

Aufgaben zu Kapitel 7:

Aufgabe 1:

In einer Klinik sollen zwei verschiedene Therapiemethoden miteinander verglichen werden. Zur Messung des Therapieerfolges werden die vorhandenen Symptome einmal vor Beginn der Therapie, nach der Therapie und drei Monate später auf einer intervallskalierten Skala erfasst. Die Patienten werden zufällig der Therapie A oder der Therapie B zugeordnet. Vorstudien haben ergeben, dass der Zusammenhang zwischen wiederholten Messungen der Symptome relativ gering ist ($r = 0,2$).

- Wie viele Patienten werden benötigt, um einen Effekt von $f = 0,25$ einer therapeutischen Intervention mit 80%iger Sicherheit zu entdecken, unabhängig von der Art der Therapie?
- Wie viele Patienten sollten teilnehmen, um mit einer Sicherheit von 95% mittelgroße Unterschiede zwischen den Therapieformen in der Art ihres Therapieerfolgs zu entdecken, falls sie existieren?
- Welche Patientenzahl wird benötigt, um mit einer Wahrscheinlichkeit von 80% auszuschließen, dass die Patienten in einer der Therapiearten mehr Symptome aufweisen als in der anderen, unabhängig vom Messzeitpunkt? Hier wird nur einer großer Effekt als inhaltlich relevant erachtet.

Aufgabe 2:

In einer messwiederholten Varianzanalyse gibt SPSS folgenden Output aus:

Mauchly-Test auf Sphärizität^b

Maß: MASS_1

Innersubjekteffekt	Mauchly-W	Approximiertes Chi-Quadrat	df	Signifikanz	Epsilon ^a		
					Greenhouse-Geisser	Huynh-Feldt	Untergrenze
FaktorA	,560	9,692	5	,085	,781	,905	,333
FaktorB	1,000	,000	0	.	1,000	1,000	1,000
FaktorA * FaktorB	,002	108,534	5	,000	,345	,347	,333

Prüft die Nullhypothese, daß sich die Fehlerkovarianz-Matrix der orthonormalisierten transformierten abhängigen Variablen proportional zur Einheitsmatrix verhält.

a. Kann zum Korrigieren der Freiheitsgrade für die gemittelten Signifikanztests verwendet werden. In der Tabelle mit den Tests der Effekte innerhalb der Subjekte werden korrigierte Tests angezeigt.

b.

Design: Intercept
Innersubjekt-Design: FaktorA+FaktorB+FaktorA*FaktorB

Aufgaben mit SPSS und G*Power

Rasch, Friese, Hofmann & Naumann (2014). *Quantitative Methoden. Band 2* (4. Auflage). Heidelberg: Springer.

Tests der Innersubjekteffekte

Maß: MASS_1

Quelle		Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Signifikanz	Partielles Eta-Quadrat	Nichtzentralitäts-Parameter	Beobachtete Schärfe ^a
FaktorA	Sphärizität angenommen	,043	3	,014	3,378	,025	,158	10,133	,733
	Greenhouse-Geisser	,043	2,343	,018	3,378	,037	,158	7,914	,651
	Huynh-Feldt	,043	2,716	,016	3,378	,029	,158	9,172	,700
	Untergrenze	,043	1,000	,043	3,378	,083	,158	3,378	,413
Fehler(FaktorA)	Sphärizität angenommen	,229	54	,004					
	Greenhouse-Geisser	,229	42,174	,005					
	Huynh-Feldt	,229	48,880	,005					
	Untergrenze	,229	18,000	,013					
FaktorB	Sphärizität angenommen	,043	1	,043	15,372	,001	,461	15,372	,959
	Greenhouse-Geisser	,043	1,000	,043	15,372	,001	,461	15,372	,959
	Huynh-Feldt	,043	1,000	,043	15,372	,001	,461	15,372	,959
	Untergrenze	,043	1,000	,043	15,372	,001	,461	15,372	,959
Fehler(FaktorB)	Sphärizität angenommen	,050	18	,003					
	Greenhouse-Geisser	,050	18,000	,003					
	Huynh-Feldt	,050	18,000	,003					
	Untergrenze	,050	18,000	,003					
FaktorA * FaktorB	Sphärizität angenommen	,006	3	,002	,720	,544	,038	2,161	,193
	Greenhouse-Geisser	,006	1,036	,006	,720	,412	,038	,746	,128
	Huynh-Feldt	,006	1,042	,006	,720	,412	,038	,751	,128
	Untergrenze	,006	1,000	,006	,720	,407	,038	,720	,127
Fehler(FaktorA*FaktorB)	Sphärizität angenommen	,150	54	,003					
	Greenhouse-Geisser	,150	18,648	,008					
	Huynh-Feldt	,150	18,764	,008					
	Untergrenze	,150	18,000	,008					

