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, das der Zusammenhang zwischen wiederholten Messungen der Symptome relativ gering ist (r = 0,2).

a) Wie viele Patienten werden benötigt, um einen mittleren Effekt einer therapeutischen Intervention mit 80%iger Sicherheit zu entdecken, unabhängig von der Art der Therapie?

b) Wie viele Patienten sollten teilnehmen, um mit der sehr hohen Sicherheit von 95% mittelgroße Unterschiede zwischen den Therapieformen in der Art ihres Therapieerfolgs zu entdecken, falls sie existieren?

c) 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	Approximierte s Chi-Quadrat	df	Signifikanz	Epsilon[a] Greenhous e-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

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

a) Welche Faktoren sind messwiederholt, welche nicht?

b) Wie viele Stufen haben die einzelnen Faktoren?

c) Für welche Faktoren tritt wahrscheinlich eine Verletzung einer Annahme der messwiederholten Varianzanalyse auf?

d) Warum kann für einen der Faktoren keine Verletzung der Sphärizitätsannahme auftreten?

e) Wie viele Versuchspersonen wurden untersucht?

f) Welches Korrekturverfahren der Freiheitsgrade sollte man für die Testung der einzelnen Faktoren und der Wechselwirkung verwenden und warum?

g) Welche Effekte sind signifikant?

h) Wie groß war die Teststärke, in dieser Untersuchung einen mittleren Effekt der Wechselwirkung zu finden, falls er existieren würde (r = 0,73)?

i) Ist das von SPSS angegebene Eta-Quadrat für den Faktor A mit einem großen Effekt aus einer Studie ohne Messwiederholung vergleichbar?

j) 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

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

Lösungen

Aufgabe 1

In GPower müssen Sie für die folgenden Berechnungen die Option „Other F-Tests“ auswählen. Da GPower hier nicht die Möglichkeit bietet, eine a priori Stichprobenumfangsplanung durchzuführen, müssen Sie die Anzahl der Versuchspersonen sowie die Anzahl der Fehlerfreiheitsgrade sukzessive anpassen. Die von GPower angegebenen Größen von kleinen, mittleren und großen Effekten sollten Sie ignorieren, da in der messwiederholten Varianzanalyse die Korrelation zwischen den Messzeitpunkten einen entscheidenden Einfluss auf die Größe des Effekts hat. Deshalb ist die direkte Orientierung an Konventionen von Effektgrößen im messwiederholten Fall nicht möglich, wir greifen stattdessen auf die Konventionen für unabhängige Stichproben zurück.

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-meswiederholten 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.

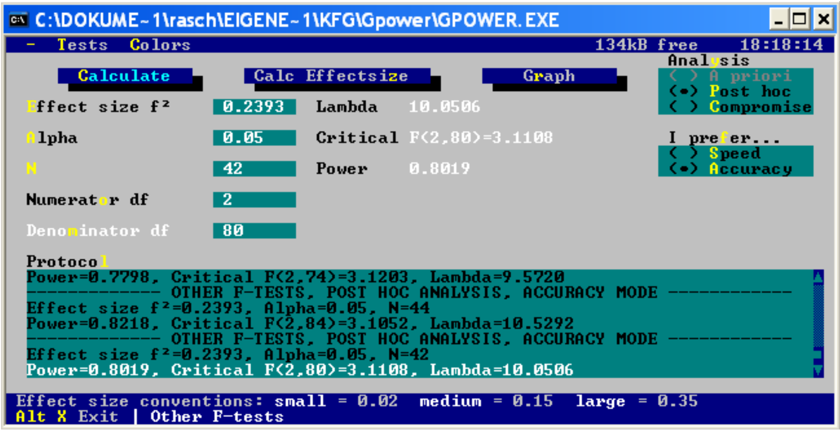
a) Für die Stichprobenumfangsplanung des Haupteffekts des messwiederholten Faktors müssen Sie zunächst die inhaltlich relevante Effektgröße für unabhängige Stichproben ($\Omega^2 = 0{,}06$ bzw. $\Phi^2 = 0{,}0638$) in das von GPower verlangte f²’ umrechnen.

$$f'^2 = \frac{p}{1-\bar{r}} \cdot \Phi^2 = \frac{3}{1-0{,}2} \cdot 0{,}0638 = 0{,}2393$$

Die Zählerfreiheitsgrade sind $df_A = p - 1 = 3 - 1 = 2$.

