimal“ ☺.

² In der Enzyklopädie der Psychologie werden alle denkbaren Themenkomplexe so kompakt wie nur irgend möglich dargestellt, sodass man hier einen guten Einblick in den Stand der Forschung erhält. Aber: „1000e Seiten ohne Bilder; wirklich hässlich und nicht schön zu lesen“.

2.3.1. Informationen im Internet

Relevante Suchstellen sind unter anderem:

- Verbände (z.B. APA, DGPs, BDP. DGPs ist super, v.a. hinsichtlich der Fachgruppen. Die beiden Verbände BDP und DGPs machen unterschiedliche Sachen und bekämpfen sich auch oft).
- Überblicksseiten zu psychologischen Fachbereichen, 5 Jahre alte Linksammlung enthält nur noch:
 - www.neuropsychologycentral.com
- Buch- und Bibliothekskataloge
 - www.buchhandel.de: Verzeichnis lieferbarer Bücher
 - www.ubka.uni-karlsruhe.de/hylib/virtueller_katalog.html: Virtueller Katalog Karlsruhe als Meta-Katalog für Bibliotheks- und Buchhandelskataloge
- Institute
- Sonstiges: Siehe F14 und 15 für Links zu Listen psychologischer Institute, Stellenbörse PsychJob, Psychologischer Kongresse, Psychologenzwitsche (toll für Präsentationen vor fachfremdem Publikum) etc.

Anmerkung: Google. Kulturelles Wissen ist im Internet frei verfügbar und kann daher auch mit Google erfasst werden. Wissenschaftliche Informationen sind dagegen nicht frei verfügbar, sodass Google in diesem Zusammenhang nicht immer sinnvoll einsetzbar ist. Wenn man aber mit Google sucht, gibt es einige nette Strategien, etwa das Eingrenzen der Suche auf .pdf-Dateien oder über den Befehl *define* nach Wortdefinitionen zu suchen.

2.3.2. Bewertung von Quellen

Die Validität von Quellen ist äußerst bedeutsam. Wie im Kontext der Metaanalyse besprochen sind neben den klassischen Validitätsbereichen (interne, externe, Konstruktvalidität und Validität des statistischen Schlusses) vor allem 4 weitere Kriterien relevant:

- Authentizität (authenticity)
- Glaubwürdigkeit (credibility)
- Repräsentativität (representativeness)
- Bedeutung/Sinngehalt (meaning)

Wichtige Hinweise zur Bewertung von Quellen sind daher unter anderem im **Text selbst** zu finden. Hier sind Aktualität und inhaltliche Struktur der Quelle von besonderer Bedeutung, aber auch Umfang, Aktualität und Internationalität des Literaturverzeichnisses sind von Bedeutung.

Neben der Quelle selbst lässt sich auch der **Autor** bewerten, etwa hinsichtlich seiner fachlichen Reputation (Zitierindizes³) oder der Zugehörigkeit zu einer Schule: „Ein guter Autor kann nur gute Texte schreiben“.

³ Beispiele sind der SSCI („Wie oft kommt ein Autor vor?“), der Impact Factor IF („Wie oft wird ein durchschnittlicher Artikel des jeweiligen *Journals* zitiert?“). Psychologische Zeitschriften haben dabei einen eher niedrigen IF („Die Psychologie ist träge.“), Human Factors und Ergonomics beispielsweise etwa IF = 0.8. Andere Lebenswissenschaften bringen es gut auf IF = 15.

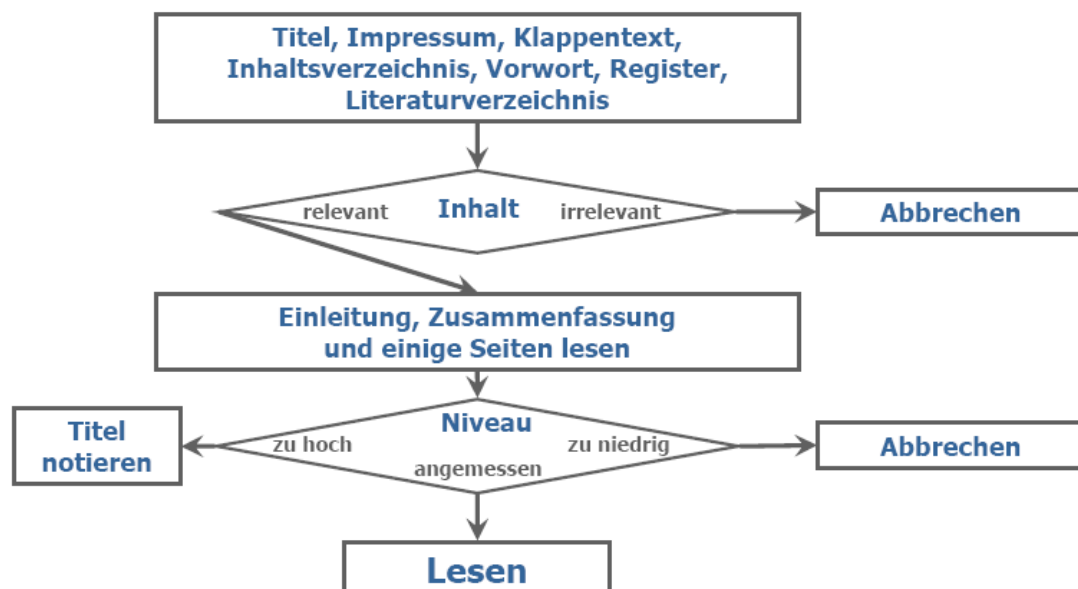
Desweiteren liefert auch der **Veröffentlichungskontext** wie Zielpublikum, Renommée von Verlag und Zeitschrift oder die Auflage der Quelle wichtige Informationen.

Eine gute Zeitschrift mit hoher Auflage wird beispielsweise einen Artikel genau prüfen, bevor er veröffentlicht wird. Das macht ihn glaubwürdiger, aber die Artikel dieser Zeitschrift werden weniger aktuell sein. Auch sollte man im Kopf behalten, dass manche Zeitschriften verlangen, dass ihre Artikel von eingereichten Manuskripten zitiert werden, bei Ergonomics beispielsweise mindestens 3 Stück.

Insgesamt sollte dabei folgendermaßen vorgegangen werden:

- Von jüngeren zu älteren Publikationen
- Publikationen von Wissenschaftlern für Wissenschaftler bevorzugen
- Zeitschriften und Bücher sind besser als Papers und das Internet
- Autoren hinterfragen (Stellung, Schule, Arbeitsgebiete)
- Guter Stil ist Indikator für gute Arbeit

Rost (2003) hat ein Schema hierfür ein Schema für die Relevanzprüfung von Quellen erstellt:



2.4. Informationsintegration

Nach der Literatursuche müssen relevante Informationen aus den gefundenen Quellen extrahiert und kodiert werden. Relevante Informationen sind nach Sonntag (2006):

- Ziel und Fragestellung der Untersuchung
- Theoretischer Ansatz
- Hypothesen
- Untersuchungsdesign
- Variablen
- Statistische Verfahren
- Zentrale Ergebnisse
- Stärken und Schwächen
- Bedeutung für die eigene Forschungsfrage

Eine Quelle sollte dabei zunächst überflogen werden. Danach sollten alle Quellen in einer Tabelle organisiert werden und erst dann kategorisiert werden.

Stichworte zur Charakterisierung des Quelleninhalts beinhalten häufig das Problem von ungenauen oder missverständlichen Bezeichnungen (F26).

2.4.1. Visualisierung von Quelleninhalten

Zur Visualisierung von Quelleninhalten stehen verschiedene Methoden zur Verfügung. Die wichtigsten sind: Ablaufdiagramme, Mind-Maps, Hierarchische Darstellungen, Kreuztabellen und Kausaldiagramme.

2.4.1.1. Ablaufdiagramme

Ablaufdiagramme bestehen aus einer schematischen Darstellung der Abfolge der zentralen Inhalte und orientieren sich dabei bspw. an der Gliederung des Texts bzw. wichtigen verwendeten Konzepten und Begriffen.

Die Erstellung eines Ablaufdiagramms setzt jedoch eine einfache, eindimensionale Abfolge der einzelnen Inhalte voraus.

2.4.1.2. Mind-Maps

Mind-Maps ermöglichen die Darstellung komplexer Zusammenhänge. Dabei führen Verästelungen von einem zentralen Punkt aus in verschiedene Richtungen (Gedanken).

Die Schwerpunkte der Argumentation werden so verdeutlicht. Auch sind Änderungen und Ergänzungen leicht möglich, ohne die Gesamtstruktur zu verändern, was Mind-Maps vor allem bei schlecht strukturierten Quellen attraktiv macht.

2.4.1.3. Hierarchische Darstellungen

Hierarchische Darstellungen stellen feste hierarchische Beziehungen zwischen den Inhalten her und sind somit nur begrenzt anwendbar. In der Regel werden sie als Baumdiagramme dargestellt.

2.4.1.4. Kreuztabellen

Bei zwei oder mehr Gliederungskriterien bieten sich Kreuztabellen an. Gezielte Leerräume können hier beispielsweise hervorgehoben werden.

2.4.1.5. Kausaldiagramme

Kausaldiagramme sind die komplexeste Art der grafischen Darstellung. Das Ziel ist die umfassende Darstellung von Zusammenhängen zwischen Variablen, bspw. die Visualisierung einer Theorie.

Die Diagramme der LISREL-Verfahren sind hierbei nur eine von vielen Alternativen.

2.4.2. Kodierung

Eine umfassendere Darstellung der Kodierung findet sich in der Ausarbeitung zur Metaanalyse (Methoden der Analyse qualitativer Daten). Ausgehend von einem Kodierplan werden hier möglichst mehrere Kodierer vorbereitet, welche

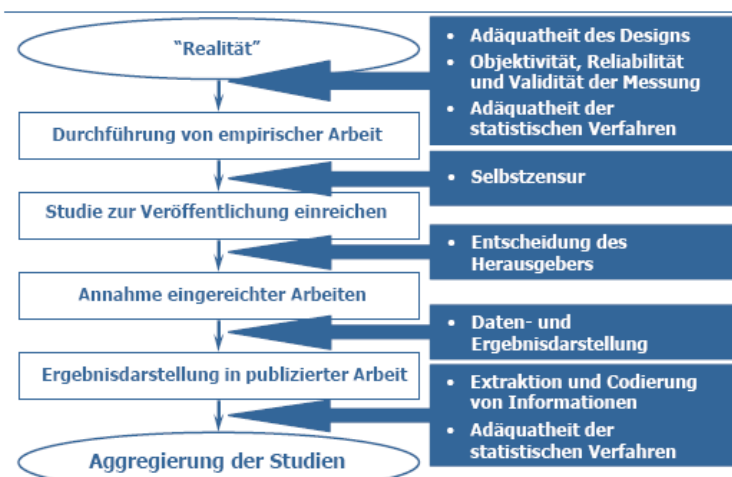
bei einer genügend hohen Beobachterübereinstimmung eine systematische Kodierung gewährleisten.

Hinsichtlich der Auswahl der zu kodierenden Variablen lassen sich folgende Merkmale unterscheiden:

- **Substanzielle Merkmale der Primärstudie:** Inhaltlich relevante Merkmale der Untersuchungsgruppe, der Behandlung, der Rahmenbedingungen sowie der Messinstrumente.
- **Verzerrende Merkmale der Primärstudie:** Qualitätsmerkmale des Versuchsdesigns.
- **Extrinsische Merkmale der Primärstudie:** Generelle Eigenschaften (Autoren, Publikationsjahr) und editoriale Merkmale (Titel der Quelle).

Idealerweise werden mindestens zwei unabhängige Beurteiler verwendet und die Beurteilung ist blind gegenüber Methoden- und Theorieteil. Der Genauigkeit der Beobachtung zuträglich ist zudem die Entwicklung und Validierung eines Kodierhandbuches, spezifischer Kodierblätter und eine Beurteilerschulung.

2.4.3. Anmerkung I: Primärstudien



Neben den genannten Kriterien zur Bewertung von vorhandenen Studien sollte auch immer berücksichtigt werden, dass vor der Veröffentlichung einer Studie ein Prozess steht, der die vorgefundene Information entschieden beeinflussen kann.