a. Unter Verwendung von Alpha = ,05 berechnet

Tests der Zwischensubjekteffekte

Maß: MASS_1

Transformierte Variable: Mittel

Quelle	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Signifikanz	Partielles Eta-Quadrat	Nichtzentralitäts-Parameter	Beobachtete Schärfe ^a
Konstanter Term	12,957	1	12,957	140,184	,000	,886	140,184	1,000
Fehler	1,664	18	,092					

a. Unter Verwendung von Alpha = ,05 berechnet

- Welche Faktoren sind messwiederholt, welche nicht?
- Wie viele Stufen haben die einzelnen Faktoren?
- Für welche Faktoren tritt wahrscheinlich eine Verletzung einer Annahme der messwiederholten Varianzanalyse auf?
- Warum kann für einen der Faktoren keine Verletzung der Sphärizitätsannahme auftreten?
- Wie viele Versuchspersonen wurden untersucht?
- Welches Korrekturverfahren der Freiheitsgrade sollte man für die Testung der einzelnen Faktoren und der Wechselwirkung verwenden und warum?
- Welche Effekte sind signifikant?
- Wie groß war die Teststärke, in dieser Untersuchung einen mittleren Effekt der Wechselwirkung zu finden, falls er existieren würde ($r = 0,73$)?
- Ist das von SPSS angegebene Eta-Quadrat für den Faktor A mit einem großen Effekt aus einer Studie ohne Messwiederholung vergleichbar?
- Was gibt der Output „Tests der Zwischensubjekteffekte“ hier an?

Aufgabe 3:

In einer Untersuchung zum Einfluss der Belohnung auf die Reaktionszeit haben sich folgende Daten ergeben. Die Tabelle gibt die mittleren Reaktionszeiten (in Millisekunden) der einzelnen Versuchspersonen in den verschiedenen experimentellen Bedingungen wieder. Im Unterschied zu der entsprechenden Aufgabe zu Kapitel 5 nimmt eine Versuchsperson an allen vier Bedingungen teil.

Keine Belohnung	5€	10€	20€
534 ms	493 ms	372 ms	420 ms
260 ms	215 ms	210 ms	299 ms
237 ms	283 ms	295 ms	308 ms
437 ms	321 ms	319 ms	222 ms
353 ms	258 ms	311 ms	265 ms
523 ms	439 ms	329 ms	273 ms
635 ms	248 ms	259 ms	320 ms
274 ms	275 ms	219 ms	267 ms
320 ms	342 ms	234 ms	201 ms
302 ms	230 ms	190 ms	240 ms

- Geben Sie die Daten in SPSS ein.
- Rechnen Sie eine einfaktorielle Varianzanalyse mit Messwiederholung ($\alpha = 5\%$). Unterscheiden sich die Gruppen signifikant voneinander? Wie groß ist der Effekt?
- Wie groß sind die Mittelwerte der einzelnen experimentellen Bedingungen?
- Welche Gruppen unterscheiden sich signifikant voneinander?
- Berechnen Sie die mittlere Korrelation zwischen den einzelnen messwiederholten Faktorstufen.
- Wie groß war die Teststärke dieser Untersuchung, einen mittleren Effekt (für unabhängige Stichproben) zu finden?
- Vergleichen Sie die gewonnenen Ergebnisse mit der Analyse aus Kapitel 5. Was ist der Grund für die unterschiedlichen Ergebnisse?

Lösungen

Aufgabe 1

In G*Power müssen Sie für die folgenden Berechnungen die Optionen „F Tests“, „ANOVA: Repeated measures, within factors“ und „A priori“ auswählen.

Die Untersuchung umfasst einen messwiederholten Faktor A mit $p = 3$ Stufen (vor der Therapie, direkt nach der Therapie und drei Monate später) und einen nicht-messwiederholten Faktor B mit $q = 2$ Stufen (Therapie A vs. Therapie B). Die mittlere Korrelation zwischen den Stufen des messwiederholten Faktors wird mit $r = 0,2$ angenommen.

- a) Nach Eingabe der Daten in G*Power ergibt sich eine optimale Patientenzahl von insgesamt 44, um einen Effekt von $f = 0,25$ bei einer mittleren Korrelation von $r = 0,2$ des Therapieerfolgs mit einer Wahrscheinlichkeit von mindestens 80% zu entdecken, falls er existiert. In jeder der beiden Therapiearten sollten also mindestens 22 Patienten teilnehmen.