Die Nennerfreiheitsgrade für den messwiederholten Faktor berechnen sich aus $df_{A\times Vpn} = q \cdot (p-1) \cdot (n-1)$ und müssen in GPower sukzessive mit der Versuchspersonenzahl angepasst werden. (Achtung: Die Gesamtversuchspersonenzahl N ergibt sich aus dem Produkt der Anzahl der Stufen des nicht-messwiederholten Faktors und der Anzahl n der Personen in einer dieser Stufen ($N = q \cdot n$). Die Anzahl der Stufen p des messwiederholten Faktors spielen hier keine Rolle, weil dieselben Versuchspersonen mehrere Werte abgeben. Die Nennerfreiheitsgrade für den messwiederholten Faktor sind hier größer als die Gesamtanzahl an Versuchspersonen. Als Berechnungshilfe kann folgende Überlegung dienen: Die Nennerfreiheitsgrade sind um $q \cdot (p-1)$ kleiner als das $(p-1)$fache der Versuchspersonenzahl. Für das Beispiel: $df_{Res} = N \cdot 2 - 4$.

Nach Eingabe der Daten in GPower ergibt sich eine optimale Patientenzahl von insgesamt 42 (!), um einen mittleren Effekt (Konvention für unabhängige Stichproben) 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 21 Patienten teilnehmen.



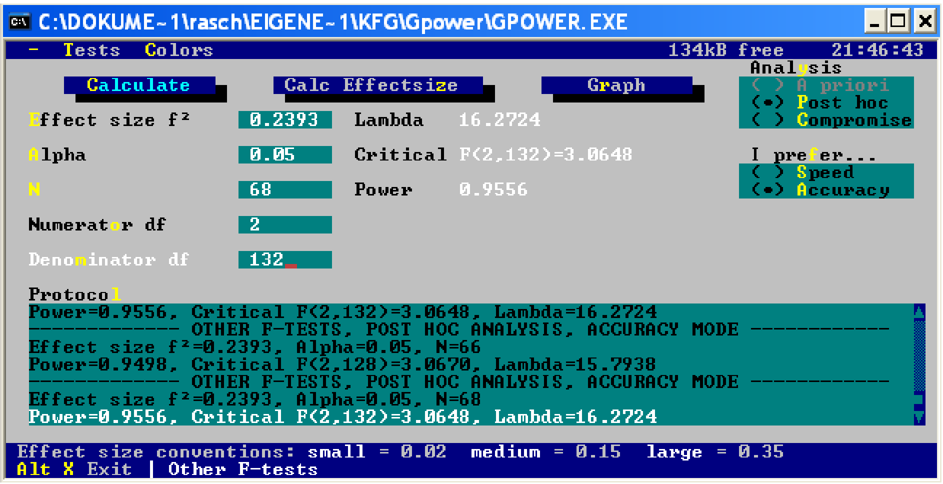
b) Hier ist nach einer Stichprobenumfangsplanung der Wechselwirkung zwischen den verschiedenen Messzeitpunkten und der Therapieform gefragt. Die Formel für die Umrechnung eines mittelgroßen Effekts für unabhängige Stichproben ($\Omega^2 = 0{,}06$ bzw. $\Phi^2 = 0{,}0638$) in das von GPower verlangte f'^2 ist mit der Formel für den messwiederholten Haupteffekt identisch ($r = 0{,}2$).

$$f'^2 = \frac{p}{1-\bar{r}} \cdot \Phi^2 = \frac{3}{1-0{,}2} \cdot 0{,}0638 = 0{,}2393$$

Die Zählerfreiheitsgrade der Wechselwirkung ergeben sich zu $df_{A\times B} = (p-1)\cdot(q-1) = 2$.

Die Nennerfreiheitsgrade sind für die Wechselwirkung wie auch für den messwiederholten Haupteffekt A $df_{A\times Vpn} = q\cdot(p-1)\cdot(n-1)$.

