

Lösungen zu den Übungsaufgaben aus dem Lehrbuch „Psychologische Methodenlehre / Statistik“ von Rainer Leonhart

Kapitel 1

1. Was bedeutet es, dass die Psychologie eine empirische Wissenschaft ist?

Die Psychologie baut auf Erkenntnissen auf, die durch verschiedene Formen der Datenerhebung gewonnen wurden. So werden beispielsweise durch Experimente theoriegeleitete definierte Hypothese bestätigt (oder verworfen) und die auf dieser empirischen Grundlage gewonnenen Erkenntnisse gehen als Fakten in die Psychologie ein. Hierbei müssen im Allgemeinen empirische Methoden verwendet werden. Die psychologische Methodenlehre umfasst eine Vielzahl von Werkzeugen zum Erkenntnisgewinn und ist somit die Grundlage für alle anderen Disziplinen der Psychologie (z. B. Sozialpsychologie, Allgemeine Psychologie).

2. Wie wird beim sog. wissenschaftlichen Zirkel vorgegangen?

Der sog. wissenschaftliche Zirkel veranschaulicht den zirkulären Prozess des Erkenntnisgewinns in einem Prozess in mehreren (vielen) Schleifen, der dem Erkenntnisgewinnen dient. Wissenschaftliches Arbeiten durchläuft hierbei (immer wieder) die folgenden Stufen:

- Aufnahme und Zusammenstellung der relevanten Fakten zu einem Thema bzw. einer Fragestellung,
- Definition einer in der Empirie überprüfbaren Theorie,
- Ableitung von Folgerungen (Hypothesen) aus dieser Theorie,
- Überprüfung der Folgerungen an der Realität (Hypothesenprüfung), beispielsweise durch ein Experiment,
- Ergebnis der Hypothesenprüfung als neue Fakten zur Kenntnis nehmen,
- zurück zum Anfang des Zirkels, da nun neue Erkenntnisse die Datenbasis, die Faktenlage, verändert haben.

3. Was besagt das Induktionsproblem?

Bei der Induktion werden über viele Einzelbeobachtungen hinweg Informationen gesammelt, anhand derer anschließend eine Allgemeinaussage abgeleitet wird. Problematisch an diesem Vorgehen ist, dass die so aufgestellten Regeln nie endgültig beweisbar sind, da dazu eigentlich alle möglichen Fälle beobachtet werden müssten.

4. Nennen Sie die drei Prinzipien des Positivismus.

Die drei Prinzipien des Positivismus sind:

- Philosophie muss als exakte Wissenschaft betrieben werden (d.h. es wird ein naturwissenschaftliches, experimentelles Vorgehen vorgeschlagen);
- wissenschaftliche Erkenntnis wird durch Logik gewonnen (d.h. Sätze und Regeln müssen mit den Grundlagen der Logik bearbeitet werden können);
- metaphysische Themen werden ausgeblendet (d. h. nur beobachtbare und objektiv erfassbare Eigenschaften werden untersucht bzw. sind Gegenstand der Wissenschaft).

5. Was ist ein Protokollsatz?

Protokollsätze sind alle Sätze, die durch unmittelbare Wahrnehmung zustande kommen oder bestätigt werden können. Ein Beispiel für einen Protokollsatz wäre: „Das Kind hat mit dem roten Ball gespielt.“

6. Diskutieren Sie vor dem Hintergrund der Hypothesen von Ludwik Fleck die möglichen Probleme der Vermittlung von Methodenkenntnissen in der Psychologie.

Nach Ludwik Fleck wird auch die Methodenausbildung durch den vom „Denkkollektiv“ erzeugten, charakteristischen „Denkstil“ geprägt. Somit wird in der methodischen Ausbildung nur ein bestimmter Teil aller möglichen Methoden gelehrt bzw. vertieft. Dies lenkt das wissenschaftliche Vorgehen und das wissenschaftliche Denken der „Wissenschaftlergemeinschaft“ in eine bestimmte Richtung.