The screenshot shows the G*Power software interface. At the top, 'Test family' is set to 'F tests' and 'Statistical test' is 'ANOVA: Repeated measures, within factors'. The 'Type of power analysis' is 'A priori: Compute required sample size - given alpha, power, and effect size'. The 'Input Parameters' section includes: 'Determine =>' (checked), 'Effect size f' (0.25), 'alpha err prob' (0.05), 'Power (1-beta err prob)' (0.80), 'Number of groups' (2), 'Number of measurements' (3), 'Corr among rep measures' (0.2), and 'Nonsphericity correction epsilon' (1). The 'Output Parameters' section includes: 'Noncentrality parameter lambda' (10.3125000), 'Critical F' (3.1051566), 'Numerator df' (2.0000000), 'Denominator df' (84.0000000), 'Total sample size' (44), and 'Actual power' (0.8133255).

- b) Hier ist nach einer Stichprobenumfangsplanung der Wechselwirkung zwischen den verschiedenen Messzeitpunkten und der Therapieform gefragt. Dies lässt sich unter „Statistical test“ bei G*Power angeben. Es ergibt sich eine optimale Stichprobengröße von 68 Patienten (34 pro Therapieform), um einen mittleren Effekt der Wechselwirkung mit einer Wahrscheinlichkeit von mindestens 95% zu finden, falls er existiert.

Aufgaben mit SPSS und G*Power

Rasch, Friese, Hofmann & Naumann (2014). *Quantitative Methoden. Band 2* (4. Auflage). Heidelberg: Springer.

The screenshot shows the G*Power interface for an ANOVA: Repeated measures, within-between interaction. The 'Test family' is set to 'F tests' and the 'Type of power analysis' is 'A priori: Compute required sample size - given α , power, and effect size'. The 'Input Parameters' section includes: Effect size $f = 0.25$, α err prob = 0.05, Power ($1 - \beta$ err prob) = 0.95, Number of groups = 2, Number of measurements = 3, Corr among rep measures = 0.2, and Nonsphericity correction $\epsilon = 1$. The 'Output Parameters' section includes: Noncentrality parameter $\lambda = 15.9375000$, Critical F = 3.0647607, Numerator df = 2.0000000, Denominator df = 132, Total sample size = 68, and Actual power = 0.9516947.

- c) In der Stichprobenumfangsplanung für den nicht-messwiederholten Faktor B macht sich die Korrelation zwischen den Messzeitpunkten negativ bemerkbar: Je höher die Korrelation, desto mehr Versuchspersonen werden benötigt. In der Frage wird ein großer Effekt für unabhängige Stichproben $f = 0,40$ ($\Omega^2 = 0,14$) als inhaltlich relevant erachtet und wiederum eine Korrelation von $r = 0,2$ zwischen den Messzeitpunkten angenommen. Wenn man die entsprechenden Daten bei G*Power eingibt, ergibt sich, dass mindestens 26 Patienten (13 pro Gruppe) benötigt werden, um einen großen Effekt bei einer mittleren Korrelation von $r = 0,2$ mit mindestens 80%iger Sicherheit auszuschließen, falls ein nicht signifikantes Ergebnis auftritt.

The screenshot shows the G*Power interface for an ANOVA: Repeated measures, between factors. The 'Test family' is set to 'F tests' and the 'Type of power analysis' is 'A priori: Compute required sample size - given α , power, and effect size'. The 'Input Parameters' section includes: Effect size $f = 0.40$, α err prob = 0.05, Power ($1 - \beta$ err prob) = 0.80, Number of groups = 2, Number of measurements = 3, and Corr among rep measures = 0.2. The 'Output Parameters' section includes: Noncentrality parameter $\lambda = 8.9142857$, Critical F = 4.2596773, Numerator df = 1.0000000, Denominator df = 24.0000000, Total sample size = 26, and Actual power = 0.8171095.

