

Anleitung zur Gestaltung des Ergebnisteils einer empirischen psychologischen Untersuchung

Formale Aspekte dieser Anleitung beziehen sich auf die Richtlinien der „American Psychological Association-APA“. Fragen, die sich über diesen Leitfaden hinausgehend zu den Richtlinien der APA ergeben, lassen sich durch Nachschlagen im „Publication Manual“ der APA (mehrere Exemplare müssten in der Fakultätsbibliothek stehen) klären oder durch einen Blick auf Zeitschriften, die von der APA herausgegeben werden (z.B. „Journal of Experimental Psychology“, „Developmental Psychology“ usw.).

Warum diese Anleitung? Nicht, weil diese und nur diese Vorgaben die einzig akzeptablen sind und auch nicht, weil es wichtig ist, dass alle studentischen Arbeiten bis auf den letzten Punkt formal übereinstimmen, sondern einfach deshalb, weil es unendlich viel Zeit spart, ein vernünftiges Schema vorzugeben, anstatt jede Arbeit mit ihrem Verfasser in all ihren Details zu besprechen.

1 Gruppenvergleiche

1.1 Beschreibende Darstellungen

Es darf nicht sein, dass zwar die Ergebnisse eines Signifikanztests berichtet werden, aber keine deskriptiven Kennwerte.

Bei letzterem empfiehlt sich folgende Vorgangsweise:

1.1.1 Wenn nicht allzu viele Bedingungskombinationen (Faktoren mal Stufen der einzelnen Faktoren) vorliegen, dann ist eine **Tabelle** zu bevorzugen (Messwerte können exakt abgelesen werden, Streuungsmaße lassen sich einfach mitberichten [=> Abschätzung von Effektstärken möglich]). Tabellen sollten nicht nur Maße der zentralen Tendenz (z.B. arithmetisches Mittel) enthalten, sondern auch Dispersionsmaße (Standardabweichung am gängigsten; Range ist meist recht uninteressant). In dieser Hinsicht auch auf Übereinstimmung zwischen den eingesetzten Signifikanztests und den berichteten deskriptiven Kennwerten achten (wenn z.B. U-Test statt t-Test, dann (auch) Mediane und nicht nur arithmetisches Mittel).

Beispieltabelle

Tabelle 12

Mittelwerte und Standardabweichungen leseschwacher Kinder und schriftsprachlich unauffälliger Kinder bei phonologischer Segmentation sowie Benennungsschnelligkeit

Test	Leseschwache		Kontrollkinder	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Segmentationsfähigkeit ^a (Fehler)				
Alliterationserkennung (Max = 16)	2.9	2.7	2.5	2.8
Reimerkennung (Max = 15)	2.2	2.6	2.4	2.8
Benennungsschnelligkeit (s/Item)				
Objekte ^a	1.55	0.33	1.33	0.20
Ziffern ^b	0.76	0.19	0.61	0.11

^a Am Beginn der 1. Klasse erhoben.

^b Am Ende der 2. Klasse erhoben.

Referenz im Text

Alles, was über den reinen Text hinausgeht (Tabellen, Abbildungen, Bilder, ...) bedarf einer Referenz im Text. Wenn die Tabelle selbsterklärend ist, kann dies z.B. in Verbindung mit dem Bericht eines Ergebnisses geschehen („Wie Tabelle 2 zeigt, wiesen die lese-rechtschreibschwachen Kinder auf“). Wenn die Tabelle erklärt werden muss, dann z.B.: „Tabelle 2 zeigt die Ergebnisse der lese-rechtschreibschwachen Kinder sowie der Kontrollkinder in den durchgeführten kognitiven Tests. Die in eckigen Klammern ausgewiesenen Werte stellen die Prozentränge der Gruppenmittelwerte innerhalb der Normstichprobe dar.“ (Erklärungen zur Tabelle können auch über Fußnoten erfolgen; s. weiter unten).

Tabellenüberschrift

Jede Tabelle bedarf einer Überschrift. Diese umfasst „Tabelle [fortlaufende Nummer]“ und darunter die Beschreibung des Tabelleninhalts. Bei letzterem gilt folgendes Prinzip: Wenn nicht sonst zu viel zu sagen ist, dann hinschreiben, welche Kennwerte (z.B. Mittelwert, Standardabweichung, Median, ...) berichtet werden; darüber hinaus erwähnen, welche abhängigen Variablen; ev. noch Gruppen von VPs benennen (s. oben).