7. Durch was wird nach Lakatos der Kern jeder Theorie geschützt?

Durch einen Gürtel aus modifizierbaren Schutzhypothesen, welche die Kernhypothese unangreifbar machen.

8. Was ist bei einem vollstandardisierten Interview vorgegeben?

Bei einem vollstandardisierten Interview werden die exakten (schriftlich fixierten) Fragen und Antwortmöglichkeiten vorgegeben sowie ein für alle Personen identischer Ablauf der Befragung. Ebenfalls ist auch der einführende Text für die Befragenden (Hinführung zum Interview) festgelegt. Darüber hinaus ist der Ablauf der Befragung für alle Personen identisch, was mit einschließt, dass der einleitende Text für die Befragenden (Hinführung zum Interview) festgelegt ist.

9. Was muss bei einem Experiment gegeben sein?

Bei einem Experiment werden Daten unter kontrollierten Untersuchungsbedingungen erhoben. Der Untersuchungsleiter kann hierbei Zeitpunkt und Ausprägung der unabhängigen Variablen nach dem Versuchsplan variieren (selbstverständlich nur innerhalb der ethischen Grenzen). Gemessen wird der Effekt dieser Manipulation (Veränderung) auf die abhängige Variable. Im Allgemeinen sind Experimente immer replizierbar.

10. Wie ist der Effekt eines Treatments definiert?

Der Effekt eines Treatment wird erfasst über die Differenz in der abhängigen Variablen zwischen den Werten der Probanden, die einem Treatment ausgesetzt sind, verglichen mit den Ergebnissen von Probanden, die keinem Treatment ausgesetzt sind.

11. Wann kann von Kausalität ausgegangen werden?

Nach John Stuart Mill kann von Kausalität ausgegangen werden, wenn:

- dem Effekt die Ursache zeitlich vorausgeht,
- Ursache und Effekt in einer Beziehung zueinander stehen (Korrelation),
- es keine alternative Erklärung für den gefundenen Effekt gibt (es darf nicht zwei oder mehr alternative, nicht widerlegbare Erklärungsansätze geben).

12. Was unterscheidet ein Experiment und ein Quasi-Experiment?

Die Zuordnung von Probanden zu Untersuchungsbedingungen ist bei Quasi-Experimenten im Gegensatz zu Experimenten nicht zufällig.

13. Welche Formen der Validität gibt es?

- Validität statistischer Schlussfolgerungen
- Interne Validität
- Externe Validität
- Konstruktvalidität (Inhaltsvalidität, Kriteriumsvalidität)

Kapitel 2

1. Nennen Sie jeweils zwei psychologische Beispiele für latente und manifeste Merkmale.

Latente Merkmale: Studieneignung, Hochbegabung, Ängstlichkeit.

Manifeste Merkmale: Körpergröße, Augenfarbe, Geschlecht, Reaktionszeit.

2. Warum ist die Definition des Skalenniveaus vor dem Beginn der Datenerhebung relevant und wie kann bei der statistischen Auswertung mit dem Problem eines zu geringen Niveaus umgegangen werden?

a) Beispielsweise wird bei der Entwicklung eines Fragebogens mit der Anzahl der vorgegebenen Kategorien das mögliche Skalenniveau schon eingeschränkt. So kann bei nur zwei Antwortmöglichkeiten (stimme zu – stimme nicht zu) nicht von einem Merkmal auf Intervallskalenniveau gesprochen werden. Das Skalenniveau spiegelt den Informationsgehalt erhobener Variablen wider und bestimmt dadurch, wie differenziert eine Fragestellung untersucht bzw. beantwortet werden kann. Daher ist es sinnvoll, Merkmale so in Variablen zu überführen, dass das maximal mögliche Skalenniveau erreicht wird. Mit der Festlegung vor einer Untersuchung, wie Merkmale gemessen werden, wird also auch das Skalenniveau bestimmt.

b) Wenn die Voraussetzung für ein höheres (parametrisches) Verfahren nicht gegeben sind, muss ein testschwächeres (nonparametrisches) Verfahren eingesetzt werden. Die Anwendung bestimmter statistischer Kennwerte (z.B. Berechnung des Mittelwertes) wäre bei diesem Beispiel nicht möglich.