Durch sukzessives Anpassen der Anzahl der Versuchspersonen und der Nennerfreiheitsgrade 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.



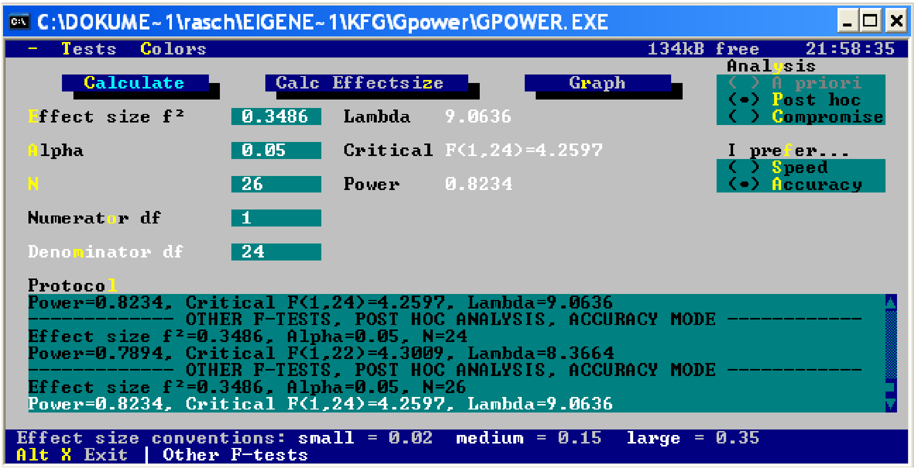
c) In der Stichprobenumfangsplanung für den nicht-messwiederholten Faktor B macht sich die Korrelation zwischen den Messzeitpunkten negativ bemerkbar: Je höher die Korrelation, desto mehr Versuchspersonen werden benötigt. In der Frage wird ein großer Effekt für unabhängige Stichproben ($\Omega^2 = 0{,}14$ bzw. $f^2 = 0{,}1627$) als inhaltlich relevant erachtet und wiederum eine Korrelation von r = 0,2 zwischen den Messzeitpunkten angenommen. Das von GPower geforderte f'² berechnet sich für den nicht-messwiederholten Faktor wie folgt:

$$f'^2 = \frac{p}{1+(p-1)\cdot\bar{r}}\cdot\Phi^2_{\text{unabhängig}} = \frac{3}{1+(3-1)\cdot 0{,}2}\cdot 0{,}1627 = \frac{3}{1{,}4}\cdot 0{,}1627 = 0{,}3486$$

Zählerfreiheitsgrade: $df_B = (q-1) = 1$

Nennerfreiheitsgrade: $df_{Vpn\ in\ S} = q \cdot (n-1)$

Es werden mindestens 26 Patienten (13 pro Gruppe) benötigt, um einen großen Effekt (für unabhängige Stichproben) bei einer mittleren Korrelation von r = 0,2 mit mindestens 80%iger Sicherheit auszuschließen, falls ein nicht signifikantes Ergebnis auftritt.



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 Sphärzität 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) Für den messwiederholten Faktor B kann keine Verletzung der Sphärizität auftreten, da dieser Faktor nur zwei Stufen hat. 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.