Räumliche Anordnung

Wenn möglich, dann Versuchspersonengruppen als Spalten und abhängige Variablen als Zeilen.

Wenn viel Platz, dann Mittelwert und Standardabweichung in separaten Spalten, wenn wenig Platz, dann Standardabweichungen in runden Klammern hinter Mittelwert.

Bei der Beschriftung von Bedingungskombinationen aus Messwiederholungsfaktoren wird Redundanz möglichst vermieden (hierarchischer Aufbau). Also nicht z.B. in einer Zeile „Benennungsschnelligkeit-Objekte“ und in der nächsten Zeile „Benennungsschnelligkeit-Ziffern“, sondern Überschriftenzeile „Benennungsschnelligkeit“ und dann darunter - vom linken Rand eingezogen - „Objekte“ und in nächster Zeile „Ziffern“.

Bezeichnungen innerhalb einer Tabelle müssen weitestgehend selbsterklärend sein, d.h. Abkürzungen von Variablennamen sollten wenn möglich vermeiden werden. Wenn sie aus Platzgründen erforderlich sind, dann diese unterhalb der Tabelle auflösen (z.B. „BENSCH = Benennungsschnelligkeit, OBJ = Objekte, ZIF = Ziffern“).

Tabelle als möglichst autonom

Eine Tabelle soll im günstigsten Falle zu verstehen sein, ohne dass im Text nachgelesen werden muss (eine Ausnahme stellt nur die Frage dar, wie die einzelnen Aufgaben oder Tests genau ausgesehen haben; hier muss der Leser unter Umständen dann im Methodenteil nachlesen).

Was sollte man gegebenenfalls noch an Information in die Tabelle oder in die Tabellenerläuterung unterhalb der Tabelle hineinpacken:

Maßeinheit (bei jedem Test sind verschiedene Maßeinheiten denkbar und seien es nur „Anzahl richtiger Antworten“ vs. „Anzahl falscher Antworten“)

Theoretische Maximalwerte einer Aufgabe (ist hinfällig, wenn z.B. % richtiger Antworten oder Werte standardisierter Variablen berichtet werden); falls der Wertebereich nicht bei 0 beginnt, dann auch theoretischen Minimalwert berichten

Messzeitpunkte (wenn es mehrere davon gibt)

Inhaltliches

Innerhalb ein und derselben Tabelle ist es wünschenswert, die Maßeinheit möglichst einheitlich zu halten (z.B. nicht einmal Fehler und einmal Richtige zählen; nicht einmal Absolutwerte richtiger Antworten und einmal % richtiger Antworten)

Prozentwerte haben den Vorteil, dass der Schwierigkeitsgrad einer Aufgabe unmittelbar zu ersehen ist. Darüber hinaus lassen sich Messwiederholungseffekte sofort inhaltlich interpretieren (z.B. wenn Mittelwert richtiger Antworten in einer Aufgabe aus 10 Items bei 8 Items liegt (80%) ist die Aufgabe leichter als eine Aufgabe mit durchschnittlich 50 richtigen Items von 100 (50%).

Auf die Darstellung von Prozentwerten sollte verzichtet werden, wenn die Anzahl der beurteilten Ereignisse recht gering ist (z.B. bei einem Test mit 8 Items entspricht eine Antwort einem Ausmaß von 12.5 %; unter Verwendung von Prozentwerten wird damit ein Auflösungsgrad der Messung vorgetäuscht, der nicht wirklich gegeben ist)

Standardwerte (Messwerte, die an Mittelwert und Standardabweichung relativiert werden) haben den Vorteil, dass Gruppenmittelwerte als eine Art von Effektstärken interpretiert werden können. Gleichzeitig werden Messwiederholungseffekte zum Verschwinden gebracht.

1.1.2 Wenn *viele Bedingungskombinationen* gegeben sind und vor allem dann, wenn es dabei Interaktionseffekte gibt, dann sollte die Verwendung von Diagrammen erwogen werden. Die Effekte lassen dabei rascher ersehen.