3. Gegeben sei die folgende Reihe von Messwerten: 2, 3, 4, 4, 4, 5, 5, 6, 6, 7

Erstellen Sie eine Tabelle der Häufigkeitsverteilung. Bestimmen Sie Mittelwert, Median und Modalwert.

Merkmalsausprägung	Anzahl	Prozent	Kumulative Prozent
2	1	10,0	10,0
3	1	10,0	20,0
4	3	30,0	50,0
5	2	20,0	70,0
6	2	20,0	90,0
7	1	10,0	100,0
Total	10	100,0	

Mittelwert = 4.6

Median = 4.5

Modalwert = 4.0

4. Wie ist der Unterschied zwischen einer Normierung und einer Normalisierung definiert?

Bei einer Normierung wird mittels einer linearen Transformation ($y = a \cdot x + b$) eine vorliegende Verteilung in eine standardisierte Verteilung umgewandelt – meist durch z-Transformation (Mittelwert = 0 und Streuung = 1). Dadurch werden die Werte von Personen aus verschiedenen Stichproben oder Erhebungen mit verschiedenen Messinstrumenten vergleichbar. Im Gegensatz hierzu wird bei der Normalisierung eine nichtlineare Flächentransformation durchgeführt. Somit sind die ursprünglichen Verhältnisse der Abstände zwischen den Werten der Rohdaten nicht mehr mit den Abständen der Verhältnisse bei den transformierten Daten vergleichbar.

5. Worin unterscheiden sich die Wahrscheinlichkeitsbegriffe nach Laplace und nach Bernoulli? Wann wird welche Definition eingesetzt.

Bei der Wahrscheinlichkeit nach Laplace wird die Anzahl der möglichen Ausgänge in ein Verhältnis zur Menge der günstigen Ereignisse gesetzt, während bei der Ermittlung der Wahrscheinlichkeit nach Bernoulli durch die Verwendung des Limes (Grenzwert) eine Schätzung der wahren Wahrscheinlichkeit über die relative Häufigkeit stattfindet. Die Wahrscheinlichkeit nach Laplace wird verwendet, wenn vor der Durchführung eines Zufallsexperiments alle möglichen Ausgänge bekannt sind. Eine Schätzung der Wahrscheinlichkeit nach Bernoulli ist hingegen erst nach sehr vielen Durchgängen eines Zufallsexperiments möglich.

6. Worin unterscheidet sich eine Klumpenauswahl von einer mehrstufigen Zufallsauswahl?

Bei der Klumpenauswahl werden in den letzten Teilstichproben alle Elemente erhoben, während bei der mehrstufigen Zufallsauswahl auch innerhalb dieser Teilstichproben noch zufällig Elemente gezogen werden. Somit findet bei der Klumpenauswahl in den letzten Teilstichproben eine Totalerhebung statt, bei der mehrstufigen Zufallsauswahl nicht.

7. Sie möchten für ein Referat eine kleine Umfrage zur Einstellung von Studierenden zum aktuellen Veranstaltungsangebot machen. Hierzu befragen Sie alle Studierenden in Ihrem Seminar. Um was für eine Form der Stichprobenziehung handelt es sich?

Es handelt sich um eine Ad-hoc-Auswahl, da keine zufällige Ziehung stattfindet.

8. Standardfehler und Standardabweichung des Mittelwertes. Welcher von beiden Kennwerten wird stärker von der Stichprobe beeinflusst?

Der Standardfehler wird stärker beeinflusst. Bei unendlich großer Stichprobe nähert er sich dem Wert Null an.

9. In welchem Intervall liegt mit einer Wahrscheinlichkeit von 95% der Populationsmittelwert, falls in einer Stichprobe (N = 40) die folgenden deskriptiven Werte erhoben wurden: Mittelwert = 20, Standardabweichung = 5.

