Lösungen zu den Übungsaufgaben in Kapitel 14

(1) In den Übungsaufgaben zu Kapitel 13 hatten Sie bereits mit der Datei »daten_kap13.txt« gearbeitet. Hier handelte es sich um ein fiktives Vier-Gruppen-Experiment. Stellen Sie sich vor, die vier Bedingungen (es handelte sich um eines von vier Musikstücken, die die Versuchspersonen während der Bearbeitung eines Konzentrationstests anhören mussten) seien nicht interindividuell, sondern vielmehr intraindividuell variiert worden. Die Versuchspersonen müssen also nacheinander vier Musikstücke hören und bearbeiten währenddessen je einen Konzentrationstest. Die erreichte Punktzahl in diesem Test fungiert als abhängige Variable.

Lesen Sie diese Datendatei in ein Statistikprogramm (z. B. R, SPSS) oder ein Tabellenkalkulationsprogramm (z. B. Excel) ein. Hier müssen Sie darauf achten, dass das Eingabeformat bei messwiederholten Designs anders ist als bei nicht-messwiederholten Designs: Die Stufen des messwiederholten Faktors stehen nun in unterschiedlichen Spalten und nicht mehr in unterschiedlichen Zeilen. Beantworten Sie die folgenden Fragen:

(a) Prüfen Sie, ob die Voraussetzungen für die Anwendung des *F*-Tests für die einfaktorielle Varianzanalyse mit Messwiederholung erfüllt sind.

Die Voraussetzungen für die Anwendung des *F*-Tests bei der einfaktoriellen Varianzanalyse mit Messwiederholung lauten wie folgt:

(1) Die zufälligen Personeffekte π_m sind unabhängig und identisch normal verteilt mit $N(0, \sigma_\pi^2)$.

(2) Die Residuen ε_{mj} sind unabhängig und identisch normal verteilt mit $N(0, \sigma_\varepsilon^2)$.

(3) Die Kovarianz der Personeffekte und der Residuen ist gleich 0: $\text{Cov}(\pi_m, \varepsilon_{mj}) = 0$.

Aus diesen Annahmen folgt, dass die Varianz von X_{mj} der Varianz $\sigma_{X_j}^2$ entspricht mit $\sigma_{X_j}^2 = \sigma_\pi^2 + \sigma_\varepsilon^2$.

Die Kovarianzmatrix hat für vier Bedingungen also die folgende Form:

$$\mathbf{\Sigma}_X = \begin{pmatrix} \sigma_\pi^2 + \sigma_\varepsilon^2 & \sigma_\pi^2 & \sigma_\pi^2 & \sigma_\pi^2 \\ \sigma_\pi^2 & \sigma_\pi^2 + \sigma_\varepsilon^2 & \sigma_\pi^2 & \sigma_\pi^2 \\ \sigma_\pi^2 & \sigma_\pi^2 & \sigma_\pi^2 + \sigma_\varepsilon^2 & \sigma_\pi^2 \\ \sigma_\pi^2 & \sigma_\pi^2 & \sigma_\pi^2 & \sigma_\pi^2 + \sigma_\varepsilon^2 \end{pmatrix}$$

Der *F*-Test kommt aber auch dann noch zu robusten Ergebnissen, wenn die Matrix der Varianzen und Kovarianzen nicht die hier dargestellte compound symmetry-Form aufweist, sondern lediglich sphärisch (oder zirkulär) ist. D. h. die Varianzen der sechs möglichen Differenzvariablen, die hier gebildet werden können, müssen auf Populationsebene identisch sein. Werten wir die Daten mit SPSS aus, erhalten wir die folgende Tabelle für die Sphärizitätsdiagnostik:

Mauchly-Test auf Sphärizität[b]

Maß:MASS_1

Innersubjekt-effekt	Mauchly-*W*	Approximiertes Chi-Quadrat	*df*	Sig.	Epsilon[a]		
					Greenhouse-Geisser	Huynh-Feldt	Untergrenze
Bedingung	,945	1,009	5	,962	,965	1,000	,333

Im Allgemeinen gilt die Sphärizitätsannahme als verletzt, wenn $\hat{\varepsilon}_{GG}$ kleiner ist als 0,75. In unserem Fall ist $\hat{\varepsilon}_{GG} = 0{,}965$. Darüber hinaus zeigt der Mauchly-Test auf Sphärizität an, dass die Abweichung der

Kovarianzmatrix von einer sphärischen Matrix nicht signifikant ist (p = 0,962). Wir können also davon ausgehen, dass die Sphärizitätsannahme hier nicht verletzt ist.

(b) Prüfen Sie mit Hilfe einer einfaktoriellen Varianzanalyse mit Messwiederholung, ob sich die vier Bedingungsmittelwerte signifikant voneinander unterscheiden. Der *F*-Test wird auf einem Signifikanzniveau von α = 5 % durchgeführt.