[Vgl. Darstellung zur Metaanalyse].

2.4.4. Anmerkung II: Wikipedia

Auf Wikipedia und dem Wiki-Brockhaus-Vergleich soll hier nicht weiter herumgeritten werden, aber es gibt einen kleinen Hinweis von Ingo. Scheinbar sollte man einem Wikipedia-Artikel nur dann glauben, wenn mindestens 50-60 Autoren mitgewirkt haben. Daher ist Wikipedia auch unschlagbar bei Lokalpatriotismus und Sex 😊.

3. Berichtslegung

3.1. Schriftliche Darstellung wissenschaftlicher Arbeiten

3.1.1. Struktur von Berichten

Was braucht ein Bericht?

- Titel
- Einleitung
- Theorieteil (evtl. Hypothesen/Fragestellung gesondert)
- **Ergebnisse** (evtl. Auswertungsstrategie gesondert)
- Diskussion
- Literaturverzeichnis
- Anhang (Testunterlagen, Instruktionen, Originaldaten, Syntax, Tabellen und Graphen, evtl. Ergebnisse die im Ergebnisteil nicht dargestellt wurden, weil exemplarisch nur Item 1 und 2 von 50 ausgewertet wurden).

3.1.1.1. Diplomarbeit

Die Diplomarbeit gliedert sich in 9 Teile (Ellgring, 2000):

- Zusammenfassung (1-2 Seiten)
- Einleitung (2 Seiten)
- Stand des Problems (Theoretischer Teil; 10-20 Seiten)
- Spezifische Fragestellung und Hypothesen (1/2-2 Seiten)
- Methoden (10-20 Seiten)
- Ergebnisse (25-40 Seiten)
- Diskussion (5-10 Seiten)
- Literaturverzeichnis nach APA-Standard
- Anhang

Der (bzw. jeder) Hauptteil sollte dabei nach der **4-3-2-Regel** gegliedert werden. Er sollte sich aus 4 Kapiteln zusammensetzen, die jeweils mindestens 3 Unterkapitel mit je mindestens 2 Gliederungspunkten aufweisen.

Unterkapitel müssen dabei logisch strukturierte Teilaspekte des Kapitelthemas behandeln. Dabei ist ein Oberpunkt immer nur dann ein Oberpunkt, wenn er mindestens in zwei Teilaspekte untergliedert werden kann. Gliederungsebenen mit nur einem Gliederungspunkt darf es nicht geben. Auch ist es nicht empfehlenswert, mehr als 3 Hierarchieebenen zu verwenden (im Notfall evtl. 4).

Der Gliederungsaufbau sollte dabei themengerecht und übersichtlich erfolgen. Die Ausführlichkeit der Behandlung einzelner Hierarchiestufen sollte dabei einheitlich erfolgen. Was das Layout angeht (Schriftgröße, Zeilenabstand etc.) hat jeder Lehrstuhl andere Vorstellungen – hier sollte man sich also mit seinem Betreuer absprechen. Generell gilt es aber, möglichst sparsam mit Sonderformatierungen (F, K, U) oder gar Kombinationen davon umzugehen.

3.1.1.2. Publikation: Empirische Arbeit

Die Gliederung einer Publikation sollte hingegen aus folgenden Punkten bestehen:

- Abstract
- Einleitung, Literaturüberblick und Kurzdarstellung der Fragestellung
- Methodisches Vorgehen
- Statistisches Vorgehen und Ergebnisse
- Diskussion der Ergebnisse (bei kurzen Artikeln auch in Kombination mit dem Ergebnisteil möglich)

3.1.2. Stil und Argumentation

Die drei wichtigsten Kriterien zur Art des Schreibens sind **Genauigkeit**, **Klarheit** und **Verständlichkeit**. Vor allem bzgl. der Klarheit sollte dabei ein einfacher, direkter Stil gewählt werden. Das heißt auch: Kein interner Thesaurus! Ein gutes Beispiel sind die Begriffe 1\$-Bedingung vs. 20\$-Bedingung bei Festinger. Häufig sind die Begriffe am besten, die man auch bei der Kommunikation mit dem Betreuer verwendet.

Auch sollte der Artikel gut organisiert werden (Sanduhrprinzip, übersichtliche Darstellung) und die innere Logik der Argumentation deutlich werden. Hier ist ein häufiger Fehler, Gedanken nicht zu Ende zuführen (**ABAB-Fehler**).

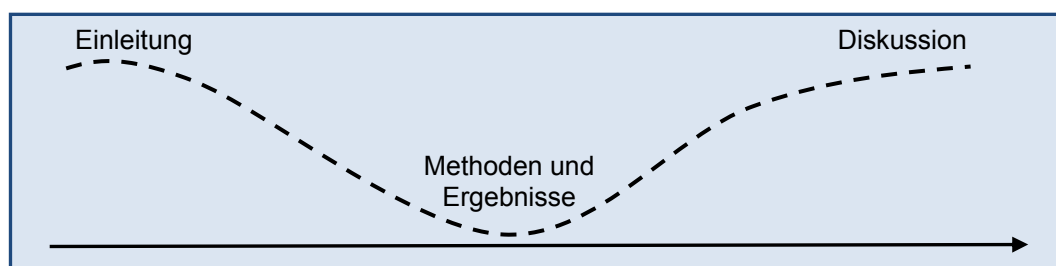
Nach Madigan et al. (1995) lassen sich einige charakteristische Merkmale psychologischer Fachtexte identifizieren:

- Psychologische Aufsätze stammen häufig von mehreren Autoren.
- Psychologen zitieren viel.
- Psychologen zitieren nicht wörtlich, sondern paraphrasieren.
- Psychologen benutzen kaum Fußnoten, dafür aber umso mehr Unter-Überschriften.
- Psychologen widersprechen gerne anderen Autoren, aber sie bleiben dabei immer freundlich.
- Theoretische Aussagen, selbst wenn sie empirisch geprüft wurden, werden häufig vorsichtig ausgedrückt (Konjunktiv).
- Die Struktur des Experimentalberichts ist deduktiv (vom Gesetz zum Einzelfall = von der Theorie zur Methode).

3.1.2.1. Form eines Artikels

Die Einleitung beginnt breit, wird aber schnell spezifischer. Nach der Zitierung einiger empirischer Argumente wird schließlich die eigene Studie eingeführt.

Am spezifischsten sind Methoden- und Ergebnisteil. Die Diskussion beginnt mit den Implikationen der Studie und wird danach wieder breiter (Auch: Sanduhrmetapher).



Die einzelnen Teile sollen nun etwas detaillierter betrachtet werden.

3.1.2.2. Teil 1: Einleitung

Es sollte nicht mitten im Thema, sondern besser mit einer formalen oder theoretischen Einleitung begonnen werden. Die Einleitung sollte insgesamt gut verständlich sein (kein psychologischer Jargon!) und theoretische Argumente sollten durch Beispiele veranschaulicht werden. Aber: Nicht im ersten Satz mit irgendwelchen Jahreszahlen anfangen...man sollte nach dem ersten Satz Lust haben, den Artikel zu lesen und nicht abgeschreckt werden ☺.

Mögliche **Einstiege** sind:

- Persönliche Aufhänger
- Aktuelle Bezüge
- Historischer Bezug
- Provokation
- Ankündigung
- Betonung des Nutzens
- Erzählen einer Geschichte
- Herausstellen einer konkreten Empfehlung am Anfang
- Teilhabe an Erkenntnisprozess

Gerade die Ankündigung oder aber auch die Betonung des Nutzens bieten sich für wissenschaftliche Artikel dabei weniger an. Schöner sind hier die erstgenannten Einstiege. Für Präsentationen bieten sich auch provokante Fragestellungen an – diese locken Publikum.

Nach dem Einstieg ins Thema sollte eine kurze Zusammenfassung der vorliegenden Studien und Theorien gegeben werden, am besten in chronologischer Reihenfolge. Für korrekte Zitierung siehe Kapitel 3.1.3.1, S. 23. Für jedes größere Argument bietet sich dabei ein schematischer Aufbau der eigenen Argumentation an:

Hinführung zum Thema → Behauptung → Beleg → Einschränkung
→ Ergebnis (eigene Wertung)

Zum Abschluss der Einführung sollte schließlich eine kurze Übersicht über die eigene Studie gegeben werden. Dieser beinhaltet einen kurzen Abriss über das methodische Vorgehen sowie evtl. eine Einführung in die Hypothesen.

Zur Bewertung der Einleitung liegt eine Kriterienkatalog von Sonntag (2006) vor:

- Kommt die Darstellung ohne langwierige Ausführungen schnell und direkt zur Fragestellung?
- Wird die Forschungsfrage direkt benannt?
- Wird beschrieben, warum die Forschungsfrage relevant ist?
- Fokussiert die Darstellung ausschließlich auf Konstrukte und Prozesse, die für die Untersuchung unmittelbar relevant sind?
- Werden Ausführungen zu Konstrukten und Prozessen, die nur lose mit dem Thema zu tun haben, vermieden?
- Werden die zentralen Konstrukte definiert?
- Werden die Hypothesen explizit benannt?
- Werden die Hypothesen ausreichend begründet?

3.1.2.3. Teil 2: Methoden

Das Grundprinzip des Methodenteils sollte darin bestehen, den Leser durch den Versuch zu führen, als ob er selbst teilgenommen hätte. Zunächst wird jedoch ein kurzer Überblick über die methodischen Aspekte der Studie gegeben.

Dazu gehören Beschreibung der Stichproben (N, n, Männer/Frauen, Alter, Studenten?, Kontaktaufnahme, Vergütung, Selektionskriterien), Beschreibung des Settings (Materialien, Versuchsplanung) und Beschreibung der AVn.

Danach folgt eine Beschreibung des Vorgehens aus Sicht der Teilnehmer inklusive Zusammenfassungen der Instruktionstexte, Präsentation von Beispielen und Kopien des Stimulusmaterials. Schließlich wird das Vorgehen nochmals mit Blick auf das Ziel der Studie kurz zusammengefasst.

Richtlinien für eine genaue Abfolge der verschiedenen Aspekte gibt es dafür nicht; hier hat man also ein wenig Spielraum.

3.1.2.4. Teil 3: Ergebnisse

Im Ergebnisteil werden die aufbereiteten Daten und deren statistische Auswertung präsentiert. Dabei gilt der Grundsatz: „Give the forest first and then the trees.“. Es sollte also mit zentralen Ergebnissen anfangen und erst später auf periphere Ergebnisse eingegangen werden.

Bevor inferenzstatistische Kennwerte dargestellt werden, *müssen* deskriptive Kennwerte im Text bzw. in einer Tabelle mitgeteilt werden. Die Darstellung inferenzstatistischer Kennwerte erfolgt dabei immer nach folgendem Schema:

Symbol (Freiheitsgrad/e) = Wert, p = Wert ohne führende Null.

Das Signifikanzniveau wird dabei immer exakt angegeben.

Beispiele:

t-Test:

$$t(50) = 3.521, p = .01$$

Varianzanalyse:

$$F(3, 37) = 1.253, p = .12$$

Chi2-Test (zusätzlich Stichprobengröße in Klammern):

$$\chi^2(3, N = 125) = 12.10, p = .01$$

Anmerkung: Tabellen und Abbildungen. Tabellen und Abbildungen können das Lesen erschweren, wenn zu viele davon verwendet werden. Zudem müssen sie selbsterklärend sein, da Leser zuerst darauf schauen.

Sie müssen immer nummeriert werden und es muss im Text darauf Bezug genommen werden. Häufig sind die Standardeinstellungen von Statistikprogrammen zur Präsentation ungeeignet. Eine geschickte Alternative ist das Grafikpaket SigmaPlot von Stata.

Bei umfangreichen Auswertungen bietet es sich zudem an, einige Grafiken und Tabellen in den Anhang auszulagern.

3.1.2.5. Teil 4: Diskussion

Bei kurzen Artikeln kann die Diskussion auch in Kombination mit dem Ergebnisteil geführt werden. Immer sollte sie mit der Einleitung eine Einheit bilden, anhand derer der Leser die Studie weitgehend verstehen können soll.