KI = Stichprobenkennwert $\pm 1,96 \cdot$ Standardfehler (SE)

Standardfehler (SE) = Standardabweichung / Wurzel (N)

SE = 5 / Wurzel (40) = 0,79

KI = 20 \pm 1,96 \cdot 0,79 = 20 \pm 1,55

Kapitel 3

1. Was besagt der p-Wert?

Der p-Wert beschreibt die Wahrscheinlichkeit eines gefundenen Ergebnisses in einer Stichprobe, unter der Voraussetzung, dass in der Population die Nullhypothese gültig ist.

2. Nennen Sie Vorteile und Nachteile des ungerichteten Hypothesentestens.

Vorteil: Vorhandene Unterschiede werden unabhängig vom Vorzeichen des Unterschiedes gefunden. Somit können auch „überraschende“ Ergebnisse geprüft werden, bzw. der negative Ausgang einer Studie wird nicht unterdrückt.

Nachteil: Falls die Mittelwertsdifferenz „in die richtige Richtung geht“, ist das Verfahren testschwächer.

3. Sie haben einem Kollegen ein unwirksames Medikament zur Prüfung gegeben. Nach einer Versuchsreihe behauptet dieser allerdings, dass dieses Medikament eine Wirkung hätte. Welchen Fehler hat er begangen?

Einen Alpha-Fehler (Ablehnung der gültigen Nullhypothese).

4. Definieren Sie in statistisch-mathematischer Notation eine Hypothese, welche besagt, dass Alkoholkonsum die Reaktionszeit erhöht.

Es sei μ_1 die mittlere Reaktionszeit in der Population der Testpersonen mit Alkoholkonsum und es sei μ_2 die mittlere Reaktionszeit in der Population der Testpersonen ohne Alkoholkonsum, dann gilt:

$H_0: \mu_1 = \mu_2$ (Nullhypothese)

$H_1: \mu_1 > \mu_2$ (gerichtete Alternativhypothese, niedrigere Werte in der zweiten Gruppe)

Bei einem Alpha-Niveau von 5%.

5. Welchen Unterschied gibt es zwischen Standardabweichung und Standardfehler des Mittelwertes bezüglich des Einflusses der Stichprobengröße?

Die Standardabweichung wird kaum von der Stichprobengröße beeinflusst, während der Standardfehler mit zunehmender Stichprobengröße immer kleiner wird. Geht die Stichprobengröße gegen unendlich, nähert sich der Standardfehler dem Wert Null an.

6. Von welchen Faktoren hängt der Fehler zweiter Art ab?

- Höhe des Alpha-Niveaus
- Einseitige vs. zweiseitige Testung
- Homogenität der Merkmalsverteilung
- Stichprobenumfang
- Größe des Effekts
- Abhängige vs. unabhängige Stichprobe
- Teststarke Prüfverfahren

7. Worin unterscheiden sich parametrische und nichtparametrische Testverfahren?

Die nichtparametrischen Verfahren unterscheiden sich von den parametrischen Verfahren darin, dass die untersuchten Variablen keiner theoretischen Prüfverteilung entsprechen müssen. Sie sind somit an weniger Voraussetzungen gebunden, aber im Allgemeinen auch immer testschwächer als die parametrischen Verfahren.

8. Sie möchten den Erfolg einer neuen Unterrichtsform überprüfen. Welcher der beiden Fehler des Hypothesenprüfens wäre hierbei schwerwiegender?

Der Alpha-Fehler, Ablehnung der „richtigen“ Nullhypothese bei gültiger Nullhypothese, würde dazu führen, dass finanzielle Mittel für Einführungen neuer Unterrichtsformen verwendet werden, die vermutlich an anderer Stelle eingespart werden müssten. Der Beta-Fehler, die Beibehaltung der „falschen“ Nullhypothese bei gültiger Alternativhypothese, würde dafür sorgen, dass „alles beim Alten bleibt“. Somit würden mögliche und sinnvolle Veränderungen nicht durchgeführt werden. Welcher von beiden Fehlern der schwerwiegender ist, ist somit eine inhaltliche Fragestellung, die nicht eindeutig beantwortet werden kann.