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

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

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

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

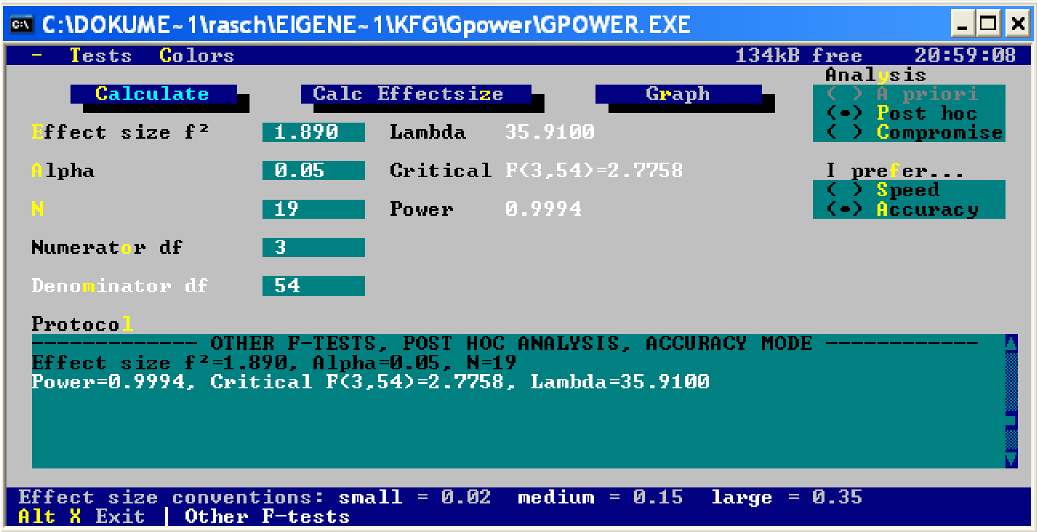
f) Für den Faktor B ist keine Korrektur der Freiheitsgrade notwendig, da keine Verletzung der Sphärizität auftreten kann (siehe Aufgabe d). Für die Wechselwirkung sollte das mit Greenhouse-Geisser bezeichnete Korrekturverfahren verwendet werden, da hier eine starke Verletzung der Sphärizität vorliegt ($\varepsilon_{\text{Greenhouse-Geisser}} < 0{,}75$). Für den Haupteffekt A könnte auch die liberalere Freiheitsgradkorrektur nach Hyundt-Feldt gewählt werden ($\varepsilon_{\text{Greenhouse-Geisser}} > 0{,}75$). 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) Berechnung des von GPower verlangten f'² für einen mittleren Effekt der Wechselwirkung für unabhängige Stichproben ($\Omega^2 = 0{,}06$ bzw. $\Phi^2 = 0{,}0638$) und einer mittleren Korrelation zwischen den Bedingungskombinationen von r = 0,73:

$$f'^2 = \frac{p\cdot q}{1-\bar{r}}\cdot\Phi^2 = \frac{4\cdot 2}{1-0{,}73}\cdot 0{,}0638 = 1{,}890$$

Einsetzen in GPower (N = 19; $df_{A\times B} = (p - 1) \cdot (q - 1) = 3$; $df_{A\times B\times Vpn} = (p - 1) \cdot (q - 1) \cdot (n - 1) = 54$):



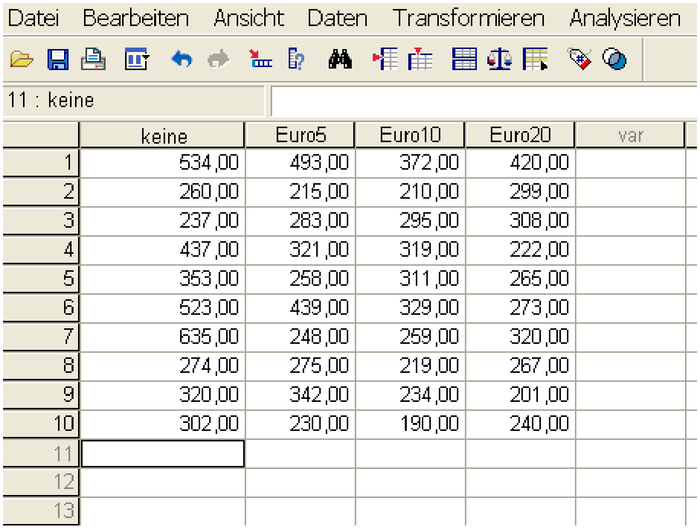
Die Teststärke, einen mittleren Effekt der Wechselwirkung (für unabhängige Stichproben) in der Untersuchung zu finden, war ausgesprochen hoch ($1 - \beta = 0{,}9994$). Die Nullhypothese, dass kein mittlerer Effekt (für unabhängige Stichproben) existiert, kann also mit sehr hoher Sicherheit angenommen werden.