Am Anfang der Diskussion ist ein deutliches Statement zugunsten oder gegen die Hypothese sehr geschickt. Zudem sollten die Ergebnisse mit bereits vorliegenden Ergebnissen anderer Studien verglichen werden.

Eine Diskussion sollte außerdem Überlegungen zur Generalisierbarkeit der Befunde auf Basis von Stichprobenmerkmalen und Merkmalen der verwendeten Methoden enthalten. Auch sollten offene bzw. unbeantwortete Fragen angesprochen werden.

Und abschließend: Die Diskussion sollte kurz sein und bestenfalls mit einem „BANG!“ beendet werden 😊 ([near cosmic significance](#)).

3.1.2.6. Titel und Abstract

Titel und Abstract [werden in Datenbanken aufgenommen](#). Daher ist eine genaue und präzise Formulierung notwendig, die den Inhalt des Artikels repräsentiert. Beide sollten daher erst nach Beendigung des Schreibens formuliert werden. [Zusätzlich: [Angabe von Keywords](#)].

Der [Titel sollte maximal 10-12 Wörter](#) umfassen und eigenständig verständlich sein. Aus ihm sollte hervorgehen, welche theoretischen Aspekte oder Variablen untersucht wurden. Empfehlungen finden sich auf F41.

Das Abstract sollte nicht länger als 120 Wörter sein (APA-Empfehlung). Kurz und präzise soll dabei auf Fragestellung, Stichprobe, experimentelles Vorgehen, statistisches Vorgehen, Ergebnisse, Schlussfolgerungen und Implikationen eingegangen werden.

Eine Checkliste zur Berichtslegung findet sich auf den Folien 43-45. Eine kleine Anmerkung von Ingo: Im Studium hört man das Wort Berichtslegung quasi nicht, sobald man aber selbst Forschungsberichte verfasst, sind die letzten Wochen eines Projekts von diesem Schlagwort geprägt.

3.1.2.7. Anmerkung: Stilistische Tipps

- Vermeidung von umgangssprachlichen Wörtern und Wendungen.
- Korrekter Gebrauch von *das selbe* und *das gleiche*.
- Vermeidung von Meta-Kommentaren.
- Synonyme verwirren!
- Ähnliche Einleitungssätze verwenden.
- Fachbegriffe können treffender/präziser sein als Alltagssprache oder mit bestimmten relevanten Konzepten einhergehen. Ist keines der beiden Argumente relevant: Alltagssprache.
- Verwendung der Ich/Wir-Perspektive anstelle von „der Autor...“. Aber: Ich/Wir-Sätze sollten vermieden werden (passivische Konstruktion). Auch sollte der Leser im Hintergrund gehalten werden.
- Zeit: Literaturübersicht, Methoden in der Vergangenheit, Ergebnisse und Diskussion in der Gegenwart.

Neben diesen allgemeinen Tipps gibt es auch einige Hinweise zur spezifischen Bedeutung einzelner Ausdrücke:

It has been long known	I haven't bothered to check the references
It is known	I believe
It is believed	I think
There has been some discussion	Nobody agrees with me
Some samples were chosen	The others didn't make sense
Typical results are shown	The best results are shown
The results are inconclusive	The results seem to disprove my hypothesis
Additional work is required	Someone else can work out the details
The results will be in a subsequent report	I might get around to this sometime, if pushed/funded

(http://www.xs4all.nl/~jcdverha/scijokes/8_2.html#subindex; im LP-Ordner)

3.1.3. Zitierung und Literaturverzeichnis

[Nach den DGPs-Richtlinien (1997)].

3.1.3.1. Zitierung im Text

Regeln für die Zitierung im Text [Beispiele finden sich auf den Folien 53ff]:

- **Einzelautor:**
 - Nach einer belegenden Aussage: (Name, Jahr)
 - Ist der Name Bestandteil des Texts wird das Jahr unmittelbar danach in Klammern gebracht: Schon Name (Jahr)...
 - Sind Name und Jahr Bestandteil des Texts entfällt der Hinweis in Klammern. [Problem: Was ist mit Müller (1986a)...das würde ja dann untergehen...].
- **Zwei Autoren:**
 - Wie bei Einzelautor; beide Autoren werden immer zitiert. Wenn die Autoren in Klammern stehen werden die Namen mit einem kfm. & getrennt.
- **Mehr als 2, weniger als 6 Autoren:**
 - Beim ersten Bezug im Text werden sämtliche Autoren aufgeführt. Die Namen werden dabei durch Kommata getrennt: Name, Name, Name und Name (Jahr).
 - Nachfolgende Angaben verwenden nur noch den Namen des Erstautors und den Zusatz et al.
- **6 oder mehr Autoren:**
 - Auch beim erstmaligen Bezug im Text nur Anführung des ersten Namens und der Angabe et al.

Empfehlungen: Falls bei Abkürzung der Literaturangabe identische Quellenangaben auftreten sind stets alle Autoren zu nennen [Bsp.: Müller et al. (1994) mit zwei verschiedenen et al.'s].

Auch sollten Autoren insgesamt besser in Klammern als im Text aufgeführt werden. Eine explizite Aufnahme in den Text ist vor allem dann sinnvoll, wenn auf die Forscher selbst fokussiert werden soll.

3.1.3.2. Literaturverzeichnis

- **Zeitschriftenartikel:**
 - Autor(en) (Jahr). Titel des Artikels. *Name der Zeitschrift, Band, Seitenangaben.*
 - Bsp.: Borg, I. (1984). Das additive Konstantenproblem der multidimensionalen Skalierung. *Zeitschrift für Sozialpsychologie, 15*, 248-253.
- **Bücher:**
 - Autor(en) (Jahr). *Buchtitel.* Verlagsort: Verleger.
 - Bsp.: Szagun, G. (1980). *Sprachentwicklung beim Kind.* München: Urban & Schwarzenberg.
- **Beiträge in Herausgeberbänden:**
 - Autor(en) (Jahr). Beitragstitel. In Herausgebername(n) (Hrsg.), *Buchtitel* (Seitenangaben). Verlagsort: Verleger.
 - Bsp.: Döbert, R. & Nummer-Winkler, G. (1984). Abwehr- und Bewältigungsprozesse in normalen und kritischen Lebenssituationen. In E. Olbricht & E. Todt (Hrsg.), *Probleme des Jugendalters. Neuere Sichtweisen* (S. 259-295). Berlin: Springer-Verlag.
- **Literatur aus dem Internet:**
 - Bislang keine allgemein verbindlichen Richtlinien. Vorschlag:
 - Autor(en). Titel des Artikels (Web-Dokument). Verfügbar unter (Available): Pfadangabe (Datum des Zugriffs).
 - Bsp.: Vollrath, M. & Totzke, I. (2000). In-Vehicle Communication and Driving: An Attempt to Overcome their Interference (Web-Dokument). Driver Distraction Internet Forum. Available: <http://www-nrd.nhtsa.dot.gov/departments/nrd-13/driverdistraction/Papers.htm> (22.10.2001).

Anmerkung: Die angegebenen Richtlinien beziehen sich auf die Manuskripterstellung, wenn Paper vom Herausgeber nicht gesetzt werden. Falls das Paper doch gesetzt wird, dann wird eine *kursive* Schriftart durch Unterstreichungen ersetzt.

3.1.4. Positionierung von Veröffentlichungen

Die Wahl der Zeitschrift für die Publikation richtet sich sowohl nach inhaltlich-fachlichen Kriterien, als auch des **Impact-Factors** einer Zeitschrift. Der Impact-Factor wird im jährlichen Journal Citation Report veröffentlicht und beruht u. a. auf dem Social Science Citation Index (SSCI).

$$\text{IF 2002} = \frac{\text{Zitate im Jahr 2002 aus einer Zeitschrift}}{\text{Anzahl der veröffentlichten Artikel in einer Zeitschrift}}$$

In der Regel werden dabei nur etwa 15-20% der eingereichten Paper direkt akzeptiert. Im Normalfall werden also Korrekturen vor der Neuverlage des Pa-

pers notwendig. Die häufigsten Kritikpunkte sind hierbei Uneindeutigkeiten und Erklärungslücken.

Daher sollte ein Paper auch vor der erstmaligen Einsendung einige Tage liegen gelassen werden und erst dann überarbeitet werden, wobei Unklarheiten sofort beseitigt werden sollten. Auch bietet sich eine Korrektur durch Kollegen an.

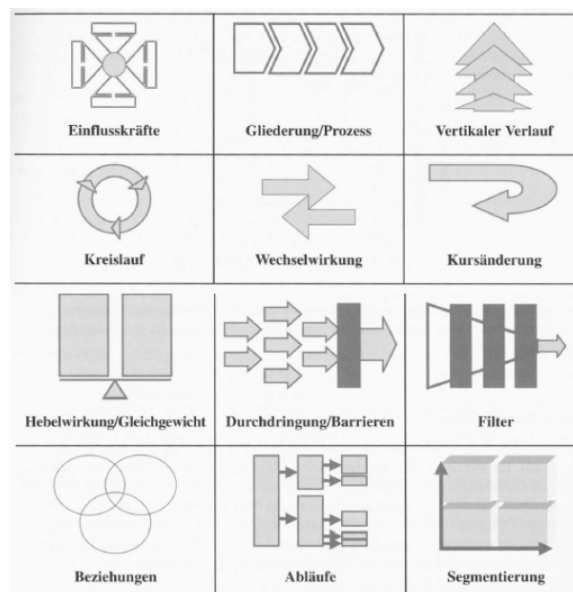
Literatur:

- Sehr gut verständliche und nachvollziehbare **Tipps** zur Publikation: **Bem, D.J. (2002). Writing the empirical journal article.** In J.M. Darley, M.P. Zanna, & H.L. Roedinger III (Eds), *The complete academic: A career guide.* Washington, DC: American Psychological Association.
- **Zitierregeln für den deutschsprachigen Raum**; Richtlinien zur formalen Manuskriptgestaltung: Deutsche Gesellschaft für Psychologie (1997). *Richtlinien zur Manuskriptgestaltung.* Göttingen: Hogrefe.
- **Zitierregeln für den englischsprachigen Raum**, Übersicht über Gestaltung und Organisation einer Publikation und Tipps zum sprachlichen Ausdruck: American Psychological Association (2001). *Publication Manual of the American Psychological Association.* Washington, DC: American Psychological Association.
- Madigan, R., Johnson, S. & Linton, P. (1995). The language of psychology – APA style as epistemology. *American Psychologist*, 50 (6), 428-436.

3.2. Visuelle Darstellung wissenschaftlicher Arbeiten

Visualisieren = Veranschaulichen von Inhalten. Die Visualisierung soll dabei **motivieren, das Verstehen erleichtern und das Behalten unterstützen.**

Dabei stehen verschiedenste Möglichkeiten zur Veranschaulichung von Material zur Verfügung. Eine kleine Auswahl wichtiger Diagramme ist auf der folgenden Seite dargestellt.



Desweiteren finden sich auf F6 einige Anmerkungen zum Anwendungsbereich verschiedener graphischer Darstellungen. Vor allem für die Rezeption empirischer Arbeiten ist hierbei wichtig im Hinterkopf zu behalten, dass graphische Darstellungen auch immer ein hohes Missbrauchspotential bieten. Einfache Beispiele sind:

- Stauchung oder Entzerrung der X-Achse
- Darstellung lediglich eines Teils der Y-Achse (mit Sprung //)

Auch interessant sind Spielereien mit den Daten selbst, etwa die Errechnung hochpräziser Angaben aus eigentlich unpräzisen Daten, Verwendung von Anteilen ohne Anzahlen oder auch Spielereien mit Wachstum und Wachstumsraten:

„Getting information from a table is like extracting sunlight from a cucumber“
(Farquhar & Farquhar, 1891).

3.2.1. Eigene Visualisierungen

Die Art der Darstellung sollte auf Basis von Überlegungen zur Hauptaussage der Grafik geschehen. Entsprechend sollten auch Achsenskalierung und Beschriftung so gewählt werden, dass die Aussage für den Leser leicht verständlich ist. Die Vermischung verschiedener Aussagen in einer Grafik sollte generell vermieden werden (vgl. 5.1.2.6).