9. Warum wird der optimale Stichprobenumfang als optimal bezeichnet?

Weil die Stichprobe dann so groß ist, dass ein erwarteter oder größerer Effekt auch statistisch belegt werden kann, während ein kleinerer Effekt statistisch nicht bedeutsam wird.

Kapitel 4

1. Wie lautet der Varianzadditionssatz?

Der Varianzadditionssatz besagt, dass für eine Variable Z, die durch Addition der beiden Variablen X und Y entsteht, gilt: $s^2_z = s^2_x + s^2_y + 2 \cdot s_{xy}$. Die Varianz s_{xy} wird als Kovarianz (covxy) bezeichnet.

2. Welchen Einfluss hat die Streuung einer Variablen auf die Korrelation dieser Variablen X mit einer zweiten Variablen Y?

Die Einengung der Streuung einer Variablen (auch „restriction of range“ genannt) bewirkt, dass der wahre Zusammenhang zwischen beiden Merkmalen in der Population unterschätzt wird.

3. Definieren Sie ein Beispiel für eine positive und ein Beispiel für eine negative Korrelation.

Positive Korrelation: Vorbereitungszeit und Anzahl der richtig gelösten Aufgaben in einer Klausur.
Negative Korrelation: Vorbereitungszeit und Anzahl der Fehler in einer Klausur.

4. Welche Werte können Korrelation und Kovarianz minimalst bzw. maximalst annehmen?

Korrelationen sind nur im Intervall von -1 bis +1 möglich (standardisierte Werte), während Kovarianzen unendlich große Werte annehmen können.

5. Wie könnte eine Semipartialkorrelation auch mittels versuchsplanerischer Methoden ermittelt werden?

Durch Konstanthaltung der Drittvariablen. So könnten beispielsweise in einer Studie nur Personen mit gleichem IQ-Wert berücksichtigt werden.

6. Ein Kollege berichtet Ihnen, dass er aufgrund seiner Daten einen Determinationskoeffizienten von .64 ermittelt hat. Welche Aussagen können Sie über den Zusammenhang zwischen beiden Variablen machen und welche nicht?

Es besteht eine Korrelation von -.8 oder +.8. Somit kann die Größe der Korrelation, nicht aber deren Vorzeichen, bestimmt werden.

Kapitel 5

1. Welches Problem tritt auf, wenn man bei einem Untersuchungsdesign mit einem zweistufigen Faktor statt einer Varianzanalyse einen t-Test rechnet?

Da nur ein Faktor mit zwei Stufen vorliegt, ergibt sich kein Problem. Hier gilt für die statistischen Kennwerte $F = t^2$, d.h. Varianzanalyse und t-Test kommen zum selben Ergebnis.

2. Geben Sie die folgenden Freiheitsgrade an:

- a) $df_{total} =$
- b) $df_{within} =$
- c) $df_{between} =$

- a) $N-1$
- b) $N-p$
- c) $p-1$

3. Welchen Nachteil kann die Definition eines Kontrasts mit sich bringen?

Es können nur jene Mittelwertsdifferenzen statistisch untersucht werden, die innerhalb der Kontraste definiert worden sind.

4. Definieren Sie bei einer Untersuchung mit 5 Gruppen die folgenden 4 Kontraste:

- a) Vergleich der ersten 4 Gruppen mit der letzten Gruppe.
- b) Vergleich von Gruppe 1 und 2 mit Gruppe 3 und 4.
- c) Vergleich von Gruppe 3 und 4.

Sind diese Kontraste unabhängig? Können noch weitere unabhängige Kontraste gebildet werden? Wenn ja, welche?

- a) 1 1 1 1 -4
- b) 1 1 -1 -1 0
- c) 0 0 -1 1 0

Zum Nachweis der Unabhängigkeit aller Kontraste müssen diese jeweils paarweise auf ihre Unabhängigkeit überprüft werden.