Aufgabe 2:

- a) Faktor A und Faktor B sind messwiederholt. Zusätzliche nicht-messwiederholte Faktoren treten in dieser Analyse nicht auf.
- b) Faktor A hat $p = 4$ Stufen ($df_A = 3$), Faktor B hat $q = 2$ Stufen ($df_B = 1$).
- c) Der Mauchly Test auf Sphärizität kann zur Überprüfung einer möglichen Verletzung der mathematischen Voraussetzung der Sphärizität nur unter Vorbehalten herangezogen werden. Allerdings liefert er in diesem Beispiel für die Wechselwirkung einen sehr geringen p-Wert unter der Nullhypothese, die Annahme der Sphärizität ist also mit großer Sicherheit verletzt. Für den Haupteffekt A wird der Mauchly Test nicht signifikant ($p = 0,085$). Trotzdem ist es hier ratsam, auf Grund der relativ geringen Versuchspersonenzahl ebenfalls von einer Verletzung der Sphärizität auszugehen.
- d) Die Sphärizitätsannahme verlangt, dass die Varianzen der Differenzen zwischen zwei Faktorstufen homogen sind. Bei lediglich zwei Faktorstufen gibt es dementsprechend nur eine solche Varianz. Deshalb kann für den messwiederholten Faktor B mit zwei Stufen keine Verletzung der Sphärizität auftreten. Bei nur zwei Messzeitpunkten ist die Annahme der Sphärizität immer erfüllt.
- e) Die Anzahl der Versuchspersonen kann hier aus jedem der Nennerfreiheitsgrade der drei Prüfvarianzen für die einzelnen Effektarten berechnet werden.

$$df_{A \times V_{pn}} = (p - 1) \cdot (n - 1) \rightarrow 54 = (4 - 1) \cdot (n - 1) \leftrightarrow n = (54 / 3) + 1 = 19$$

$$df_{B \times V_{pn}} = (q - 1) \cdot (n - 1) \rightarrow 18 = (2 - 1) \cdot (n - 1) \leftrightarrow n = (18 / 1) + 1 = 19$$

$$df_{A \times B \times V_{pn}} = (p - 1) \cdot (q - 1) \cdot (n - 1) \rightarrow 54 = (4 - 1) \cdot (2 - 1) \cdot (n - 1) \\ \leftrightarrow n = (54 / 3) + 1 = 19$$

Es wurden insgesamt 19 Versuchspersonen untersucht, die an acht verschiedenen Messzeitpunkten jeweils einen Messwert abgegeben haben.

- f) Für den Faktor B ist keine Korrektur der Freiheitsgrade notwendig, da keine Verletzung der Sphärizität auftreten kann (siehe Aufgabe d). Für die Wechselwirkung sollte das mit Greenhouse-Geisser bezeichnete Korrekturverfahren verwendet werden, da hier eine starke Verletzung der Sphärizität vorliegt ($\epsilon_{\text{Greenhouse-Geisser}} < 0,75$). Für den Haupteffekt A könnte auch die liberalere Freiheitsgradkorrektur nach Hyundt-Feldt gewählt werden ($\epsilon_{\text{Greenhouse-Geisser}} > 0,75$). Allerdings ist es in der Praxis unüblich, in derselben Analyse zwei unterschiedliche Korrekturverfahren zu verwenden. Deshalb sollten Sie auch hier die mit Greenhouse-Geisser bezeichnete Korrektur verwenden.
- g) Der Haupteffekt des Faktors A ist signifikant ($F_{(3;54)} = 3,38; p < 0,04$). Dies gilt ebenso für den Faktor B ($F_{(1;18)} = 15,37; p < 0,002$). Die Wechselwirkung ist nicht signifikant ($F_{(3;54)} < 1$).
- h) Die Teststärke, einen mittleren Effekt der Wechselwirkung (für unabhängige Stichproben) in der Untersuchung zu finden, war nahezu 100%. Die Nullhypothese, dass kein mittlerer Effekt (für unabhängige Stichproben) existiert, kann also mit sehr hoher Sicherheit angenommen werden.

Aufgaben mit SPSS und G*Power

Rasch, Friese, Hofmann & Naumann (2014). *Quantitative Methoden. Band 2* (4. Auflage). Heidelberg: Springer.

The screenshot shows the G*Power software interface. At the top, the 'Test family' is set to 'F tests' and the 'Statistical test' is 'ANOVA: Repeated measures, within factors'. The 'Type of power analysis' is 'Post hoc: Compute achieved power - given alpha, sample size, and effect size'. Below this, there are two main sections: 'Input Parameters' and 'Output Parameters'.