Checkliste für eigene Darstellungen:

- Ist die richtige Diagrammart gewählt?
- Gibt der Titel Auskunft über WAS / WO / WANN?
- Sind überflüssige Wörter eliminiert?
- Sind Linien und Flächen gut unterscheidbar?
- Ist alles durch eine Legende identifizierbar?
- Sind Ordinate und Abszisse gut gewählt und Beschriftungen leicht lesbar?
- Sind die Proportionen Flächen-Linien-Schriften gut gewählt?
- Keine überflüssigen Details?
- Ist der dargestellte Sachverhalt trivial?
- Zeigt die Grafik das Zahlenmaterial vollständig?

3.2.2. Tipps

Es sollten maximal drei Farben pro Abbildung verwendet werden. Für zusammengehörende Sachverhalte sollte die gleiche Farbe und Form verwendet werden und die Inhalte sollten räumlich gruppiert werden (Blöcke).

Wichtiges sollte hervorgehoben werden. Insgesamt lassen sich allerdings auch Freiflächen zur Gestaltung verwenden. Interessant ist außerdem eine freie Grafik („Bildchen“), die eine stimulierende Wirkung haben kann.

Generell gilt: Im Weglassen liegt die Kunst (KISS – keep it simple and stupid). Wenn die Visualisierung fertig ist, sollte sie nach Möglichkeit an Bekannten und Kollegen getestet werden.

3.3.Präsentation

3.3.1. Konzeption

Die Konzeption einer Präsentation sollte mindestens in den folgenden vier Schritten erfolgen: Analyse der Adressaten, Festlegung der Ziele, Analyse der Situation, Gliederung der Präsentation.

3.3.1.1. Analyse der Adressaten

Hier ist vor allem relevant, welches Vorwissen die Zuhörer mitbringen und welche Relevanz das Thema für sie hat bzw. wie sie gegenüber der Kernaussage eingestellt sind.

Auch sollte die Auswahl der verwendeten Medien dem Hörerkreis angepasst werden, nicht nur der reine Inhalt.

3.3.1.2. Festlegung der Ziele

Mögliche Ziele sind Informationsvermittlung, Argumentation und Überzeugung, Verstärkung eines Problembewusstseins oder aber auch reine Unterhaltung.

3.3.1.3. Analyse der Situation

Zentral hierbei ist die Frage nach der verfügbaren / einzuplanenden **Zeit**. Bei freier Wahl der Vortragsdauer sollten etwa 45 Minuten gewählt werden (Aufmerksamkeit: BRAC), wobei etwa 15% Pufferzeit eingeplant werden sollten.

Die Dauer des Vortrages sollte auch für die Vorbereitung der Präsentation herangezogen werden. Hier sinkt der Aufwand Vorbereitungszeit/Minute Vortrag, je länger die Vortragszeit ist.

Neben der Zeit sollte die Analyse der Situation jedoch auch unbedingt die Punkte **Medienwahl** und **räumliche Rahmenbedingungen** umfassen.

3.3.1.4. Gliederung der Präsentation

3.3.2. Durchführung

Zu Haltung und Verhalten während der Präsentation: Dem Publikum sollte Respekt und **Wertschätzung** gezeigt werden. Noch wichtiger sind jedoch begründetes **Selbstvertrauen**, fachliche **Überzeugung** und **Enthusiasmus!**

Generell empfiehlt sich eine deutliche, laute und dabei natürliche Aussprache und ein Sprechen zum Publikum (Augenkontakt). Eine ruhige Bauchatmung hilft gegen Nervosität, genauso wie das Verteilen des Gewichts auf beide Beine. Sachliche Fehler sollten nicht nur zugegeben werden, sondern man sollte sich auch beim Zuhörer für die Richtigstellung bedanken.

Damit möglichst viele Informationen beim Zuhörer ankommen (gleichzeitiges Lesen und Zuhören ist schwer), sollte man wichtige Folien kurz bevor man sie

wirklich zeigt ankündigen, nach dem Auflegen der Folie eine kurze Pause einlegen und sie danach erklären und abschließend Kommentieren.

Zur Fokussierung der Aufmerksamkeit auf sich selbst (weg von den Folien) bieten sich die **b/w oder ./,-Tasten in Powerpoint** an.

Anmerkung 1: Animationen. Einheitliche, einfache Animationen sollten generell vorgezogen werden. Wenn Animationen verwendet werden, sollten immer Sinneinheiten animiert werden.

Anmerkung 2: Was merkt sich der Zuhörer?

- Es waren Emotionen im Spiel.
- Es ist etwas Unerwartetes passiert.
- Es war Bewegung vorhanden.
- Sie hatten selber etwas damit zu tun.
- Die Situation war außergewöhnlich.
- Es wurde eine Gewohnheit durchbrochen.
- Es hat sich was Neues ereignet.
- Es ist eine Wahrheit ans Licht gekommen.
- Es geschah etwas, was zu einer Veränderung führte.

3.3.3. Diskussion

Nach Möglichkeit sollten auch für die Diskussion Folien bzw. unterstützende Materialien bereitgestellt werden. Mögliche Diskussionsthemen sind:

- Angewandte Methode
- Gewählte Theorie
- Präsentierte Definitionen
- Getroffene Annahmen
- Gütekriterien
- Nicht beantwortete Fragen

Zum generellen Verhalten während der Diskussion ist vor allem für den Umgang mit Fragen und Kommentaren anzumerken:

- Geduldig zuhören, danach: Denkpause
- Frage gegebenenfalls wiederholen
- Ausschließlich Beantwortung von Fragen, keine Kommentare
- Antworten an das gesamte Publikum gerichtet
- So lange bei Fragen verbleiben, bis sie beantwortet sind (evtl. beim Fragesteller nachfragen)

3.3.4. Nachbereitung

Nach der Präsentation sollte das Präsentationsmanuskript (Foliensatz, Handouts, Seminararbeit...) verschickt bzw. den Zuhörern zugänglich gemacht werden, bestenfalls mit der Bitte um Feedback.

Auch bieten sich private Gespräche mit einzelnen Zuhörern an. Eine geschickte Möglichkeit zur Evaluation einer längeren Veranstaltung ist es, leere Zettel zu verteilen und jeden Teilnehmer einen positiven Punkt (markiert mit einem +) und einen negativen Punkt (-) auf jeweils eine Seite zu notieren.

4. Datenanalyse mit Excel und SPSS

4.1. Allgemeine Anmerkungen

- Statistica gibt's auf der Rechenzentrums-Homepage zum Download
- Details zu bestimmten Verfahren finden sich in den jeweiligen Skripten
- AMOS 7 wird im Zuge der LISREL-Modelle (Skalierung) besprochen

4.2. Excel

- Darstellung: Für Zahlen gibt es verschiedene Formate (z.B. normal vs. Datum). Das Programm rechnet also evtl. mit anderen Zahlen, als man in der Darstellung sieht.
- Absolute vs. relative Bezüge. Standardmäßig sind relative Bezüge eingestellt; wird die Formel verschoben, erfolgt eine ebensolche Verschiebung der Bezugfelder. Absolute Bezüge können mit Dollarzeichen gesetzt werden (z.B. $\$C\16). Das Dollarzeichen vor dem Buchstaben setzt einen absoluten Zeilenbezug, das Dollarzeichen nach dem Buchstaben setzt einen absoluten Spaltenbezug.
- Wenn man bei Diagrammen als Fehlerindikator „Standardabweichung“ auswählt, wird die SD auf der Basis der Mittelwerte berechnet (also SS_{Treat} in der Sprache der Varianzanalyse). Sinnlos.

4.2.1. Statistische Zusatzfunktionen

Generell empfiehlt sich, die Add-Ins zu statistischen Funktionen zu aktivieren, sodass ANOVAs auch mit Excel berechnet werden können. Auch wird hierdurch die Generierung von Zufallszahlen möglich.

4.2.2. Zufallszahlen

In Excel stehen drei Möglichkeiten zur Erzeugung von Zufallszahlen zur Verfügung:

- Standardfunktion „[=Zufallszahl\(\)](#)“: Erzeugt gleichverteilte Pseudozufallszahlen zwischen 0 und 1. Achtung: Die so erstellten Zahlen ändern sich bei jeder Aktion in der Tabelle.
- Standardfunktion „[=Zufallsbereich\(\)](#)“: Wie oben, nur dass der Bereich definiert werden kann.

Beide Standardmöglichkeiten können also nur gleichverteilte Zufallszahlen erzeugen, deren Werte sich ständig ändern. Die erzeugten Daten sind also nicht reproduzierbar.

Daher ist die Verwendung der Analysefunktion „Zufallszahlengenerierung“ auch für experimentelle Designs wesentlich besser geeignet. Sie erlaubt die Erzeugung von Zufallszahlen auf Basis verschiedener Verteilungen unter der Angabe eines Ausgangswertes (Seed), sodass eine bestimmte Folge von Zufallszahlen erzeugt wird.

- Also: [Zufallszahlengenerierung unter Extras\Analysefunktionen](#).

4.2.3. ANOVA mit Excel

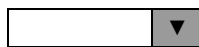
Wieder stehen 3 Möglichkeiten zur Verfügung: „Von Hand“ über Varianzschätzungen oder Winer-Kenngrößen (obsolet ☹) oder automatisch über die Analysefunktion „einfaktorielle Varianzanalyse“.

4.3. SPSS

- Jede Vpn eine Zeile (Fall bzw. Case)
- Missing Values werden durch einen Punkt im unteren rechten Bereich der Zelle dargestellt. Diese Zellen werden bei der Berechnung ausgeschlossen.



- Variablen umkodieren (z.B. String → Numerisch) unter *Transformieren*
- Manche SPSS-Module können auch mit Strings rechnen, die meisten funktionieren aber ausschließlich mit Zahlen.
- Wertelabels: Angeben, was für die Zahlencodes dargestellt werden soll (z.B. „männlich“ und „weiblich“ statt „0“ und „1“). Unter *Ansicht* kann eingestellt werden, ob Wertelabel oder Zahl dargestellt werden soll. Wie in Excel zeigt sich also, dass das was dargestellt wird und das mit dem das Programm rechnet nicht das gleiche sein muss.
- Steht für das Programm eine Zahl in einem Kästchen, so findet sich eine Labelanzeige im Auswahlmü:



- ANOVAS finden sich unter *Analysieren\Allgemeines Lineares Modell*
- Spezialität: Bei t-Tests werden Gruppenwerte vergeben. Man kann so z.B. Bedingung 1 mit Bedingung 5 vergleichen.
- SPSS-Algorithmen lassen sich mit visual basic programmieren.
- Hilfe: Command Syntax Reference, wenn's wirklich ernst wird („Was rechnet das Programm eigentlich?)
- Analyseschritte im Output protokollieren zu lassen ist ziemlich mächtig. Lässt sich unter *Extras* einstellen. Der Output lässt sich auch direkt in die Syntax einfügen.
- Tabellen im Output lassen sich per Doppelklick umformatieren

5. Quantitatives Vorgehen

5.1. Deskriptive Analyse

Dieser Teil beschäftigt sich mit Datenniveaus, Kennwerten der zentralen Tendenz, Dispersion und Verteilungsform sowie der Darstellung deskriptiver Kennwerte (Präsentationsgraphen).

Wie in jeder Abhandlung über statistische Grundbegriffe, soll auch hier die Tabelle zur den einzelnen Skalentypen nicht fehlen.

Skalentyp	Zulässige Transformation	Invariant bleiben	Beispiel
Nominal	Jede eineindeutige Funktion	Eindeutigkeit der Messwerte	Nummerierung, Geschlecht
Ordinal	Jede monoton steigende Funktion	Rangordnung der Messwerte	Härteskala, Schulnoten
Intervall	Jede positiv lineare Funktion $y = bx + c$ Mit $b > 0$	Verhältnisse der Intervalle der Messwerte	Temperatur, Nutzen
Verhältnis	Jede Ähnlichkeitsfunktion $y = bx$	Verhältnisse von Messwerten	SI-System
Absolut	Identitätsfunktion $y = x$	Messwerte	Häufigkeiten, W'keiten

Anmerkung 1: Ordered metric bezeichnet eine Mischform von Ordinal- und Intervallskala.

Anmerkung 2: Kardinalskala ist eine Bezeichnung für die Kombination von Intervall- und Verhältnisskala.