Überprüfung a und b: $1*1 + 1*1 + 1*(-1) + 1*(-1) + (-4)*0 = 0$

Überprüfung a und c: $1*0 + 1*0 + 1*(-1) + 1*1 + (-4)*0 = 0$

Überprüfung b und c: $1*0 + 1*0 + (-1)*(-1) + (-1)*1 + 0*0 = 0$

Bei 5 Gruppen können maximal 4 unabhängige Kontraste definiert werden. Somit kann ein weiterer Kontrast definiert werden (1 -1 0 0 0).

5. Ein Kollege behauptet, er würde generell auf die Berechnung von zweifaktoriellen Varianzanalysen verzichten und der Übersicht wegen stets 2 einfaktorielle Varianzanalysen ermitteln.

- a) Was „entgeht“ diesem Kollegen bei diesem Vorgehen?
- b) Warum wird die zweifaktorielle Varianzanalyse als teststärker betrachtet?

- a) Die Ermittlung von Interaktionseffekten.
- b) Durch die Hinzunahme des zweiten Faktors und des Interaktionseffekts kann mehr Varianz an der abhängigen Variablen erklärt werden. Dies führt zu einer Reduktion der Fehlervarianz, was wiederum zu teststärkerer Hypothesenprüfung führt.

6. Sie planen eine Untersuchung, bei der Sie über 2 Monate 4-mal täglich die Konzentrationsfähigkeit von depressiven Patienten erheben wollen. Welches Problem kann auftreten, wenn Sie zur Auswertung eine Varianzanalyse mit Messwiederholung heranziehen wollen?

Bei der Varianzanalyse mit Messwiederholung darf innerhalb einer Person kein fehlender Wert vorliegen, da sonst dieser Fall komplett aus der Analyse ausgeschlossen wird. Die Varianzanalyse mit Messwiederholung geht somit nach dem listweisen Fallausschlussverfahren vor. In Anbetracht der vielen Messzeitpunkte und der (Symptomatik der) Patientengruppe ist die Wahrscheinlichkeit fehlender Werte sehr hoch.

7. Bei einer Untersuchung mit 2 Gruppen und 4 Messzeitpunkten stellen Sie fest, dass die beiden Gruppen sich zum ersten Messzeitpunkt bedeutsam unterscheiden. Wie gehen Sie mit diesem Problem bei der statistischen Auswertung um?

Mittels einer Kovarianzanalyse, bei der die Daten des ersten Messzeitpunktes als Kovariate aufgenommen werden, können die Unterschiede statistisch konstant gehalten werden. Es wird dann eine Kovarianzanalyse mit 2 Gruppen und 3 Messzeitpunkten durchgeführt, in die der erste Messzeitpunkt als Kovariate eingeht.

Kapitel 6

1. Nennen Sie Formen der Validität, die bei der Diskussion einer Studie unterschieden werden.

- Validität statistischer Schlussfolgerungen: Wurden die Voraussetzungen der jeweiligen Verfahren berücksichtigt und kann in Abhängigkeit von der durchgeführten Datenanalyse die gezogene Schlussfolgerung wirklich angenommen werden?
- Interne Validität: Ist der beobachtete Zusammenhang zwischen unabhängiger und abhängiger Variable wirklich anhand der experimentellen Ergebnisse belegbar oder lässt die Studie noch alternative Erklärungsansätze zu?
- Externe Validität: Können die gefundenen Ergebnisse auch unter anderen experimentellen Bedingungen repliziert werden?
- Konstruktvalidität: Kann von den erhobenen, manifesten Variablen wirklich auf das zugrunde gelegte Konstrukt (latente Variable) geschlossen werden?

2. Was wird bei der Validität statistischer Schlussfolgerungen bewertet?

Hierbei wird diskutiert, ob die angemessenen statistischen Verfahren korrekt eingesetzt wurden, d.h., ob die Voraussetzungen (z.B. Normalverteilung, Intervallskalenniveau etc.) gegeben sind, sowie der erreichte Stichprobenumfang. Eventuell können alternativ mehrere Verfahren eingesetzt werden, wobei eines hiervon optimal wäre.