Beispiel eines Liniendiagramms

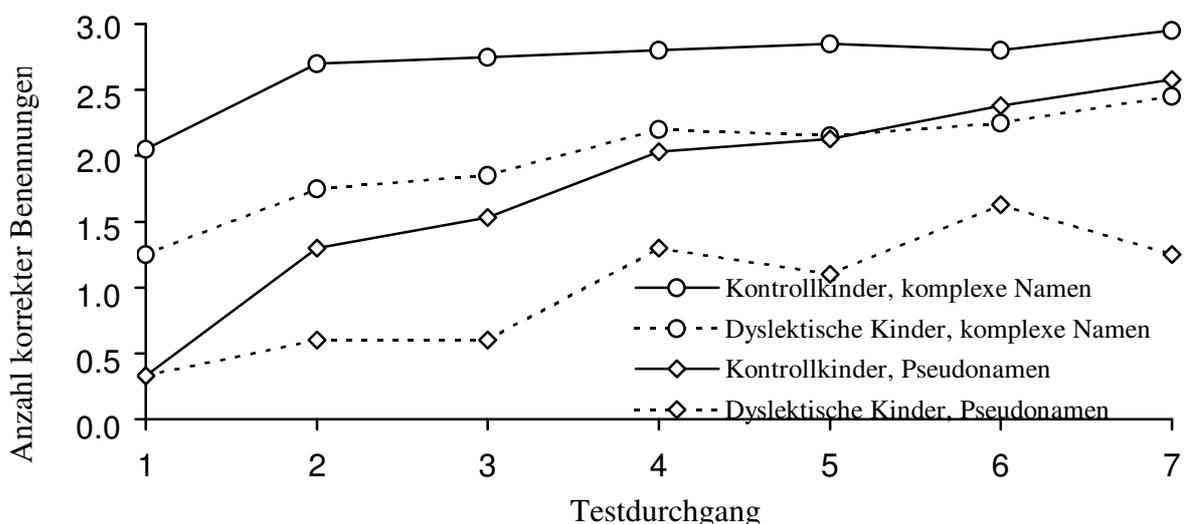


Abbildung 1. Anzahl korrekter Benennungen über die Testdurchgänge hinweg und separat für Probandengruppen und Lernbedingungen.

Beispiel eines Balkendiagramms

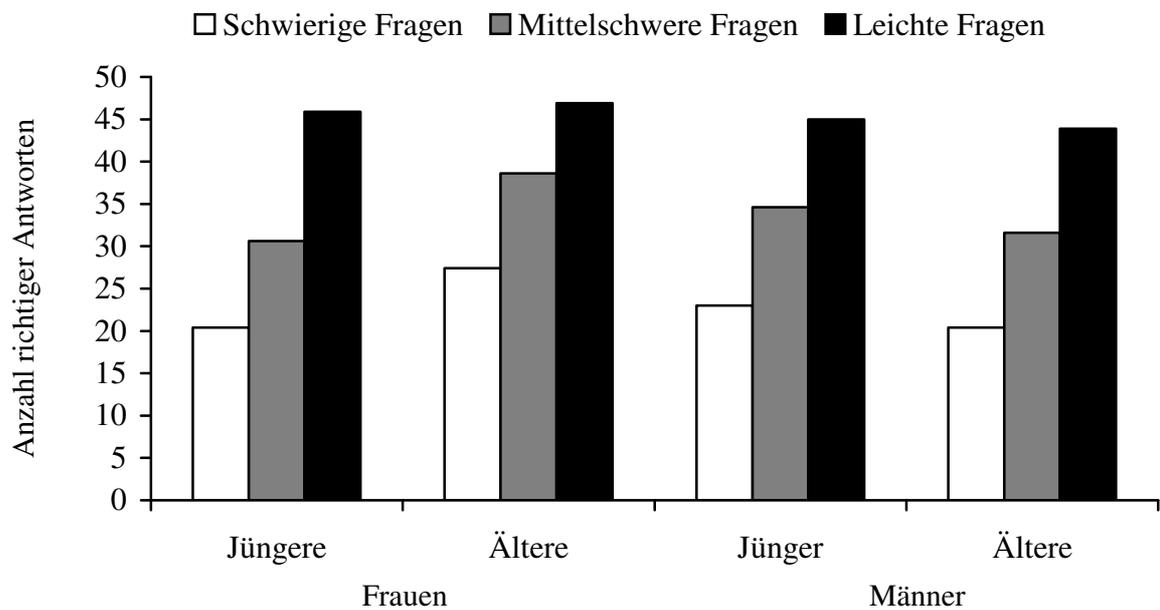


Abbildung 4. Anzahl richtiger Antworten zu politischen Wissensfragen getrennt nach Geschlecht und Alter der befragten Personen.

Erläuterungen

Auch Diagramme müssen wieder im Text erwähnt werden.