Input Parameters		Output Parameters		
Determine =>	Effect size f	0.25	Noncentrality parameter λ	35.1851852
	α err prob	0.05	Critical F	2.0830373
	Total sample size	19	Numerator df	7.0000000
	Number of groups	1	Denominator df	126
	Number of measurements	8	Power (1 - β err prob)	0.9961424
	Corr among rep measures	0.73		
	Nonsphericity correction ϵ	1		

- i) Die empirische Effektstärke η^2 aus einer messwiederholten Untersuchung kann nicht mit der Effektstärke aus einer Analyse ohne Messwiederholung verglichen werden, da die Anzahl der Faktoren, die Stufen der Faktoren und die Korrelation zwischen den Messwiederholungen einen Einfluss auf die Größe des Effekts haben. Deshalb ist es auch nicht möglich, die empirische Effektstärke einer messwiederholten Analyse anhand der Konventionen von Cohen (1988) zu bewerten, da diese für unabhängige Stichproben formuliert wurden.
- j) Der Output „Test der Zwischensubjekteffekte“ ist nur dann für die Interpretation der Daten interessant, wenn ein nicht-messwiederholter Faktor in der Analyse vorhanden ist. In dem Beispiel zeigt diese Tabelle nur an, dass der Gesamtmittelwert aller Versuchspersonen über alle Bedingungskombinationen signifikant von Null verschieden ist (entspricht einer Testung der Größe der Konstanten in einer multiplen Regression).

Aufgabe 3:

- a) Für die Eingabe der Daten müssen Sie in SPSS vier neue Variable benennen. Eine Versuchsperson gibt in dieser Untersuchung vier Werte ab, die alle in derselben Zeile stehen müssen, um die Messwiederholung deutlich zu machen. Nach der Eingabe sollte Ihr SPSS Datenfenster in etwa wie folgt aussehen:

	RZ_keine	RZ_5Euro	RZ_10Euro	RZ_20Euro
1	534	493	372	420
2	260	215	210	299
3	237	283	295	308
4	437	321	319	222
5	353	258	311	265
6	523	439	329	273
7	635	248	259	320
8	274	275	219	267
9	320	342	234	201
10	302	230	190	240

- b) Zur Berechnung einer einfaktoriellen Varianzanalyse mit Messwiederholung gehen Sie auf „Analysieren“ → „Allgemeines Lineares Modell“ → „Messwiederholung“. Geben Sie im folgenden Fenster einen Namen für den messwiederholten Faktor und die Anzahl an Stufen ein. Fahren Sie fort mittels „Definieren“. Geben Sie hier die vier Variablen der Messzeitpunkte ein (Reihenfolge beachten).

Messwiederholung: Faktor(en) definieren

Name des Innersubjektfaktors:

Anzahl der Stufen:

Hinzufügen

Ändern

Entfernen

Maßbezeichnung:

Hinzufügen

Ändern

Entfernen

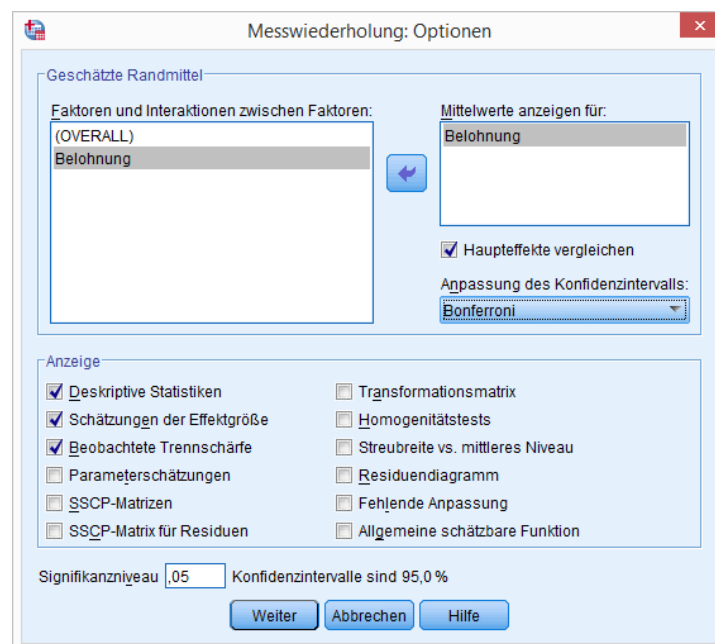
Definieren Zurücksetzen Abbrechen Hilfe

Aufgaben mit SPSS und G*Power

Rasch, Friese, Hofmann & Naumann (2014). *Quantitative Methoden. Band 2* (4. Auflage). Heidelberg: Springer.