5.1.1. Numerische Kennwerte

Lediglich Erwähnung finden die **Kennwerte der zentralen Tendenz**, **Dispersion**, **Schiefte** (Abweichung von der Symmetrie) und **Exzess** (Steilheit, Gipflichkeit, Kurtosis).

5.1.1.1. Maße der zentralen Tendenz

Das **arithmetische Mittel** benötigt Intervallniveau, **geometrisches Mittel** (für Mittelung von sich multiplikativ verhaltenden Größen) und **harmonisches Mittel** (z.B. durchschnittliche Geschwindigkeit, Dichte oder Überlebenszeit) benötigen Verhältnisskalen.

Zur Abschätzung der Genauigkeit einer Schätzung sollte hier zusätzlich immer der Standardfehler SE berechnet werden.

$$SE = \hat{\sigma}_{\bar{x}} = \frac{\hat{\sigma}}{\sqrt{n}} \quad (\text{Punktschätzung})$$

Während die Berechnung des Standardfehlers die Genauigkeit der Punktschätzung eines Populationsmittelwertes ausgibt (Schätzung eines Populationsparameters durch einen Einzelwert), lässt sich auch das jeweilige Ver-

trauensintervall (Konfidenzintervall KI; engl.: CI) auf Basis des Prinzips der Intervallschätzung berechnet werden.

$$KI = \bar{X} \pm z_{(\alpha/2)} \cdot SE \quad (\text{Intervallschätzung})$$

Dabei steht $z_{(\alpha/2)}$ für ein beliebiges kritisches Quantil (etwa auch aus der t-Verteilung, wenn Stichprobenkennwerte nicht Normalverteilt oder $N \leq 30$).

Schließlich sollen auch **Median** (md) und Modalwert / **Modus** (mo) nicht übergangen werden.

5.1.1.2. Maße der Dispersion

Die **Standardabweichung (Varianz)** ist hier definitiv das wichtigste Maß. Fall eine Normalverteilung angenommen werden kann, lässt sich die Varianz folgendermaßen interpretieren:

- 68.3% der Messwerte liegen im Intervall $m \pm 1 s$
- 95.5% der Messwerte liegen im Intervall $m \pm 2 s$
- 99.7% der Messwerte liegen im Intervall $m \pm 3 s$

Alternativen bei niedrigeren Datenniveaus sind **Bereichsmaße** wie (gestutzte) Streubreite oder Interquartilsbereich für ordinalskalierte Daten oder das **Informationsmaß h** (Entropie) für nominalskalierte Daten.

Anmerkung: Die Streubreite lässt sich durch die Varianz der Messwerte schätzen:

- N < 10: Varianz * 2
- N = 10-20: Varianz * 3
- N ≈ 25: Varianz * 4
- N ≈ 100: Varianz * 5
- N > 100: Varianz * 6

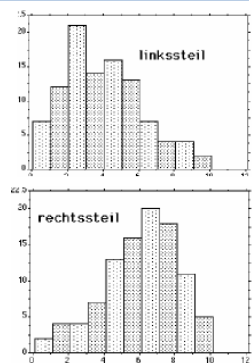
5.1.1.3. Schiefe

Die Bestimmung der Schiefe ist über einen Schiefekoeffizienten möglich. Bei einem positiven Schiefekoeffizienten wird die Verteilung als positiv schief bzw. rechtschief bzw. linkssteil bezeichnet:

Positiv schief → $mo < md < m$

Bei einem negativen Schiefekoeffizienten wird von negativ schief, linksschief bzw. rechtssteil gesprochen.

Negativ schief → $mo > md > m$



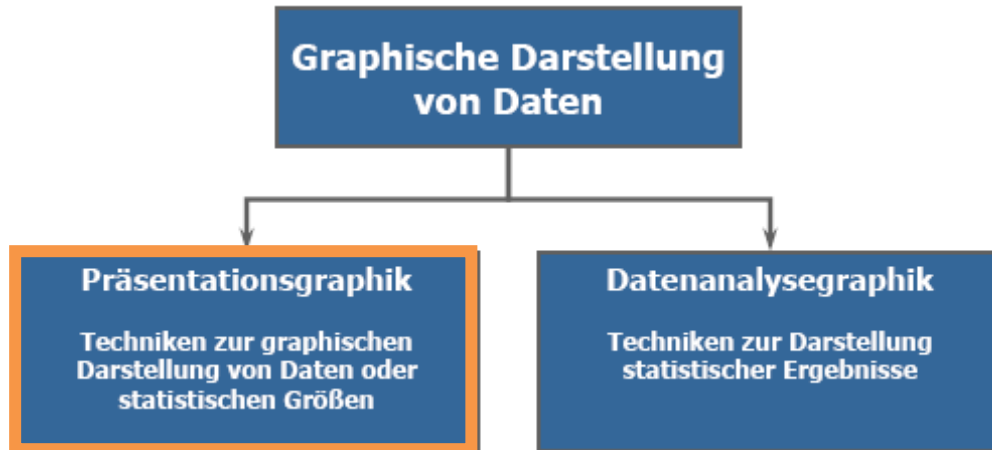
5.1.1.4. Exzess

Der Exzess gibt den Grad der Abweichung von der Normalverteilung an. Ein Exzess von 0 entspricht der Normalverteilung, negative werte sprechen für eine breitgipflige Verteilung, positive Werte für eine schmalgipflige Verteilung.

Falls der Exzess numerisch zwischen -1 und +1 liegt, kann von einer geringen Abweichung von der NV gesprochen werden. Bei extremeren Werten von einer starken Abweichung. Selbstverständlich kann der Exzess nur bei unimodalen Verteilungen sinnvoll verwendet werden.

5.1.2. Graphische Darstellung: Präsentation

Die Behandlung graphischer Elemente erfolgt hier in zwei getrennten Kapiteln. Zunächst soll die graphische Darstellung quantitativer Daten zur Präsentation deskriptiver Kennwerte betrachtet werden. Unter 5.2 wird der Einsatz von graphischen Techniken zur Analyse betrachtet.



Grundsätzlich wird für jede Art der grafischen Darstellung auf eine Anzahl verschiedener Darstellungstypen zurückgegriffen:

- Balkendiagramm (horizontal)
- Stabdiagramm (vertikal)
- Kreisdiagramm
- Linien diagramm (Abszisse nach Fällen sortiert)
- Flächendiagramm
- Scatterplot (Abszisse nach Größe sortiert)
- XYZ-Grafiken

Hinsichtlich ein-, zwei- und dreidimensionalen Darstellung muss jeweils der Unterschied zwischen Proportionalität von Linien, Linien oder Flächen bzw. Linien, Volumen oder Flächen beachtet werden.

5.1.2.1. Stabdiagramm

Richtlinien: Die Ordinate muss bei 0 beginnen und alle Stäbe müssen gleich breit sein. Bei wenigen Stäben sollte der Abstand der Stäbe die Hälfte der Stabbreite betragen, bei vielen Stäben sollte kein Abstand gelassen werden.

Die Unterteilung der Abszisse soll durchgehend (gleichabständig) sein. Sprünge müssen entsprechend markiert werden. Auch die Ordinate sollte nach Möglichkeit nicht unterbrochen werden, es sei denn, es liegen durchgehend extrem hohe Werte vor. Auch ein Unterbrechen der Stäbe ist streng genommen nicht zulässig.

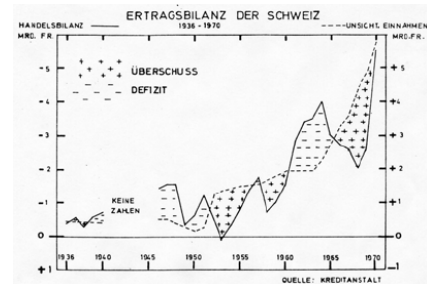
Mittelwerte immer mit Dispersionsmaß (und Angabe welches verwendet wurde). 3D ist schlecht.

Interessant sind **Komponentendarstellung**, welche den direkten Gruppenvergleich erschwert, dafür aber Aussagen über die Summe enthält, und **Verhältnisdarstellung** (Gegenteil).

5.1.2.2. Flächendiagramm

Auch hier ist zwischen **Komponentendarstellung** (Hauptaussage: Summe) und **Verhältnisdarstellung** (Hauptaussage: Anteile der Mengen) zu unterscheiden.

Ein weiterer Typ ist das **Zonendiagramm**, welches sich zur Darstellung von Überschüssen und Verlusten eignet.

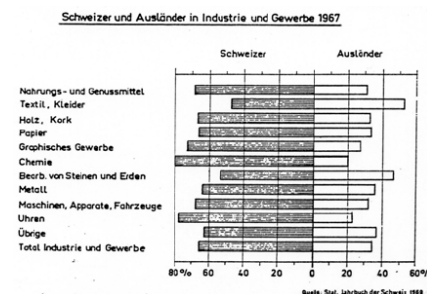


5.1.2.3. Balkendiagramm

Das Balkendiagramm eignet sich vor allem zur Darstellung zweier unterschiedlicher Maßzahlen, wenn unterschiedliche Konzentrationen oder Anteile betont werden sollen. Hier muss die Abszisse in beide Richtungen gleich skaliert werden.

Zwei-Komponenten-Balkendiagramme dienen der Darstellung einer Gesamtmasse, die in zwei Komponenten gegliedert ist. Sehr schön für die Darstellung von Korrelationen.

Auch hier muss die Abszisse in beiden Richtungen gleich skaliert werden.



5.1.2.4. Kreis-Diagramm

Werden verschiedene Kreisdiagramme verwendet, sollte die Gesamtgröße des Kreises proportional zur Größe der jeweiligen Bezugsgruppe gewählt werden (Achtung: Durchmesser vs. Fläche). Die Flächenproportionen entsprechen den Verhältnissen der einzelnen Subgruppen.

5.1.2.5. Liniendiagramm

Liniendiagramme sind vor allem bei vielen Datenpunkten angebracht und sind auch beim Vergleich mehrerer Reihen noch übersichtlich. Werden verschiedene Datenreihen mit unterschiedlichen Einheiten eingefügt, so kann die Skalierung der jeweiligen Y-Achse zu sehr unterschiedlichen Eindrücken führen.

Die Abszisse sollte in jedem Fall durchgehend skaliert sein und keinen Wechsel etwa zwischen linear und logarithmisch beinhalten.

Signifikante Unterschiede werden mittels **Sternchen** (und eckiger Klammer) dargestellt.

5.1.2.6. Anmerkungen

Attraktiv ist die Darstellung von Differenzen als Abweichung von einer horizontalen Vergleichslinie statt zwei Messreihen, da diese übersichtlicher und eingängiger sind. Bei geringen Unterschieden sollten generell Punktdiagramme herangezogen werden.