Die Titulierung erfolgt, anders als bei Tabellen, unterhalb der Darstellung. Die Nummerierung erfolgt fortlaufend für Diagramme, d.h. die Nummerierung von Tabellen und Diagrammen ist voneinander unabhängig (beide beginnen mit 1).

1.1.3 Wenn *ganz wenige Bedingungskombinationen* gegeben sind, dann können z.B. Mittelwert und Standardabweichungen im Text mitberichtet werden (z.B. „Mit einem Mittelwert von 7.3 ($SD = 2.1$) richtigen Antworten schnitten die Mädchen besser ab als die Buben, die im Mittel nur 3.1 ($SD = 1.8$) richtige Antworten geben konnten, ...“).

1.2 Inferenzstatistische Ergebnisse (Signifikanztests)

Tabelle oder Text?

Was es dezidiert nicht gibt, sind SPSS-Outputs als Anhang oder ausgeschnitten als Pseudotabellen sowie Ergebnisse von Inferenztests in selbstgemachten Tabellen. Inferenzstatistische Ergebnisse sind im Fließtext (Doppelpunkte möglichst vermeiden) zu berichten (z.B.: „Wie Tabelle 2 zeigt, wiesen die lese-rechtschreibschwachen Kinder eine deutlich verringerte Benennungsschnelligkeit bei Objektabbildungen auf, $t(27) = 2.2, p = .04$ “).

Wie genau muss ich beschreiben, welchen Test bzw. welches Design ich gerechnet habe?

Wenn völlig evident ist, was berechnet wurde, braucht nichts Näheres dazu ausgeführt werden (z.B. in obigem Fall, wo genau 1 abhängige Variable im Satz spezifiziert wurde und es nur zwei Versuchspersonengruppen gibt, ist klar, dass es sich um einen t-Test für unabhängige Stichproben [im Unterschied zu t-Test für abhängige Stichproben] handelt, und es ist klar, welche Gruppen verglichen wurden [in Ermangelung einer dritten Gruppe]). Bei komplizierteren Analysen, ist einleitend darzustellen, was berechnet wurde, z.B. „Die Daten zur Benennungsschnelligkeit wurden anhand einer zweifaktoriellen Varianzanalyse mit der Lesefähigkeit (lese-rechtschreibschwach vs. schriftsprachlich unauffällig) als Gruppenfaktor, der Stimulusart (Objekte vs. Ziffern) als Messwiederholungsfaktor und dem Zeitbedarf als abhängige Variable analysiert.“

In welcher Abfolge werden Effekte einer ANOVA berichtet?

Mit Effekten sind hier die Auswirkungen aller systematischen Varianzquellen gemeint (Haupteffekte, Interaktionen), und zwar unabhängig davon, ob sich die Effekte dann als signifikant herausstellen (also Effekt im algebraischen und nicht empirischen Sinne). Es ist wünschenswert, die Effekte in hierarchischer Abfolge zu berichten (erst die Haupteffekte, dann Interaktionseffekte 1. Ordnung, Interaktionseffekte 2. Ordnung, ...). Achtung: sofern Messwiederholungsfaktoren im Spiel sind, liefert SPSS einen Output, der nach anderen Gesichtspunkten geordnet ist, also Abfolge nicht einfach übernehmen.

Muss ich alle Effekte einer ANOVA berichten?

Von oben beschriebenen Standardvorgehen kann abgewichen werden, wenn bestimmte Effekte aus inhaltlichen Gründen überhaupt nicht interessant sind (ev. Haupteffekte bestimmter

Messwiederholungen; hier interessiert vielleicht nur die Interaktion mit einem Gruppenfaktor). Zur Frage, ob nur signifikante Effekte berichtet werden oder alle Effekte, gibt es keine verbindliche Antwort. Eine Möglichkeit ist, nur die signifikanten Ergebnisse mit Testgrößen und Irrtumswahrscheinlichkeiten zu berichten, und darüber hinaus zu sagen, dass kein anderer Haupteffekt oder keine andere 2er Interaktion, signifikant ist (ev. sagen, dass alle F-Werte kleiner oder maximal ? sind). Wenn aber das Interesse z.B. auf einen bestimmten Haupteffekt hinausläuft, dann ist es recht merkwürdig, wenn dieser im Falle einer Insignifikanz nicht berichtet wird (manchmal ist es auch oder gerade ein insignifikantes Ergebnis das interessiert). Wenn sonst keine Entscheidungskriterien vorliegen, bietet sich ein pragmatischer Weg an: insignifikante Ergebnisse werden dann berichtet, wenn das Design nur wenige Effekte beinhaltet. Wenn das Design viele Faktoren umfasst, dann werden insignifikante Testergebnisse meist nicht mit Testgrößen und Irrtumswahrscheinlichkeiten berichtet.