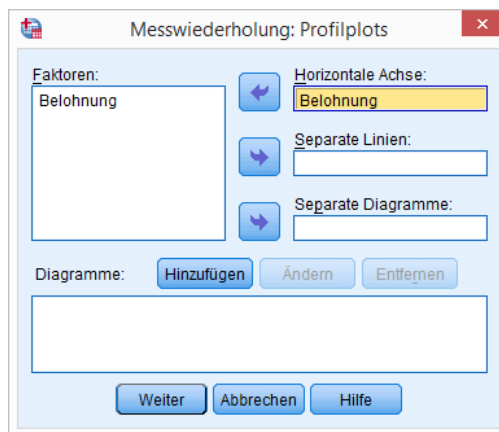
Unter „Optionen“ bewegen Sie den Faktor in das Feld „Mittelwerte anzeigen für“, aktivieren das Kästchen „Haupteffekte vergleichen“ und wählen „Bonferroni“ aus. Aktivieren Sie zusätzlich „Deskriptive Statistiken“, „Schätzungen der Effektgröße“ und „Beobachtete Trennschärfe“.



Nachdem Sie mit „Weiter“ zu dem vorangegangenen Fenster zurückgekehrt sind, können Sie noch über „Diagramme“ eine graphische Darstellung der Mittelwerte definieren.

Aufgaben mit SPSS und G*Power

Rasch, Friese, Hofmann & Naumann (2014). *Quantitative Methoden. Band 2* (4. Auflage). Heidelberg: Springer.



Nach dem Drücken von „Weiter“ starten Sie die Analyse mit „OK“.

Innersubjektfaktoren

Maß: MEASURE_1

Belohnung	Abhängige Variable
1	RZ_keine
2	RZ_5Euro
3	RZ_10Euro
4	RZ_20Euro

Deskriptive Statistiken

	Mittelwert	Standardabweichung	H
Reaktionszeit: keine Belohnung	387,50	136,778	10
Reaktionszeit: 5 Euro Belohnung	310,40	91,409	10
Reaktionszeit: 10 Euro Belohnung	273,80	60,061	10
Reaktionszeit: 20 Euro Belohnung	281,50	61,437	10

Multivariate Tests^a

Effekt	Wert	F	Hypothesen-df	Fehler df	Sig.	Partielles Eta hoch zwei	Dezent. Parameter	Beobachtete Trennschärfe ^c
Belohnung	Pillai-Spur	,563	3,010 ^b	3,000	7,000	,104	,563	,461
	Wilks-Lambda	,437	3,010 ^b	3,000	7,000	,104	,563	,461
	Hotelling-Spur	1,290	3,010 ^b	3,000	7,000	,104	,563	,461
	Größte charakteristische Wurzel nach Roy	1,290	3,010 ^b	3,000	7,000	,104	,563	,461

a. Design: Konstanter Term
Innersubjekt-design: Belohnung

b. Exakte Statistik

c. Berechnet mit alpha = ,05

Aufgaben mit SPSS und G*Power

Rasch, Friese, Hofmann & Naumann (2014). *Quantitative Methoden. Band 2* (4. Auflage). Heidelberg: Springer.

Mauchly-Test auf Sphärizität^a

Maß: MEASURE_1

Innersubjekteffekt	Mauchly-W	Näherungsw eise Chi- Quadrat	df	Sig.	Epsilon ^b		
					Greenhouse- Geisser	Huynh-Feldt (HF)	Untergrenze
Belohnung	,381	7,445	5	,193	,654	,835	,333

Testet die Nullhypothese, dass die Fehlerkovarianzmatrix der orthonormalisierten transformierten abhängigen Variablen proportional zu einer Identitätsmatrix ist.

a. Design: Konstanter Term
Innersubjektdesign: Belohnung

b. Kann für die Anpassung der Freiheitsgrade für die gemittelten Tests auf Signifikanz verwendet werden. Korrigierte Tests werden in der Tabelle 'Tests der Innersubjekteffekte' angezeigt.

In der ersten Tabelle können Sie kontrollieren, in welcher Reihenfolge Sie die einzelnen Stufen des messwiederholten Faktors den entsprechenden Variablen zugeordnet haben. Die Mittelwerte aus Tabelle 2 sollten Ihnen bereits aus den SPSS Aufgaben zu Kapitel 5 bekannt sein. Den multivariaten Test können Sie für die momentanen Zwecke ignorieren. Der Mauchly-Test auf Sphärizität zeigt keine signifikante Verletzung der Sphärizitätsannahme an. Es ist allerdings trotzdem ratsam, von einer Verletzung der Annahme auszugehen, da der Mauchly-Test bei nur zehn Versuchspersonen keine hinreichende Teststärke besitzt, um eine Verletzung der Sphärizitätsannahme zu entdecken. Das unter Greenhouse-Geisser angegebene ϵ ist kleiner als 0,75. Sie sollten für die Signifikanztestung die mit Greenhouse-Geisser bezeichnete Korrektur der Freiheitsgrade verwenden.