Zur Wahl der Grafiken (vgl. 3.2.1):

Grundtypen	Signalwörter	Beispiel
Struktur-Vergleich er zeigt den Anteil an einer Gesamtheit	Anteil, Prozentsatz	„Bundesbürger/-innen unter 16 Jahren verbringen 20% ihrer Freizeit vor dem Fernseher.“
Rangfolge-Vergleich er stellt Objekte bewertend gegenüber	größer, kleiner, besser, schlechter usw.	„Bei Wahlen zum Vorsitz erzielten alle Kandidaten etwa die gleiche Stimmenzahl.“
Zeitreihen-Vergleich er zeigt Veränderungen über die Zeit	Rückgang, Schwankung, Stagnation usw.	„Die Arbeitslosigkeit wird in den nächsten zwei Jahren zunehmen.“
Häufigkeits-Vergleich er zeigt die Auftretenshäufigkeit eines Objekts in verschiedenen Größenklassen	Bereich, Häufigkeit, Konzentration, Verteilung usw.	„Die meisten Menschen mit niedrigem Schulabschluss lesen Boulevard-Zeitungen.“
Korrelations-Vergleich er zeigt den Zusammenhang zwischen Variablen	relativ zu ..., steigt mit ..., verändert sich parallel zu ..., variiert entsprechend ... usw.	„Billigere Eintrittspreise bei Fußballspielen bedeuten nicht unbedingt eine Zunahme der Zuschauerzahlen.“

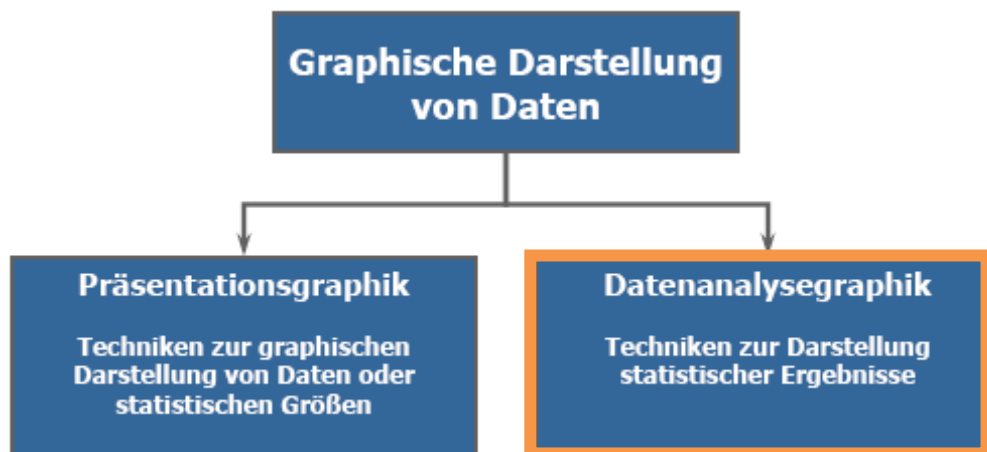
	Struktur	Rangfolge	Zeitreihe	Häufigkeit	Korrelation
Kreis					
Balken					
Säule					
Kurve					

Diagramm-Arten und Vergleichsart

- Überlege, was die Hauptaussage der Grafik sein soll und entscheide dich dann für eine Art der Darstellung.
- Wähle Achsenskalierung und Beschriftung derart, dass diese Aussage dem Leser leicht verständlich ist.
- Vermeide eine Vermischung verschiedener Aussagen in einer Grafik.

5.2. Graphische Analyse

In diesem Kapitel soll der Einsatz von graphischen Darstellungen zur Analyse von Daten behandelt werden.



Zur Wiederholung: **Präsentationsgraphen** dienen der Darstellung von Ergebnissen. Sie sind u. a. abhängig von den statistischen Annahmen der dargestellten Statistiken abhängig. Zentral ist hierbei die Verständlichkeit der Graphen.

Datenanalysegraphen sind hingegen meist unabhängig von statistischen Annahmen, da die Beobachtungen selbst dargestellt werden. Als „Wegwerf-Plots“ werden sie i. d. R. nur vom Auswerter selbst angeschaut. Man sollte dabei im Hinterkopf behalten, dass jeder Plottyp nur einen Teil der Strukturaspekte der Daten hervorheben kann.

Vorgehen bei graphischen Datenanalysen:

- Identifikation von Einzelheiten bzw. Regelmäßigkeiten (Strukturen, Mustern) in den Daten.
- Auswahl eines adäquaten Modells zur Beschreibung (Verteilung).
- Anpassung des Modells.
- Beurteilung der Modellanpassung (Residualanalysen)

Residualanalysen gehen von einem einfachen additiven Modell bzgl. der Entstehung von Daten aus:

$$\text{Data} = \text{fit} + \text{residuals}$$

Alle modellkonformen Aspekte (fit) entsprechen dabei bedeutsamen Aspekten der Daten (sofern das Modell korrekt ist), die Residuals sollten also keine erkennbare Struktur aufweisen (reasonably irregular).

Durch die Residualanalyse werden im Modell unberücksichtigte Variablen oder andere Formen des Zusammenhangs aufgedeckt, sodass sich gegebenenfalls neu spezifizierte Modelle aus den Residuen ergeben.

Datenanalysegraphen können dabei auf verschiedenen Ebenen eingesetzt werden. Die **Datenkontrolle** bezieht sich auf die Untersuchung einzelner Variablen, es sind aber auch Datenanalysegraphen für den **Vergleich empirischer Verteilungen** und zur Prüfung von **Voraussetzungen** und **Verteilungen** möglich. Auch können Datenanalysegraphen zur **Kategorisierung** von Informationen dienen.

5.2.1. Datenkontrolle

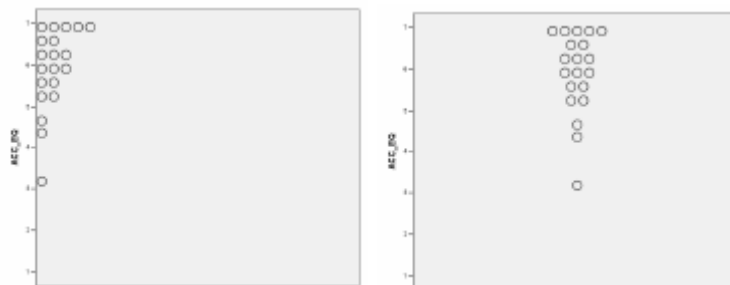
Bei der Untersuchung einzelner Variablen können verschiedene Fragen beantwortet werden. Bspw. werden **Ausreißer** aufgedeckt, genauso wie **Häufungen** von Beobachtungen in Teilen des Wertebereichs, **Fehlen** bestimmter Ausprägungen oder auch die **Verteilungsform** der Variable.

Eingesetzt werden hier vor allem Scatterplots (inkl. Q- und Q-Q-Plots), Box-Plots, Histogramme und Stem-and-Leaf-Plots.

5.2.1.1. Scatterplots (Dot-Plots)

Eindimensionale Scatterplots – die Darstellung der Ausprägungen der beobachteten Variablen an einem Zahlenstrahl – bietet sich vor allem bei $N < 100$ an. Allerdings wird hier die Anzahl der Beobachtungen mit identischen Ausprägungen nicht ersichtlich.

Alternativ bieten sich auch **Stacked Dot-Plots** (links) oder **Symmetrische Dot-Plots** (rechts) an, die bis ca. $N = 300$ sinnvoll verwendet werden können. Problematisch ist jedoch u. U. eine hohe Kurtosis (steilgipflige Verteilungen).



Bei einem **einfachen Scatterplot (Index-Plot)** wird die Kriteriumsvariable gegen die Abfolge der Fälle im Datensatz angetragen. Diese Darstellung ist besonders geeignet, um ungewöhnliche Beobachtungen / Probanden zu finden.

Q-Plots sortieren die Beobachtungen einer Variable nach deren Größe. Die Variablenausprägungen werden gegen den Anteil der Beobachtungen angetragen, die kleiner sind als sie selbst (also als Quantilwert). Bei derarti-

gen Darstellungen sind Extremwerte leicht erkennbar, sowie quantilbezogene Kennwerte (md, q1, q3) direkt aus der Grafik abzulesen. Lokale Dichten sind als starke Steigungen direkt erkennbar; identische Werte bilden senkrechte Bereiche im Plot.

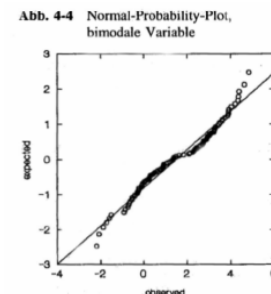
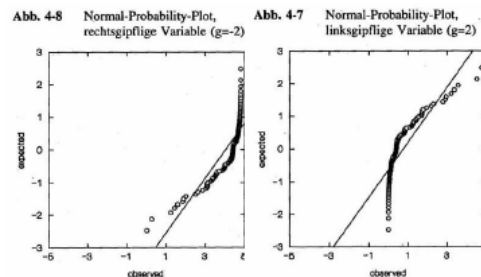
Theoretische Q-Q-Plots tragen die beobachteten Werte gegen erwartete Werte einer Verteilung an. Liegen die Plotpunkte eng auf der Ursprungsggeraden, so stimmen theoretische und empirische Verteilung überein.

Bilden Plotpunkte senkrechte Linien, so spricht das auch hier für hohe Häufigkeiten der einzelnen Werte.

Diese hohen Häufigkeiten sprechen dabei für **schiefe Verteilungen** (wenn die erwartete Verteilungsfunktion symmetrisch ist). Links oben ist eine

linksschiefe Verteilung, rechts oben ist eine rechtsschiefe Verteilung dargestellt.

Bei **bimodalen** Verteilungen der beobachteten Variablen zeigt sich im theoretischen Q-Q-Plot dagegen ein s-förmiges Abweichungsmuster.

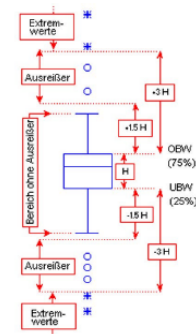


5.2.1.2. Boxplots

Boxplots erlauben Aussagen über Symmetrie und Schiefe der Verteilung sowie zu extremen Beobachtungen.

Die Fläche der Box entspricht im Normalfall 50% der Verteilung – eine kleine Fläche spricht also für eine hohe Konzentration der Werte. Die Lage des Medians gibt Auskunft über die Schiefe der Verteilung.

Bei der Erstellung eines Boxplots sollte jedoch immer darauf geachtet werden, dass die Einstellungen bzgl. Höhe der Box oder auch Definitionen von Extremwerten nicht normiert sind.



5.2.1.3. Histogramme

Histogramme sind die Standarddarstellung für Häufigkeitsverteilungen. Bei intervallskalierten Variablen ist jedoch gerade die Definition der Anzahl der Kategorien schwer (Regel von Sturges: $k = 1 + \sqrt{\ln n}$; alternativ: F19).

Attraktiv ist der Vergleich der Balkendarstellung mit dem Plot einer theoretischen Verteilung, die als Kurve in das Diagramm eingefügt werden kann.

5.2.1.4. Stem-and-Leaf-Plot

Im Stem-and-Leaf-Plot wird die Verteilung einer Variable über die Länge von Zeilen dargestellt. Die darzustellenden Zahlen werden eingeteilt in füh-

rende Ziffern (stem) und restliche Ziffern (leaf). Stem muss dabei Ziffern darstellen, die nicht im Datensatz vorkommen.

```
ACC_EQ Stem-and-Leaf Plot

Frequency      Stem & Leaf

      1.00  Extremes      (<=3.2)
      1.00      4 . 3
      1.00      4 . 6
      2.00      5 . 13
      4.00      5 . 5688
      4.00      6 . 0133
      3.00      6 . 568
      4.00      7 . 0000

Stem width:      1.00
Each leaf:      1 case(s)
```

5.2.2. Vergleich empirischer Verteilungen

Die einfachste Art des Vergleichs empirischer Verteilungen ist der Vergleich von Deskriptiva, etwa über Fehlerbalkendiagramm, Liniendiagramm oder Balkendiagramm (jeweils mit Fehlerindikatoren).

Beim Vergleich von Gruppen sind vor allem **Boxplots** attraktiv, da hier Unterschiede in Verteilungen und Lagen sichtbar werden und Ausreißer identifiziert werden können.

Q-Q-Plots können auch hier eingesetzt werden, indem die Quantile zweier Verteilungen verglichen werden. Diese Art der Darstellung ist für die Regressionsanalyse verwendbar. [Unter SPSS lassen sich derartige Q-Q-Plots unter Streu-/Scatterdiagramm ausgeben.]. Auch hier gilt: Liegen die Plotpunkte auf der Linie $X = Y$ so stimmen die Verteilungen überein.

Falls insbesondere die Abweichungen von der Regressionsgeraden interessieren, so können auch **Summen-Differenzplots** verwendet werden. Auf die Verwendung von Scatterplots soll im Folgenden noch genauer eingegangen werden.

5.2.2.1. Scatterplots

Die Streuung des Scatterplots gibt Auskunft über die Enge des Zusammenhangs zwischen den angetragenen Variablen. Dabei wird der Eindruck stark durch die Achsenskalierung beeinflusst. Regelhaft ist der beurteilte Zusammenhang ca. um 40% geringer als der tatsächliche Zusammenhang.

Zur Erleichterung der Interpretation eines Scatterplots (gerade bei nichtlinearen Zusammenhängen) kann dieser mit Informationen angereichert werden. Die einfachste Variante besteht darin, eine Linie in Form des erwarteten Zusammenhangs einzuzeichnen.

Zusätzlich können die Abweichungen von dieser Linie als Fehlerbalken visualisiert werden oder weitere Zusammenhänge eingefügt werden, um größere Bereiche abzudecken.