Wie beschreibt man einen signifikanten Effekt?

Effekte sind nicht in formaler Sprache, sondern inhalts geladen zu beschreiben!!!

Bei einer Interaktion reicht z.B. folgendes nicht aus: „Für die Benennungsschnelligkeit zeigte sich eine signifikante Interaktion zwischen der Stimulusart und dem Gruppenfaktor, ...“ Entweder man führt im Anschluss an eine solche Formulierung Näheres aus, oder - noch besser - man kann es gleich genauer ausführen, z.B. „Für die Benennungsschnelligkeit zeigte sich, dass die lese-rechtschreibschwachen Kinder bei Objektabbildungen in stärkerem Maße gegenüber den Kontrollkindern verlangsamt waren als bei den Ziffern, $F(?, ?) = ?, p = ?$ “. Dabei muss aber darauf geachtet werden, dass klar ist, dass sich das Ergebnis des berichteten Signifikanztests auf die Interaktion bezieht (und nicht z.B. auf den Gruppeneffekt für jene abhängige Variable [hier Objektbenennen], bei der der größere Gruppenunterschied gegeben ist).

Bei Haupteffekt nicht „Für die Benennungsschnelligkeit zeigte sich ein signifikanter Gruppeneffekt, $F(?, ?) = ?, p = ?$ “, sondern „Die lese-rechtschreibschwachen Kinder zeigten gegenüber den Kontrollkindern eine erhöhte Benennungszeit, $F(?, ?) = ?, p = ?$ “.

Form der Darstellung inferenzstatistischer Ergebnisse

Test	Darstellung im Text	Erläuterung
Varianzanalyse	„..., $F(1, 30) = 25.8, p < .001.$ “	In Klammer stehen der Zählerfreiheitsgrad (Freiheitsgrade des Faktors) und - danach - der Nennerfreiheitsgrad (Freiheitsgrad der Stichprobenfehlerschätzung) des F-Terms; nach der Klammer folgt die sog. Testgröße (in diesem Fall der F-Wert); nach dem „p“ steht die Irrtumswahrscheinlichkeit
t-Test	„..., $t(27) = 3.3, p = .003.$ “	In Klammer steht der Freiheitsgrad der Analyse; dann die Testgröße, dann Irrtumswahrscheinlichkeit
U-Test	„..., $U(n_1 = 14, n_2 = 15) = 63, p = .07$ “	In Klammern stehen die beiden Gruppengrößen (Anzahl der Probanden), dann U-Wert, dann Irrtumswahrscheinlichkeit
Wilcoxon	„..., $Z(N = 32) = 4.9, p < .001$ “	In Klammern steht der Stichprobenumfang (Anzahl der Probanden), dann Testgröße, dann Irrtumswahrscheinlichkeit

2 Korrelative Analysen

Beispiel für eine Interkorrelationstabelle (jede Variable mit jeder anderen korreliert)

Tabelle 11

Einfache Korrelationen zwischen verbalen Aufgaben

	(1)	(2)	(3)	(4)
(1) Alliterationserkennung				
(2) Reimerkennung	.32			
(3) Pluralbildung	.37	.36		
(4) Pseudowortnachsprechen	.35	.29	.38	
(5) Artikulationsgeschwindigkeit	.30	.21	.32	.26

Beispiel für Korrelationen mehrerer Variablen mit jeweils 1 oder mehreren anderen Variablen

Tabelle 13

Einfache Korrelationen des Pseudonamenlernens sowie phonologischer Wortrepräsentationen mit anderen Fähigkeitsbereichen

Fähigkeitsbereich	Pseudonamenlernen	Phonologische Wortrepräsentationen
Nonverbale Fertigkeiten	.03	.16
Sprachperzeption	.15	.47
Segmentationsfähigkeit	.12	.45
Buchstabenkenntnis	.37	.59
Benennungsschnelligkeit	.50	.22
Artikulationsgeschwindigkeit	.28	.38
Phonologisches Kurzzeitgedächtnis	.60	.65

Anmerkung. Alle Korrelationen sind so gepolt, dass ein positiver Wert einen positiven Zusammenhang in bezug auf die Leistung und nicht in bezug auf den Messwert ausdrückt.