3. Was wird bei der internen Validität bewertet?

Die interne Validität bezieht sich auf die Manipulation der unabhängigen Variablen und die Auswirkungen dieser Manipulation auf die abhängige Variable. Hierbei werden nur diese Zusammenhänge innerhalb des stattgefundenen Experiments diskutiert. Diskutiert wird, inwiefern der beobachtete Zusammenhang zwischen unabhängigen und abhängigen Variablen dem kausalen Zusammenhang entspricht.

4. Worauf bezieht sich die Konstruktvalidität?

Da in der Psychologie meistens latente Merkmale (Konstrukte) untersucht werden, muss immer diskutiert werden, ob diese Konstrukte auch wirklich durch die gemessenen Merkmale erfasst werden können. Hierbei gilt: Je genauer bei der Versuchsplanung die Definition zu messender Merkmale war, desto besser ist die Messung und umso besser kann auch die Konstruktvalidität diskutiert werden.

5. Was wird bei der externen Validität diskutiert?

Hier steht die Übertragbarkeit der Ergebnisse auf andere Kontexte im Fokus der Diskussion. Es wird beispielsweise diskutiert, ob die im Labor gefundenen Ergebnisse auf Alltagssituationen übertragen werden können.

6. Welche Validitätsgefährdungen sind bei statistischen Schlussfolgerungen möglich?

- Zu geringe statistische Power
- Verletzung von Voraussetzungen
- „Fischen“ nach Signifikanzen
- Unreliable Messung
- Restriction of Range
- Einflüsse bekannter Störvariablen

7. Wie kann die interne Validität gefährdet sein?

Nach Cook und Campell (1979) gibt es zur Störung der internen Validität acht Punkte (= THIS MESS):

- Testing
- History
- Instrumentation
- Statistical Regression (Regression zur Mitte)
- Maturation
- Experimental Mortality / Attrition
- Selection
- Selection-Maturation Interaction

Sowie zusätzlich noch die zeitliche Reihenfolge.

8. Was unterscheidet ein experimentelles von einem quasi-experimentellen Design?

Bei einem experimentellen Design findet im Gegensatz zum quasi-experimentellen Design eine randomisierte Zuordnung von Probanden und Untersuchungsbedingungen statt.

9. Wie wird ein Experiment definiert?

Ein Experiment liegt vor, wenn es eine aktive Manipulation der Bedingungen durch den Versuchsleiter gibt und die Zuordnung von Probanden und Untersuchungsbedingungen randomisiert ist.

10. Was sind die Vorteile, wenn zwei Prä-Messungen statt einer einzigen durchgeführt werden?

Der Einfluss der Testung (testing) kann bestimmt und berücksichtigt werden. Auch kann hierdurch belegt werden, dass Kontroll- und Experimentalgruppe vergleichbar sind.

11. Unter welchen Bedingungen empfiehlt sich ein Design mit eingeführtem und wieder abgesetztem Treatment?

Wenn sich die Wirkung des Treatments schnell „verflüchtigt“, beispielsweise der Einfluss von Medikamenten oder von Alkohol.

12. Definieren Sie Treatmentfaktoren, Versuchsfaktoren, Blockfaktoren.

Treatment-Faktoren: Die Faktorstufen bestehen aus unterschiedlichen, durch den Versuchsleiter dargebotenen Reizen, Instruktionen, Behandlungen oder Situationen (= Operationalisierungen der Ausprägungen der unabhängigen Variablen).

Versuchsfaktoren: Faktoren, die durch die Anordnung des Versuchsablaufs definiert werden (z. B. Messwiederholungsfaktoren).

Blockfaktoren: Die Probanden werden nach ihrer Ausprägung in der unabhängigen Variable nach bestimmten Kriterien zu „Blöcken“ zusammengefasst, die die Stufen des Faktors bilden (z.B. IQ von 95–99, 100–104 etc.). Hierbei handelt es sich um keine zufällige Zuweisung.