Ein Problem bei der Verwendung von Hilfslinien ist, dass die Form des Zusammenhangs von der Wahl des Modells abhängt. Daher sollte zunächst ein **Smoother** verwendet werden, um das bestmögliche Modell zu spezifizieren.

Ein Beispiel ist der LOWESS-Smoother von Cleveland (1979; Locally-weighted scatterplot smoother). Dabei wird angenommen, dass die Bedeutung eines Punktes mit dem Abstand zum Berechnungspunkt abnimmt; Ausreißer werden mit diesem Verfahren also nur gering gewichtet. Je größer der Glättungsparameter f , desto glatter wird die Funktion (i. d. R.: $0.5 < f < 0.8$).

Anmerkung 1: Bei großen Stichproben bedeutet ein einzelner Punkt möglicherweise die Zusammenfassung mehrerer Beobachtungswerte. Informationen gehen also bei großen Stichproben zwangsweise verloren.

Anmerkung 2: Erfasst eine Variable eigentlich zwei verschiedene, unabhängig bivariat normalverteilte Aspekte (es liegt also eine bivariate Normalverteilung vor), so gibt ein Scatterplot wenig Auskunft über die Form des Zusammenhangs. Eine 3D-Darstellung kann hier abhelfen.

5.2.2.2. Weitere Darstellungsformen

- 2D-Matrixplot
- 3D-Scatterplot
- 3D-Iconplots

Werden hoffentlich noch in der Vorlesung genauer behandelt.

5.2.3. Kategorisierung von Informationen

Neben der Datenkontrolle sowie dem Vergleich empirischer Verteilungen kann die graphische Datenanalyse auch zur Kategorisierung von Informationen eingesetzt werden.

Ein Beispiel ist die Verwendung von kategorialen Kontrollvariablen und die Erstellung von getrennten Diagrammen/Plots auf Basis der Ausprägung dieser Variablen. Unterschiede in Abhängigkeit von Kontrollvariablen werden auf diese Weise u. U. leichter interpretierbar als deskriptive Kennwerte.

Mögliche Darstellungen sind:

- [Mehrfachkodierte Scatterplots](#) (unterschiedliche Farben/Zeichen)
- [Kategorisierte Scatterplots](#)
- [Kategorisierte Boxplots](#) (in Zeilen oder Spalten)

5.2.4. Prüfung von Voraussetzungen / Verteilungen I: Varianzanalyse

Nach Datenkontrolle, Vergleich von empirischen Verteilungen und Kategorisierung sollen abschließend graphische Verfahren zur Prüfung bzw. Abschätzung von Voraussetzungen für Varianz- und Regressionsanalyse behandelt werden.

Die Voraussetzungen (Annahmen) der Varianzanalyse sind:

- [Normalverteilungsannahme](#): Residuen / Fehler in jeder Zelle sind normalverteilt (Kolmogoroff-Smirnov-Test)
- [Homogenitätsannahme](#): Varianzen der Residuen/Fehler sind innerhalb der Zellen gleich (Levene-Test)
- [Unabhängigkeitsannahme](#): Residuen/Fehler sind innerhalb der Zellen unabhängig voneinander (F-Test von Box).

5.2.4.1. Normalverteilungsannahme

Eine Abweichung eines Messwerts vom jeweiligen Stichprobenmittelwert (Zellenmittelwert) wird als Fehler definiert. Dabei sollten die Fehler innerhalb einer Treatmentstufe normalverteilt sein, was sich über den Kolmogoroff-Smirnov-Test (X^2 ; H_0 : Normalverteilung) prüfen lässt. Die Abweichung von der Normalverteilung ist bei schiefen Verteilungen zu vernachlässigen, genauso wie bei schmalgipfligen Verteilungen (konservativ). Kritisch sind lediglich breitgipflige Verteilungen ($\alpha_{\text{real}} > \alpha_{\text{nominell}}$).

Normal Q-Q-Plots tragen die beobachteten Werte gegeneinander an. Dabei wird eine Normalverteilung der Werte angenommen. Liegen die empirischen Werte annähernd auf der Hilfslinie, so kann von einer Normalverteilung der Werte ausgegangen werden.

Ein ähnliches Vorgehen liegt der Prüfung der **Normalverteilung der Residuen** zugrunde. Hier wird die Abweichung der Residuen von der Normalverteilung angetragen. Um einen derartigen Plot zu erstellen, müssen die Residuen zuvor mit SPSS geschätzt werden (Save-Button).

Schließlich können mit Boxplots auch **Ausreißer** gefunden werden, die in der Prüfung der Normalverteilungsannahme nicht enthalten sind, das Ergebnis der ANOVA jedoch stark beeinflussen können.

5.2.4.2. Varianzhomogenität der Residuen

Die Varianzen der Messwerte pro Treatmentstufe sollen homogen sein. Der entsprechende Levene-Test geht bzgl. seiner Nullhypothese von einer Gleichheit der Varianz über alle Gruppen hinweg aus. Dies ist besonders dann kritisch, wenn die Stichproben ungleich groß sind.

Eine graphische Prüfung der Varianzhomogenität sind **Spread-versus-Level-Plots** (mit natürlichem Logarithmus). Die zugrundeliegende Annahme ist, dass die Streuung der Messwerte unabhängig von deren Median sein sollte [Warum das?].

5.2.4.3. Unabhängigkeit der Residuen

Der Fehlerterm eines Messwerts muss unabhängig von den Störeinflüssen auf andere Messwerte sein. Diese Annahme ist u. a. bei Messwiederholungen verletzt.

Lösung: Test auf Homogenität der Varianz-Kovarianz-Matrix (F-Test von Box).

Reihenfolgeeffekte können über einen **Residuen-Index-Plot** analysiert werden. Geschickt ist hier jedoch vielmehr der **Residual*Predicted-Plot**, bei dem die Residuen gegen die vorhergesagten Werte angetragen werden. Falls ein linearer Zusammenhang auftritt liegt eine Verletzung der Homoskedastizität vor.

5.2.4.4. Diagnostische Plots

Profilplots können für den schnellen Vergleich von Mittelwerten verwendet werden und finden sich in SPSS unter dem Menüpunkt Plots der Analyse-methode (also schnell einstellbar).

Untersuchbare Hypothesen:

- **Niveauhypothese:** Gibt es gleichmäßige Unterschiede zwischen den Gruppen?
- **Parallelitätshypothese:** Laufen die Profile der Gruppen parallel?
- **Flatness-Hypothese:** Sind die Profile Flach oder gibt es Unterschiede zwischen den Variablen?

Reine **Mittelwertplots** (mit Standardabweichungen) können verwendet werden, um sich schnell ein Bild von den Daten zu machen. Signifikanzen sind hier jedoch nicht abschätzbar, weswegen anstelle der SD besser ein KI angetragen werden sollte.

Interaktionsplots schließlich erlauben Aussagen über Art einer signifikanten Interaktion zwischen den Faktoren. Eine Interaktion ist immer dann gegeben, wenn bei einem Plot die Linien nicht parallel verlaufen.

Dabei wird zwischen ordinaler, hybrider (semi-disordinaler) und disordinaler Interaktion unterschieden.

5.2.5. Prüfung von Voraussetzungen / Verteilungen II: Regressionsanalyse

Die Bedeutung von diagnostischen Plots im Kontext der Regressionsanalyse wird an den Beispieldatensätzen von Anscombe (1973; F63+) deutlich. Die zu testenden Voraussetzungen der Regressionsanalyse sind in der folgenden Tabelle dargestellt.

Voraussetzungen	Folgen bei Verletzung
1. Vollständigkeit des Modells	Modell nicht korrekt spezifiziert.
2. Keine lineare Abhängigkeit zwischen Regressoren	Multikollinearität zwischen den Regressoren.
3. Lineare Beziehung zwischen Regressor(en) und Regressand	Nichtlinearität zwischen Regressor(en) und Regressand.
4. Normalverteilung der Störgrößen bzw. Residuen	Keine Normalverteilung der Residuen.
5. Homoskedastizität der Residuen	Heteroskedastizität der Residuen.
6. Unabhängigkeit der Residuen	Autokorrelation der Residuen.

5.2.5.1. Vollständigkeit des Modells

Bei einem nicht optimalen Modell wird die Schätzung ineffizient, wobei zwischen den beiden Fällen **overfitting** (zu viele Prädiktoren) und **underfitting** (zu wenige Prädiktoren) unterschieden wird.

Diese Voraussetzung lässt sich über den Scatterplot der beobachteten Werte gegen vorhergesagte Werte prüfen. Passt das Modell nur in einem bestimmten Datenbereich, so ist es eventuell angebracht, die Daten zu Transformieren.

5.2.5.2. Lineare Abhängigkeit zwischen den Regressoren

Ein Scatterplot gibt auch hier Auskunft über die Korrelation zwischen den Regressoren. Bei annähernder Nullkorrelation ist die Voraussetzung als erfüllt anzusehen.

5.2.5.3. Lineare Abhängigkeit zwischen UV und AV

Bei Regressionen mit bis zu 10 Variablen sind bivariate Zusammenhänge noch gut über Scatterplots prüfbar. [In diesem Fall sollen sie vorliegen]. Liegen nichtlineare Zusammenhänge vor, müssen die Daten transformiert werden.

5.2.5.4. Normalverteilung der Residuen

Diese Voraussetzung ist über einen [Normal Q-Q-Plot mit „Studentized Residuals“](#) prüfbar, dabei werden die Residualwerte an der SD standardisiert, bei deren Berechnung sie selbst ausgeschlossen wurden. Werte > 2 werden als groß betrachtet.

In diesem Plot werde [Verteilungsabweichungen](#), [Klumpungen](#) und [Ausreißer](#) leicht erkenntlich.

5.2.5.5. Homoskedastizität

Um zu prüfen, ob die Fehlervarianz konstant ist [bedeutsam vor allem für die Inferenzstatistik], kann ebenfalls ein [Scatterplot mit Studentized Residuals](#) [absolut] verwendet werden, die [gegen die vorhergesagten Werte der Regressanden](#) angetragen werden.

Bei Homoskedastizität zeigt sich hierbei eine horizontale Linie. Falls nicht ist diese Voraussetzung nicht erfüllt.

5.2.5.6. Autokorrelation der Residuen

Die Residuen sollten eigentlich unkorreliert sein. Ist dies nicht der Fall, so spricht man von einer sog. Autokorrelation. [Vorlesung?]

5.3. Inferenzstatistische Analyse

5.3.1. Grundlagen der Inferenzstatistik

Begriffssammelsurium:

- Null- und Alternativhypothese
- H_0 liefert die Verteilung, H_1 die Summationsrichtung; gerichtete vs. ungerichtete, spezifische vs. unspezifische H_1
- Alpha- und Beta-Fehler
- Konvention für $\alpha = 0.05$ statt statistischer Begründung
- Alpha-Inflation und Alpha-Adjustierung
- Minimierung des Beta-Fehlers durch Alpha-Vergrößerung, Vergrößerung des Stichprobenumfangs oder versuchsplanerische und statistische Kontrolltechniken (z.B. Parallelisierung / Kovarianzanalyse).
- Warum so viele Tests? → Für jeden Versuchsplan, jedes Datenniveau und jede Art von Fragestellung ein angemessenes Verfahren.

- Bestandteile eines Hypothesentests: **Statistik** (Verteilung, df...), **Handlungsalternativen** (H_0 und H_1) und **Entscheidungsregel**.
- Hypothesentests als Wahrscheinlichkeitsaussagen (kein logisches Schließen)

5.3.2. Entscheidungsfehler und Teststärke

In einer Studie von Rosenthal und Gaito (1963) sollte $N = 19$ erfahrene Forscher empirische Studien nach ihrer Vertrauenswürdigkeit bewerten. Es zeigte sich, dass das Vertrauen umso größer war, je größer der Stichprobenumfang und je kleiner das α -Niveau. Das Vertrauen wurde vor allem dann stark gemindert, wenn ein α -Niveau von 0.05 nicht erreicht wurde (Cliff-Effekt).