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 Konvention 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:

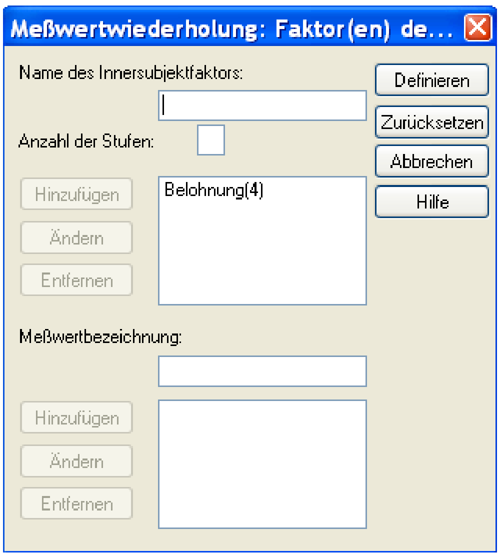
a) Für die Eingabe der Daten müssen Sie in SPSS vier neue Variable benennen. Eine Versuchsperson gibt in dieser Untersuchung vier Werte ab, die alle in derselben Zeile stehen müssen, um die Messwiederholung deutlich zu machen. Nach der Eingabe sollte Ihr SPSS Datenfenster in etwa wie folgt aussehen:

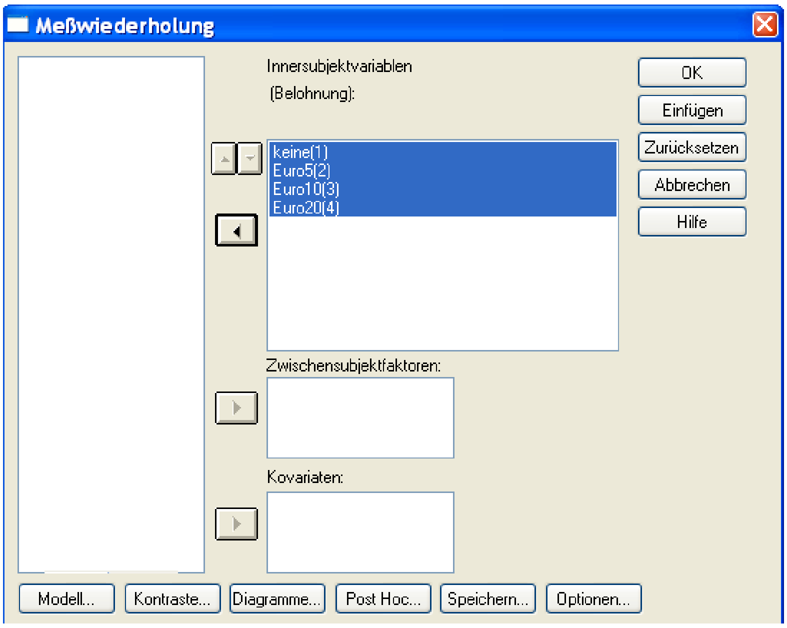


	keine	Euro5	Euro10	Euro20	var
1	534,00	493,00	372,00	420,00	
2	260,00	215,00	210,00	299,00	
3	237,00	283,00	295,00	308,00	
4	437,00	321,00	319,00	222,00	
5	353,00	258,00	311,00	265,00	
6	523,00	439,00	329,00	273,00	
7	635,00	248,00	259,00	320,00	
8	274,00	275,00	219,00	267,00	
9	320,00	342,00	234,00	201,00	
10	302,00	230,00	190,00	240,00	
11					
12					
13					