Was das Lesen einer Korrelationstabelle wesentlich vereinfacht!

Es ist ganz nett, wenn alle Korrelationen hinsichtlich des Vorzeichens gleich zu interpretieren sind. Also - wie schon bei den Gruppenvergleichen erwähnt - nicht für eine Variable mit einem Fehlerwert rechnen und für eine andere mit der Anzahl richtiger Reaktionen. Wenn z.B. eine Zeitvariable mit einer Punktwertvariable korreliert wird, dann Punktwert für Fehler heranziehen und nicht Punktwert für richtige Reaktionen (viel Zeit und viel Fehler bedeuten in Leistungstests gleichermaßen eine schlechte Leistung). Wenn dieses Prinzip nicht berücksichtigt wird, dann muss man bei jeder einzelnen Korrelation überlegt werden, welche Maßeinheit für die beiden Variablen verwendet wurden, um das Vorzeichen interpretieren zu können. Zur Not kann - um die Vorzeichen einheitlich zu halten - auch von der für die deskriptive Darstellung von Skalen- oder Gruppenwerten gewählten Maßeinheit abgewichen werden (s. Anmerkung zu obiger Tabelle).

Zusätzliche Möglichkeiten bei Interkorrelationstabellen

Es kann z.B. so verfahren werden, dass auf der einen Seite der Diagonale einfache Korrelationen berichtet werden (zero order correlations) und auf der anderen Seite Korrelationen, für die bestimmte Variablen kontrolliert wurden (über Partialkorrelation oder vorausgehende Regression der Kriteriumsvariable). Wenn Reliabilitäts- oder Validitätskoeffizienten zu Variablen vorliegen (z.B. Korrelation des Tests mit sich selbst zu zwei Messzeitpunkten [Retestrelibilität] oder testlängenkorigierte Halbierungsreliabilität), so können diese in die Diagonale geschrieben werden => Korrelation der Variable mit sich selbst in Diagonale, Korrelation mit anderen Variablen abseits der Diagonale.

Signifikanztests

Das Signifikanzniveau von Korrelationen ist vielfach nicht besonders interessant; meist kann man sich darauf beschränken, die Höhe der Korrelation zu betrachten. Wenn Signifikanztest mitberichtet werden, dann ist folgende Darstellung zu wählen: „..., $r(29) = .58, p = .001$ “. In Klammern stehen die Freiheitsgrade des Tests.

Beispieltabelle für multiple Regressionen

Tabelle 7

Ergebnisse hierarchischer Regressionsanalysen mit Intelligenz, phonologischer Sensitivität und Benennungsschnelligkeit als Prädiktoren sowie Lesegeschwindigkeit und Lesegenauigkeit als Kriteriumsmaße

Schritt	Prädiktor	Lesegeschwindigkeit		Lesegenauigkeit	
		<i>R</i>	<i>R</i> ² -Change	<i>R</i>	<i>R</i> ² -Change
1	IQ	.11	.01**	.10	.01*
2	Phonologische Sensitivität	.20	.03***	.28	.07***
3	Benennungsschnelligkeit	.31	.06***	.33	.03***

* $p < .05$ ** $p < .01$ *** $p < .001$

Bei schrittweise dargestellten Regressionsanalysen interessiert auch immer das Signifikanzniveau der einzelnen Erklärungsbeiträge!