Tests der Innersubjekteffekte

Maß: MEASURE_1

Quelle		Typ III Quadratsum me	df	Quadratische r Mittelwert	F	Sig.	Partielles Eta hoch zwei	Dezent. Parameter	Beobachtete Trennschärfe ^a
Belohnung	Angenommene Sphärizität	80855,400	3	26951,800	5,523	,004	,380	16,570	,903
	Greenhouse-Geisser	80855,400	1,962	41210,110	5,523	,014	,380	10,837	,779
	Huynh-Feldt (HF)	80855,400	2,504	32295,624	5,523	,008	,380	13,828	,855
	Untergrenze	80855,400	1,000	80855,400	5,523	,043	,380	5,523	,554
Fehler (Belohnung)	Angenommene Sphärizität	131753,100	27	4879,744					
	Greenhouse-Geisser	131753,100	17,658	7461,275					
	Huynh-Feldt (HF)	131753,100	22,532	5847,268					
	Untergrenze	131753,100	9,000	14639,233					

a. Berechnet mit alpha = ,05

Der Effekt der Höhe der Belohnung ist signifikant ($F_{(3;27)} = 5,52; p < 0,02$). SPSS gibt die Schätzung der Effektgröße mit $\eta^2 = 0,38$ an. Allerdings kann dieser Effekt nicht anhand der Konventionen von Cohen (1988) bewertet werden, da sich die Konventionen auf Effekte aus Studien ohne Messwiederholung beziehen.

c) Mittelwerte in den einzelnen experimentellen Bedingungen:

Deskriptive Statistiken

	Mittelwert	Standardabw eichung	H
Reaktionszeit: keine Belohnung	387,50	136,778	10
Reaktionszeit: 5 Euro Belohnung	310,40	91,409	10
Reaktionszeit: 10 Euro Belohnung	273,80	60,061	10
Reaktionszeit: 20 Euro Belohnung	281,50	61,437	10

Quelle: <http://www.lehrbuch-psychologie.de/qm>

© Rasch, Friese, Hofmann & Naumann

d) Für die Entscheidung, welche experimentellen Bedingungen sich signifikant voneinander unterscheiden, ist eine Post Hoc Testung erforderlich. Allerdings ist die von SPSS angebotene Post Hoc Testung mittels paarweiser Einzelvergleiche mit Bonferroni Korrektur äußerst konservativ. Zusätzlich haben Einzelvergleiche eine geringere Teststärke, da nur die Werte aus jeweils zwei experimentellen Bedingungen in den Vergleich eingehen. Mit dieser Art der Post-Hoc Analyse ergibt sich zwischen keiner der Gruppen ein signifikanter Unterschied:

Paarweise Vergleiche

Maß: MEASURE_1

(I) Belohnung	(J) Belohnung	Mittelwertdifferenz (I-J)	Standardfehler r	Sig. ^a	95 % Konfidenzintervall für Differenz ^a	
					Untergrenze	Obergrenze
1	2	77,100	38,203	,446	-51,423	205,623
	3	113,700	36,590	,075	-9,397	236,797
	4	106,000	39,574	,152	-27,134	239,134
2	1	-77,100	38,203	,446	-205,623	51,423
	3	36,600	19,105	,526	-27,671	100,871
	4	28,900	27,344	1,000	-63,091	120,891
3	1	-113,700	36,590	,075	-236,797	9,397
	2	-36,600	19,105	,526	-100,871	27,671
	4	-7,700	19,458	1,000	-73,161	57,761
4	1	-106,000	39,574	,152	-239,134	27,134
	2	-28,900	27,344	1,000	-120,891	63,091
	3	7,700	19,458	1,000	-57,761	73,161

Basierend auf geschätzten Randmitteln

e) a. Anpassung für Mehrfachvergleiche: Bonferroni.