Kapitel 7

1. Wieso wird bei der Beratung einer Ethikkommission auch ein Methodiker um eine Beurteilung des Studiendesigns gebeten?

Experimente sind nur dann ethisch vertretbar, wenn die Ergebnisse wissenschaftlich verwertbar sind. Diese Verwertbarkeit wird auch durch die statistische Validität bestimmt. Auch muss ein Statistiker bestätigen, dass der optimale Stichprobenumfang zuvor korrekt ermittelt wurde, denn nur mit angemessener Stichprobengröße können Effekte gefunden werden.

2. Welche sind die Vorteile einer schriftlichen Übereinkunft (informed consent) zwischen Versuchsleiter und Probanden anstelle einer mündlichen Aufklärung?

Das Vorgehen ist standardisiert (für alle gleich) und der Proband kann in einer von ihm gewünschten Geschwindigkeit die Informationen aufnehmen. Auch kann so die Information zur vorherigen Prüfung einer Ethikkommission vorgelegt werden.

3. Gilt das Recht auf Widerruf der Einverständniserklärung auch nach dem Abschluss einer Untersuchung? Wenn ja, wie muss mit diesem Widerruf umgegangen werden?

Es gibt dieses Recht auf Widerruf auch nach Abschluss einer Untersuchung. Dann müssen die Daten aus der Datenanalyse ausgeschlossen werden (falls dies möglich ist, da eventuell aufgrund der Anonymisierung diese Zuweisung nicht mehr möglich sein kann).

4. Worin unterscheiden sich Zufallsauswahl und zufällige Zuweisung?

Die Zufallsauswahl bezieht sich auf die Ziehung der Stichprobe und betrifft die externe Validität. Die zufällige Zuweisung bezieht sich auf die Zuordnung der Untersuchungsbedingung zum Probanden und betrifft die interne Validität.

5. Warum könnte bei einer kleinen Stichprobengröße eine einfache Randomisierung zu Problemen führen?

Es kann nicht garantiert werden, dass die Zuordnung zu den Untersuchungsgruppen im gewünschten Verhältnis geschieht, da eine Gleichverteilung auf alle Untersuchungsbedingungen nur bei unendlich vielen Ziehungen vorliegt. Dadurch können unterschiedlich große Gruppen entstehen.

6. Sie möchten ein Experiment mit vier Messungen im wöchentlichen Abstand durchführen. Sie befürchten, dass es über die Messzeitpunkte hinweg einen Schwund an Probanden geben wird. Ein Kollege schlägt vor, die Versuchspersonenbezahlung erst am Ende der Untersuchung durchzuführen, wobei nur bei vollständigen Daten ein Anrecht auf die Bezahlung besteht. Nehmen Sie hierzu auch aus ethischer Sicht Stellung.

Dieses Vorgehen ist ethisch nicht vertretbar, da der einzelne Aufwand der Probanden bezahlt werden muss. Somit muss entweder nach jeder Messung eine Auszahlung der Gratifikation erfolgen oder es muss beim Abbruch der Untersuchung eine anteilige Bezahlung angeboten werden.

7. Im Rahmen einer Studie möchten Sie eine Gruppe von schwer depressiven Patienten mit einer Kontrollgruppe aus der Bevölkerung vergleichen. Bei welcher Gruppe sollten Sie eine intensivere Probandenbetreuung durchführen?

Da in der Patientengruppe aufgrund des Krankheitsbildes vermutlich eine geringere Motivation vorliegt, sollte diese besser betreut werden.

8. Welche ersten Schritte sollten durchgeführt werden, falls in einem Datensatz fehlende Werte auftreten?

Es sollte mittels einer Missing-Data-Diagnose untersucht werden, ob es ein Muster der fehlenden Werte gibt. Liegen beispielsweise Häufungen von fehlenden Werten bei bestimmten Personen oder Variablen vor, muss dies berücksichtigt werden.