3 Wann soll man nonparametrische (verteilungsfreie) Verfahren einsetzen?

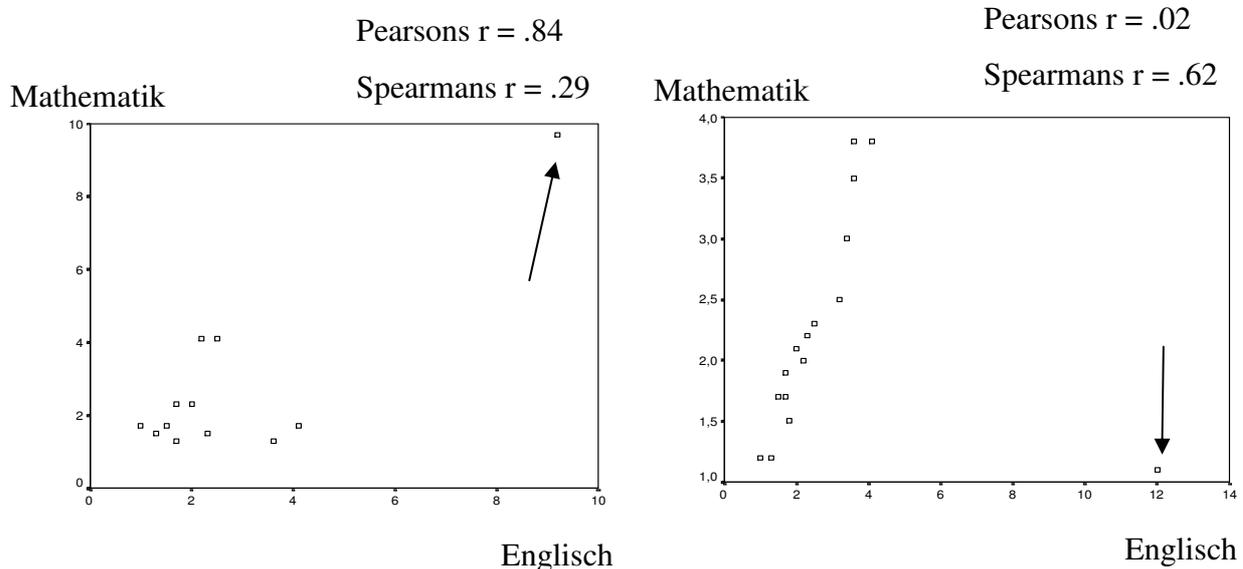
Die Verteilungs- bzw. Unabhängigkeitsvoraussetzungen die für Verfahren wie Varianzanalyse, t-Test sowie Pearsons Produkt-Momenten-Korrelation gegeben sein müssen, sollten aus dem Statistikkunterricht bekannt sein; bei Interesse können sie praktisch in jedem Statistikbuch, in dem das gewünschte Verfahren behandelt wird, nachgelesen werden. Interessanter aber ist, wie in der Praxis mit Verletzungen bestimmter Voraussetzungen umgegangen wird. Unter günstigen Voraussetzungen sind parametrische Verfahren relativ unempfindlich gegenüber Verletzungen der Voraussetzungen. Dies lässt sich daran erkennen, dass nonparametrische Verfahren (z.B. U-Test statt t-Test, Spearmans Rangkorrelationen statt Pearson) oft eine Irrtumswahrscheinlichkeit ergeben, die nicht wesentlich von jener des voraussetzungsgebundenen Verfahrens abweicht. Der einfachste Weg: wenn Zweifel über die Angemessenheit eines voraussetzungsgebundenen Verfahrens besteht, einfach beide Arten der Analysen rechnen. Wenn sich gravierende Divergenzen ergeben, dann unbedingt das nonparametrische Verfahren berichten, im anderen Fall ist es eher Geschmackssache, was man berichtet.

Trotzdem einige Anhaltspunkte

Es gibt (zumindest) vier wichtige Hinweise, die für die Verwendung nonparametrischer Verfahren sprechen:

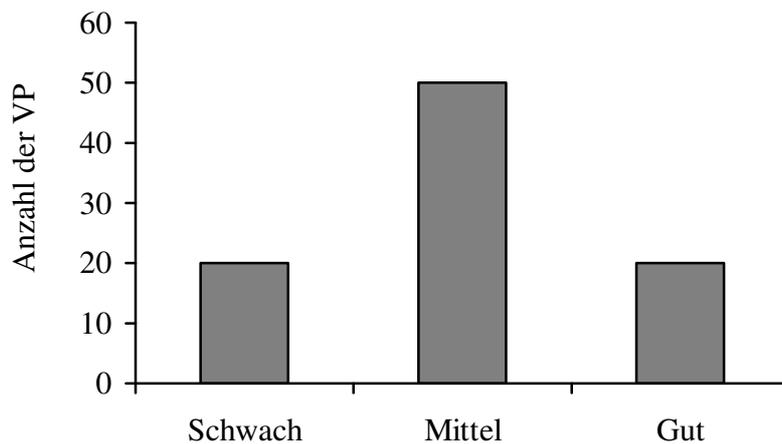
- 1) Es ist aus inhaltlichen Gründen eindeutig, dass nur Ordinalskalenniveau vorliegt (z.B. 23 Länder wurden nach ihrer Einwohnerzahl von 1-23 durchnummerieren)
- 2) Es gibt Ausreißer:

Hier Beispiele dafür, wie Ausreißer eine Korrelation in die Höhe treiben bzw. drücken können