Die Vertrauenswürdigkeit (als Synonym für Replizierbarkeit) wird also vor allem aufgrund des α -Niveaus (Signifikanzniveaus) beurteilt. Das korrekte Maß für die Vertrauenswürdigkeit ist jedoch nicht die Signifikanz, sondern die **Teststärke** ($1-\beta$), als Wahrscheinlichkeit einen Effekt zu finden, unter der Voraussetzung, dass er auch wirklich vorliegt [Anmerkung: Hatten die Forscher hierzu überhaupt Angaben? Wohl eher nicht...].

Die Teststärke wiederum hängt v. a. von folgenden Faktoren ab:

- **Effekt des Treatments**: Hohe Primärvarianz \rightarrow Effekt leicht zu zeigen
- **Variabilität des Merkmals**: Hohe Fehlervarianz \rightarrow Geringe Teststärke
- **Höhe des α -Niveaus**: Je schärfer, desto geringer Teststärke (α - und β -Fehler verhalten sich gegenläufig!)
- **Einseitige Tests** haben eine höhere Teststärke als zweiseitige
- **Stichprobengröße**: Je größer N desto höher die Teststärke

Die **Teststärke-/Gütefunktion** gibt dabei die Wahrscheinlichkeit an, eine Nullhypothese korrekterweise abzulehnen: Die Teststärke steigt mit wachsender wahrer Differenz zwischen KG und EG an.

5.3.3. Bedeutung des Stichprobenumfangs

Die Stichprobengröße wirkt sich auf die Teststärke aus, weswegen es eigentlich wünschenswert wäre immer sehr große Stpn zu untersuchen. Allerdings sprechen sowohl ökonomische Argumente (Aufwand und Kosten) dagegen, als auch die Tatsache, dass in der Praxis meist nach Mindesteffekten gesucht wird, sodass keine wirklich großen Stichproben benötigt werden.

Es gilt also, eine **optimale Stichprobengröße** für die jeweilige Studie zu ermitteln. Ein Stichprobenumfang gilt dabei dann als optimal, wenn er bei Festlegung von α , β und ε eine eindeutige Entscheidung hinsichtlich der Gültigkeit von H_0 vs. H_1 ermöglicht. Dabei gilt:

$$N = \frac{(z_{1-\alpha} - z_{\beta})^2}{\varepsilon^2} \quad \text{mit} \quad \varepsilon = \frac{(m_1 - m_2)}{sd}$$

Problematisch ist die Bestimmung der **β -Fehlerwahrscheinlichkeit**, da dies nur bei einer spezifischen H_1 möglich ist. Zur groben Abschätzung empfiehlt Cohen (1988) eine Vervierfachung der α -Fehlerwahrscheinlichkeit.

Auch die Bestimmung der **Effektgröße** ist nicht immer einfach. Bestenfalls liegt eine Metaanalyse vor, auf deren Ergebnisse zurückgegriffen werden kann. Ist

dies nicht der Fall, sollte eine Pilotstudie zur Abschätzung der Effektstärke durchgeführt werden bzw. bestimmt werden, welche minimale Effektstärke erforderlich ist.

Kleiner Effekt: $d = 0.2$ Mittlerer Effekt: $d = 0.5$ Starker Effekt: $d = .08$

Wichtig: Geringe Teststärke ist kein sicherer Hinweis auf Gültigkeit der H_0 !
Anmerkung: Für Korrelationen hängt insbesondere die Signifikanz deutlich von der Stichprobengröße ab (→ Praktische Bedeutsamkeit?). Ein Einflussfaktor auf die Höhe der Korrelation sind Reliabilität und Validität der verwendeten Verfahren. Daher sollte eine korrigierte Korrelation verwendet werden:

$$r_{\text{korr}} = \frac{r_{xy}}{\sqrt{REL_x \cdot REL_y}}$$

5.3.4. Praktische Bedeutsamkeit

Wie oben kurz erwähnt, wirkt sich die Stichprobengröße nicht nur auf die Teststärke eines Verfahrens aus, sondern auch auf die Signifikanz eines Ergebnisses. Statistische Signifikanz kann also nicht als alleiniges Kriterium zur Abschätzung der Bedeutsamkeit von Befunden herangezogen werden.

Zur Abschätzung der praktischen Bedeutsamkeit bietet sich die Berechnung eines Effektstärkemaßes an, etwa im Sinne eines **Maßes für den Anteil der aufgeklärten Varianz** oder ein **Klassifikationsmaß** bzgl. des prozentualen Zutreffens der getroffenen Klassifikationen.

5.4. „Reading Statistics“

Dieses letzte Kapitel beschäftigt sich mit der Interpretation deskriptiver, inferenzstatistischer und graphischer Informationen.

5.4.1. Interpretation deskriptiver Informationen

In diesem Abschnitt werden univariate und bivariate Darstellungen (u.a. Korrelationen) behandelt. Zudem folgt ein Ausblick zur Parameterschätzung (Vertrauensintervalle, Standardfehler).

Anmerkung: Im Text eines Artikels sind statistische Begriffe stets auszusprechen; also: „Der Mittelwert beträgt...“.

5.4.1.1. Univariate Darstellungen

Beispiele für die Verwendung von deskriptiven Kenngrößen finden sich auf den Folien 6-29.

Die am häufigsten angegebenen deskriptiven Kenngrößen sind selbstverständlich **Maße der zentralen Tendenz**. Zwar sind Mittelwert, Median und Modus formal völlig verschiedene Konzepte, im Flusstext werden sie jedoch gerne identisch bezeichnet! Bsp.: „Im Mittel“, „durchschnittlich“ oder „mittlerer Wert“. → Aufpassen!

Eine weitere Dissoziation zwischen formaler Bedeutung und sprachlicher Verwendung findet sich bei **Quartilen**. Formal sind Quartile *Punkte*, die die Stichprobe in 4 gleich große Gruppen teilt. Sprachlich wird der Begriff Quartil für Probanden*mengen* zu 25% der untersuchten Fälle verwendet.

Zur Interpretation des **Standardfehlers**: Der SE ist ein Streuungsparameter für Mittelwerte. Er kann auch als Indikator für Stichprobenfehler verwendet werden, da in 68% der Fälle der Stichprobenfehler kleiner ist als der SE.

Für **Vertrauensintervalle** siehe 5.4.1.3 (Schätzung).

5.4.1.2. Bivariate Darstellungen

Interkorrelationen zwischen verschiedenen Variablen lassen sich geschickterweise in Tabellen darstellen. Bei Korrelationen insgesamt immer zu beachten:

- Welcher Korrelationskoeffizient?
- Lineare vs. kurvenlineare Zusammenhänge?
- Einfluss von Ausreißern
- Interpretation als Anteil aufgeklärter Varianz
- Nomenklatur bei Interpretation
- Und natürlich: Korrelation und Kausalität

Zur Nomenklatur von Korrelationskoeffizienten existieren Konventionen, die sich aus den Angaben zur Interpretation der Effektstärke ableiten (jede Effektstärke und jede andere Stichprobenstatistik lässt sich in eine Korrelation umrechnen).

$R^2 = .01$	bzw. $r = .1$	Kleiner Effekt
$R^2 = .09$	bzw. $r = .3$	Mittlerer Effekt
$R^2 = .25$	bzw. $r = .5$	Großer Effekt

Problematisch an dieser Konvention: Ein kleiner „Effekt“ klärt gerade einmal 1% der Varianz auf...das ist eigentlich nichts.

5.4.1.3. Schätzung

Für eine Betrachtung des **Standardfehlers** siehe 5.4.1.1 (univariate Darstellungen).

Vertrauensintervalle werden laut Ingo auf zwei Arten interpretiert: a) Wertebereich, in dem der Populationsparameter mit der Wahrscheinlichkeit p liegt oder b) Wertebereich, in dem ein empirisch ermittelter Kennwert mit Wahrscheinlichkeit p liegen wird.

Aus QMB: Das Konfidenzintervall ist ein Bereich um eine Stichprobenstatistik, in dem die Populationsparameter liegen, die als Erzeuger der Statistik plausibel sind. Richtige Interpretation: „Wenn viele derartige Stichproben gezogen werden und KIs berechnet werden, liegt der wahre Mittelwert in 95% der Fälle innerhalb des KIs, in 5% außerhalb.“ Eine Wahrscheinlichkeitsaussage ist demnach nur für die Gesamtheit aller möglichen KIs möglich.

Also: „(...) probability statements about individual Cis can so easily be misinterpreted that they are best avoided.“ (Cummings & Finch, 2005). Mögliche Interpretationen für KIs finden sich auf F27 der QMB-Präsentation.

5.4.2. Interpretation inferenzstatistischer Informationen I: Inferenzstatistische Grundprinzipien

WDH: Bevor inferenzstatistische Größen mitgeteilt werden müssen die dazugehörigen deskriptiven Kennwerte im Text mitgeteilt werden. Die Angabe inferenzstatistischer Größen erfolgt standardisiert nach dem Schema $\text{Symbol}(df) = p\text{-Wert}$ ohne führende Null. Bei X^2 -Werten wird zusätzlich die Stichprobengröße angegeben.

Dabei sollten **Signifikanzniveaus** immer als exakte Werte angegeben werden. Wichtig: Ein Alpha-Niveau ist keine Alles-oder-Nichts-Schranke. Wird das festgelegte Alpha-Niveau knapp nicht erreicht, spricht man von:

- **marginal significance**
- **borderline significance**
- ... **approached significance** oder
- **trend toward significance**

Ist ein Ergebnis nicht auf dem 5%-Niveau signifikant sollte es also auch nicht einfach weggeschmissen werden!

Anmerkung: Gelegentlich findet man Angaben wie $p = .00$ oder $p < .000$. Das ist natürlich schwachsinn...

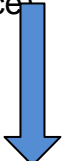
Abschließend sei noch auf den uneindeutigen Gebrauch des Wortes Hypothese hingewiesen, das nach Lust und Laune für H_0 und H_1 gleichermaßen verwendet wird.

5.4.3. Interpretation inferenzstatistischer Informationen II: Unterschiedshypothesen

Geschickt ist die Darstellung von sowohl Deskriptiva als auch deren Signifikanz in einer Tabelle (vgl. F42).

Zur ANOVA wird jede Menge gesagt...hier sei aber auf Methoden der Unterschiedsprüfung verwiesen. Immerhin stehen auf den Folien einige Beispiele zur Angabe der Kennwerte im Text einer Publikation.

Interessant ist jedoch die Angabe der Strenge verschiedener post-hoc Tests. Von liberal (progressiv) nach konservativ lässt sich folgende Liste aufstellen:

- Fisher's LSD (least significant difference)
 - Duncan's Multiple Range Test
 - Neuman-Keuls
 - Tukey
 - Scheffé
- 
- liberal
- konservativ

5.4.4. Interpretation inferenzstatistischer Informationen III: Zusammenhangshypothesen

Die Nullhypothese (engl. null hypothesis) ist bei praktisch allen Tests auf **Korrelationen**: „Die Variablen der Grundgesamtheit sind unkorreliert.“, wenn nichts anderes angegeben wird. Eine relativ häufige Alternative ist der Test, ob sich zwei Korrelationskoeffizienten signifikant unterscheiden.

Die Angabe einer Prüfgröße für Korrelationen ist in den meisten Fällen ein t-Wert. Hier ist es wichtig darauf zu achten, welcher Korrelationskoeffizient geprüft wurde. Speziell bei Korrelationen ist desweiteren der Einfluss der Stichprobengröße zu beachten. Häufig wird dabei auf den p-Wert fokussiert, obwohl eigentlich im r-Wert (R^2) mehr Informationen enthalten sind.

Bei einfacher und multipler Regression werden die einzelnen Beta-Gewichte und in Klammern der jeweilige Prädiktor angegeben. Ein Beispiel zur multiplen Regression:

„The regression equation, computed using all of the WGCTA subscales, was $Y = 2.3 + .03$ (Inference) + $.003$ (Recognition of Assumptions) + $.03$ (Deduction) + 0.6 (Interpretation) - 0.3 (Evaluation of Arguments).

In der Regressionsanalyse werden die einzelnen Prädiktoren genannt und in Klammern (beta = $.52$).

Nach Möglichkeit sollte auch immer der Determinationskoeffizient (coefficient of determination; R^2) angegeben werden. Beispiele auf F76 und F77.

5.4.5. Interpretation graphischer Informationen

Siehe 3.2 (Visualisierung).