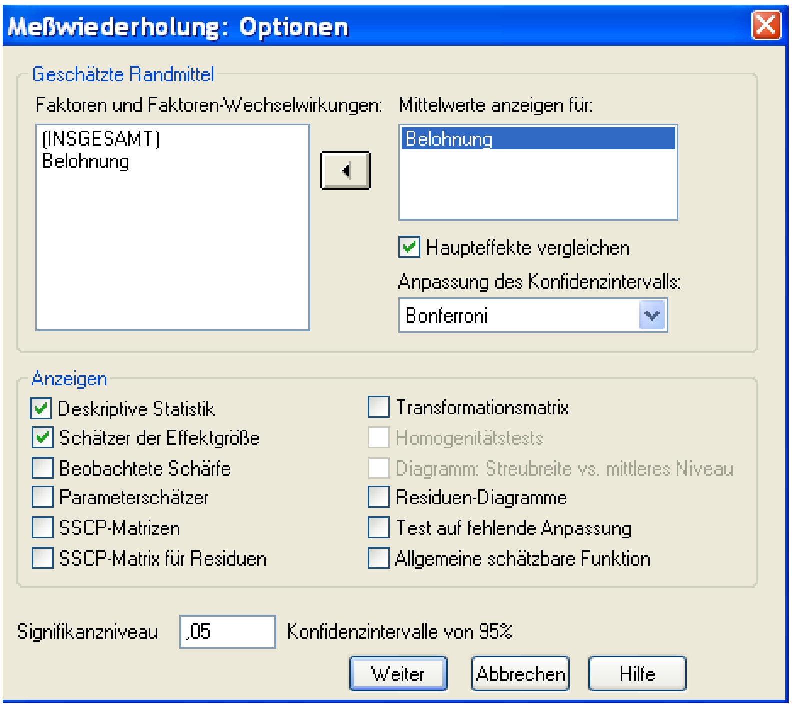
b) Zur Berechnung einer einfaktoriellen Varianzanalyse mit Messwiederholung gehen Sie auf „Analysieren“ → „Allgemeines Lineares Modell“ → „Messwiederholung“. Geben Sie im folgenden Fenster einen Namen für den messwiederholten Faktor und die Anzahl an Stufen ein.

Fahren Sie fort mittels „Definieren“. Geben Sie hier die vier Variablen der Messzeitpunkte ein (Reihenfolge beachten).

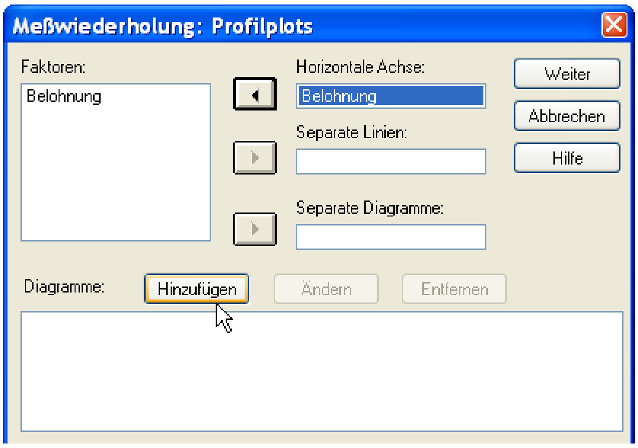




Unter „Optionen“ bewegen Sie den Faktor in das Feld „Mittelwerte anzeigen für“, aktiveren das Kästchen „Haupteffekte vergleichen“ und wählen „Bonferroni“ aus. Aktiveren Sie zusätzlich „Deskriptive Statistik“ und „Schätzer der Effektgröße“.



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



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

Innersubjektfaktoren

Maß: MASS_1

Belohnung	Abhängige Variable
1	keine
2	Euro5
3	Euro10
4	Euro20

Deskriptive Statistiken

	Mittelwert	Standardabweichung	N
keine	387,5000	136,77821	10
Euro5	310,4000	91,40897	10
Euro10	273,8000	60,06071	10
Euro20	281,5000	61,43696	10

Multivariate Tests[b]

Effekt		Wert	F	Hypothese df	Fehler df	Signifikanz	Partielles Eta-Quadrat
Belohnung	Pillai-Spur	,563	3,010[a]	3,000	7,000	,104	,563
	Wilks-Lambda	,437	3,010[a]	3,000	7,000	,104	,563
	Hotelling-Spur	1,290	3,010[a]	3,000	7,000	,104	,563
	Größte charakteristische Wurzel nach Roy	1,290	3,010[a]	3,000	7,000	,104	,563

a. Exakte Statistik

b.
Design: Intercept
Innersubjekt-Design: Belohnung

Mauchly-Test auf Sphärizität[b]

Maß: MASS_1

Innersubjekteffekt	Mauchly-W	Approximierte s Chi-Quadrat	df	Signifikanz	Epsilon[a] Greenhouse-Geisser	Huynh-Feldt	Untergrenze
Belohnung	,381	7,445	5	,193	,654	,835	,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.