Alternativ empfehlen wir, die Post Hoc Testung mit Hilfe des Tukey HSD Tests. Allerdings darf dieser Test nicht ohne weiteres angewendet werden, wenn die Zirkularitätsannahme nicht erfüllt ist (siehe Kap. 7.1.9). Wir gehen hier zur Veranschaulichung des Vorgehens von einer erfüllten Sphäritätsannahme aus. Der für den Tukey HSD Test erforderliche kritische q -Wert ist bei vier zu vergleichenden Mittelwerten, 27 Nennerfreiheitsgraden und einem Signifikanzniveau von 5% $q_{krit} = 3,9$ (Tabelle F, Band 1).

$$HSD = q_{krit} \cdot \sqrt{\frac{\hat{\sigma}_{Res}^2}{N}} = 3,9 \cdot \sqrt{\frac{4879,744}{10}} = 86,15$$

Deskriptive Statistiken

	Mittelwert	Standardabweichung	H
Reaktionszeit: keine Belohnung	387,50	136,778	10
Reaktionszeit: 5 Euro Belohnung	310,40	91,409	10
Reaktionszeit: 10 Euro Belohnung	273,80	60,061	10
Reaktionszeit: 20 Euro Belohnung	281,50	61,437	10

Die kritische Differenz von $HSD = 86,15$ wird zwischen den Gruppen „keine Belohnung“ und „10 Euro“ bzw. „20 Euro“ überschritten. Andere signifikante Unterschiede treten nicht auf. Hier zeigt sich deutlich die größere Teststärke des Tukey HSD Tests im Gegensatz zu Einzelvergleichen mit Bonferroni Korrektur.

Aufgaben mit SPSS und G*Power

Rasch, Friese, Hofmann & Naumann (2014). *Quantitative Methoden. Band 2* (4. Auflage). Heidelberg: Springer.

- f) Gehen Sie auf „Analysieren“ → „Korrelation“ → „Bivariat“ und bewegen Sie alle vier Variablen in das rechte Feld. Starten Sie die Analyse mit OK.

		Reaktionszeit: keine Belohnung	Reaktionszeit: 5 Euro Belohnung	Reaktionszeit: 10 Euro Belohnung	Reaktionszeit: 20 Euro Belohnung
Reaktionszeit: keine Belohnung	Pearson-Korrelation	1	,499	,543	,406
	Sig. (2-seitig)		,142	,105	,244
	N	10	10	10	10
Reaktionszeit: 5 Euro Belohnung	Pearson-Korrelation	,499	1	,757*	,414
	Sig. (2-seitig)	,142		,011	,234
	N	10	10	10	10
Reaktionszeit: 10 Euro Belohnung	Pearson-Korrelation	,543	,757*	1	,487
	Sig. (2-seitig)	,105	,011		,153
	N	10	10	10	10
Reaktionszeit: 20 Euro Belohnung	Pearson-Korrelation	,406	,414	,487	1
	Sig. (2-seitig)	,244	,234	,153	
	N	10	10	10	10

*. Korrelation ist bei Niveau 0,05 signifikant (zweiseitig).

Nach der Umwandlung der Korrelationen in Fishers Z-Werte ergibt sich folgender mittlerer Fishers Z-Wert (siehe Tabelle D, Band 1).

$$\bar{Z} = \frac{0,549 + 0,611 + 0,43 + 0,984 + 0,442 + 0,53}{6} = 0,591 \quad \rightarrow \quad \bar{r} = 0,53$$

Die mittlere Korrelation zwischen den messwiederholten Faktorstufen ist $r = 0,53$.

- g) Die Teststärke, einen mittleren Effekt für unabhängige Stichprobe in der Untersuchung zu finden, lag nur bei etwa 41%.

Quelle: <http://www.lehrbuch-psychologie.de/qm>

© Rasch, Friese, Hofmann & Naumann

- h) Trotz identischer Mittelwerte ist der resultierende F -Wert in der messwiederholten Analyse größer als in der nicht-messwiederholten Analyse ($F_{(3;27)} = 5,52$ vs. $F_{(3;36)} = 3,13$) und hat dementsprechend eine niedrigere Wahrscheinlichkeit unter der Nullhypothese, selbst unter Berücksichtigung der Korrektur der Freiheitsgrade ($p < 0,02$ vs. $p < 0,04$). Der Grund für die höhere Teststärke des messwiederholten Verfahrens liegt in der positiven Korrelation zwischen den wiederholten Messungen. Diese systematische Abhängigkeit zwischen den Messzeitpunkten lässt sich auf Unterschiede zwischen Versuchspersonen zurückführen, und kann als Personenvarianz einen Teil der Residualvarianz erklären, den die Varianzanalyse ohne Messwiederholung nicht erklären kann. So reduziert sich die Residualvarianz im Fall der Messwiederholung, und dies führt zu einer Erhöhung des F -Bruchs bei sonst gleichen Bedingungen.

Literatur

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NY: Erlbaum.