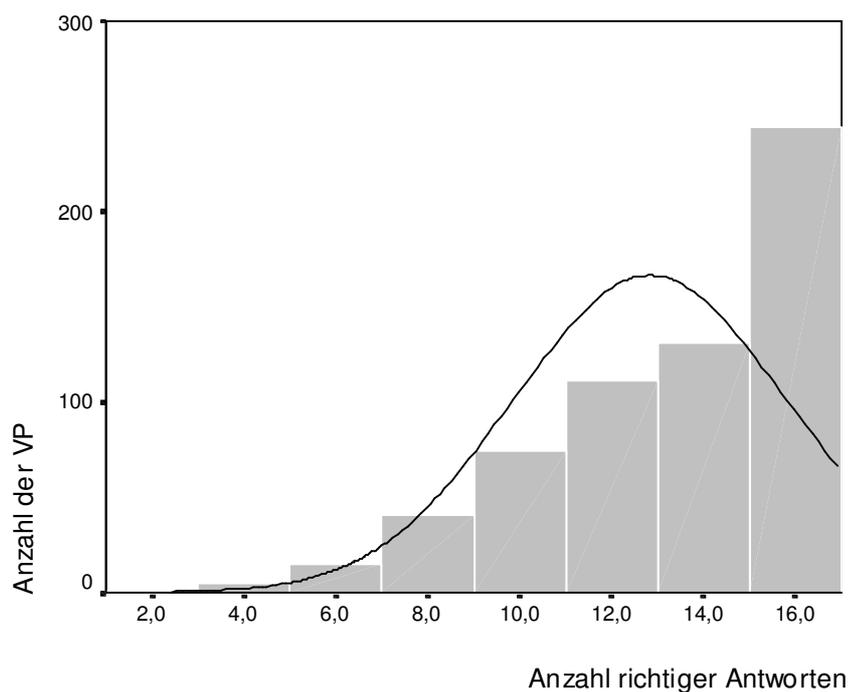


3) Sehr geringer Auflösungsgrad der Messung:

Wenn z.B. nur drei Ausprägungsgrade eines Merkmals unterschieden werden, so ist die Verwendung von voraussetzungsgebundenen Verfahren **auch dann problematisch, wenn** die Häufigkeitsverteilung **symmetrisch und deutlich unimodal** (eingipfelig) ist (s. Abbildung).



4) Die Messung weist zwar einen einigermaßen hohen Auflösungsgrad auf (z.B. Rechentest mit 17 Aufgaben), die Verteilung ist aber sehr schief. Dies bewirkt, dass faktisch dann unter Umständen nur mehr wenige verschiedene Messwerte auftreten und dass Fälle am wenig besetzten Ende der Verteilung ein wenig Ausreißercharakter haben. Hier ein noch nicht besonders dramatisches Beispiel:



4 Darstellung von Zahlen

Die Anzahl der Nachkommastellen sollen mit Bedacht gewählt werden. Wenn vor dem Komma noch einigermaßen was steht, dann reicht oft eine einzige Nachkommastelle aus, zwei sind aber praktisch immer ausreichend. Für ein bestimmtes Maß bzw. eine bestimmte Kennzahl soll die Anzahl der Nachkommastellen einheitlich gehalten werden. Eine Ausnahme stellt dabei die Irrtumswahrscheinlichkeit dar. Vernünftiges Prinzip: 2 Nachkommastellen; wenn aber durch Rundung die Erreichung eines der 3 üblichen Signifikanzniveaus ($< .05$, $< .01$, $< .001$) verschwinden würde, dann noch eine dritte Stelle (z.B. aus $.048$ nicht $.05$ machen; $.052$ kann hingegen auf $.05$ gerundet werden). Achtung: $p = .000$ wird nicht geschrieben, sondern $p < .001$!

Eine Null vor dem Komma wird nur dann geschrieben, wenn der Zahlenwert größer als 1 werden kann (z.B. „...benötigten im Durchschnitt 0.4 Sekunden.“). Bei statistischen Größen, die nicht größer als 1 werden können (z.B. Korrelationskoeffizienten, Wahrscheinlichkeiten) entfällt die Null vor dem Komma (z.B. $p = .02$).

Auch die Deutsche Gesellschaft für Psychologie verlangt, dass das Komma - wie dies im Englischen der Fall ist - als Punkt geschrieben wird und nicht - wie im Deutschen üblich - als Bindestrich.