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 Spärizität 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: MASS_1

Quelle		Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Signifikanz	Partielles Eta-Quadrat
Belohnung	Sphärizität angenommen	80855,400	3	26951,800	5,523	,004	,380
	Greenhouse-Geisser	80855,400	1,962	41210,110	5,523	,014	,380
	Huynh-Feldt	80855,400	2,504	32295,624	5,523	,008	,380
	Untergrenze	80855,400	1,000	80855,400	5,523	,043	,380
Fehler(Belohnung)	Sphärizität angenommen	131753,100	27	4879,744			
	Greenhouse-Geisser	131753,100	17,658	7461,275			
	Huynh-Feldt	131753,100	22,532	5847,268			
	Untergrenze	131753,100	9,000	14639,233			

Der Effekt der Höhe der Belohnung ist signifikant ($F_{3;27} = 5{,}52$; $p < 0{,}02$). SPSS gibt den Schätzer 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	Standardabweichung	N
keine	387,5000	136,77821	10
Euro5	310,4000	91,40897	10
Euro10	273,8000	60,06071	10
Euro20	281,5000	61,43696	10

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: MASS_1

(I) Belohnung	(J) Belohnung	Mittlere Differenz (I-J)	Standardfehler	Signifikanz[a]	95% Konfidenzintervall für die 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

Basiert auf den geschätzten Randmitteln

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 Sphärizitätsannahme 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 I).

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

Deskriptive Statistiken

	Mittelwert	Standardabweichung	N
keine	387,5000	136,77821	10
Euro5	310,4000	91,40897	10
Euro10	273,8000	60,06071	10
Euro20	281,5000	61,43696	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.

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

Korrelationen

		keine	Euro5	Euro10	Euro20
keine	Korrelation nach Pearson	1	,499	,543	,406
	Signifikanz (2-seitig)		,142	,105	,244
	N	10	10	10	10
Euro5	Korrelation nach Pearson	,499	1	,757*	,414
	Signifikanz (2-seitig)	,142		,011	,234
	N	10	10	10	10
Euro10	Korrelation nach Pearson	,543	,757*	1	,487
	Signifikanz (2-seitig)	,105	,011		,153
	N	10	10	10	10
Euro20	Korrelation nach Pearson	,406	,414	,487	1
	Signifikanz (2-seitig)	,244	,234	,153	
	N	10	10	10	10

*. Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

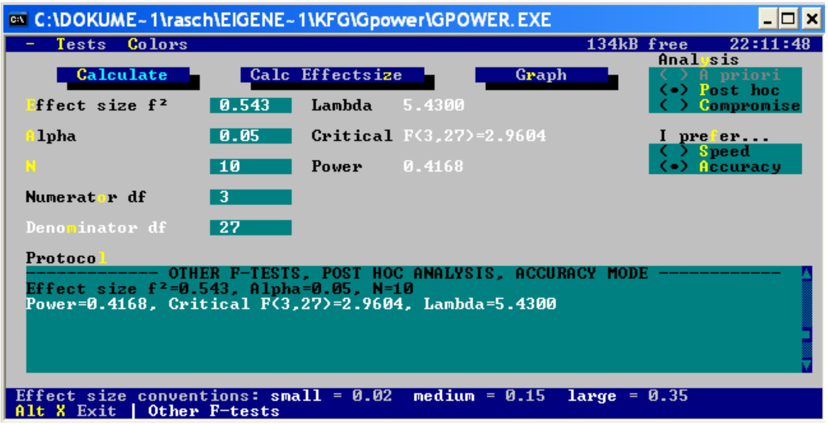
Nach der Umwandlung der Korrelationen in Fishers Z-Werte ergibt sich folgender mittlerer Fishers Z-Wert (siehe Tabelle D, Band I).

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

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

f) Berechnen Sie zuerst das von GPower verlangte f'² für einen mittleren Effekt für unabhängige Stichproben ($\Omega^2 = 0{,}06$ bzw. $\Phi^2 = 0{,}0638$) und eine Korrelation von r = 0,53.

$$f'^2 = \frac{p}{1-\overline{r}} \cdot \Phi^2 = \frac{4}{1-0{,}53} \cdot 0{,}0638 = 0{,}543$$



Die Teststärke, einen mittleren Effekt für unabhängige Stichprobe in der Untersuchung zu finden, war nur 41,68%.

g) 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 niedriger 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.