Die Ergebnisse sind in der folgenden Tafel der Varianzanalyse zusammengestellt.

Quelle der Variation	*QS*	*df*	*MQS*	*F*	*p*
Faktor *A*	2003,45	3	667,82	53,42	<0,001
Person	681,55	19	35,87		
Residuum	712,55	57	12,50		
Total	3397,55	79			

Die Mittelwerte in den vier Bedingungen unterscheiden sich also signifikant voneinander; die Nullhypothese kann verworfen werden.

(c) Wie lauten die empirische Effektstärken $\hat{\eta}^2$, $\hat{\eta}_p^2$ und $\hat{\eta}'^2$? Bewerten Sie die Größe von $\hat{\eta}'^2$ anhand der Taxonomie von Cohen. Wo liegen die Grenzen des zweiseitigen 90 %-Konfidenzintervalls für die Effektgröße η'^2?

- Effektgröße $\hat{\eta}^2$ (laut Formel F 14.8): $\hat{\eta}^2 = \dfrac{QS_{\text{zw}}}{QS_{\text{tot}}} = \dfrac{2003,45}{3397,55} = 0,60$
- Effektgröße $\hat{\eta}_{\text{p}}^2$ (laut Formel F 14.9): $\hat{\eta}_{\text{p}}^2 = \dfrac{QS_{\text{zw}A}}{QS_{\text{zw}A} + QS_{\text{Res}}} = \dfrac{2003,45}{2003,45 + 712,55} = 0,738$

Um die Effektgröße $\hat{\eta}'^2$ gemäß Formel F 14.48 schätzen zu können, benötigen wir zunächst eine Schätzung für die Intraklassen-Korrelation (siehe Formel F 14.46). Wir erhalten:

$\hat{\rho} = \dfrac{MQS_{\text{zwP}} - MQS_{\text{Res}}}{MQS_{\text{zwP}} + (J-1)\cdot MQS_{\text{Res}}} = \dfrac{35,87 - 12,5}{35,87 + 3\cdot 12,5} = 0,32$. Diesen Wert setzen wir in Formel F 14.47 ein und erhalten $f'^2 = \dfrac{\hat{\eta}^2}{(1-\hat{\rho})\cdot(1-\hat{\eta}^2)} = \dfrac{0,60}{0,68\cdot 0,40} = 2,206$. Diesen Wert setzen wir schließlich in Formel F 14.48 ein und erhalten $\hat{\eta}'^2 = f'^2 / (1 + f'^2) = 2,206/3,206 = 0,688$. Es handelt sich um einen sehr großen Effekt (vgl. Tabelle 14.3).

Bestimmung der Grenzen des zweiseitigen 90 %-Konfidenzintervalls für die Effektgröße η'^2: Mit Hilfe des Programms NDC bestimmen wir auf der Basis des empirischen *F*-Wertes (F = 53,42; df_1 = 3; df_2 = 57) folgende Unter- bzw. Obergrenze des zweiseitigen 90 %-Konfidenzintervalls für den Nonzentralitätsparameter: λ'_{u} = 99,6 und λ'_{o} = 228,08. Diese beiden Werte können nach Formel F 14.50 in die Metrik des Effektstärkenmaßes ϕ'^2 zurücktransformiert werden. So ergibt sich als Untergrenze des Intervalls $f'^2_{\text{u}} = \lambda'_{\text{u}}/(n\cdot J) = 99,6/80 = 1,245$ und als Obergrenze $f^2_{\text{o}} = \lambda'_{\text{o}}/(n\cdot J) = 228,08/80 = 2,851$. Rechnet man nun mit Hilfe von Formel F 14.48 diese Werte in die Metrik des Effektstärkenmaßes η'^2 um, erhält man als Untergrenze den Wert

$\eta_u'^2 = \phi_u'^2 / (1 + \phi_u'^2) = 1,245/2,245 = 0,555$ und als Obergrenze den Wert $\eta_o'^2 = \phi_o'^2 / (1 + \phi_o'^2) = 2,851/3,851 = 0,74$ Der wahre Populationseffekt η'^2 wird also mit einer Wahrscheinlichkeit von 90 % von einem Intervall mit den Grenzen [0,555; 0,74] überdeckt.

(d) Wie viele Versuchspersonen wären — bei einem Signifikanzniveau von α = 5 % — nötig gewesen, um einen kleinen Effekt (nach der Taxonomie von Cohen) mit einer Wahrscheinlichkeit von 1 — β = 95 % zu finden, falls es ihn gibt? Verwenden Sie für die Intraklassen-Korrelation ρ den aus den Daten geschätzten Wert.

Zur Berechnung des optimalen Stichprobenumfangs nehmen wir das Programm G*Power zu Hilfe. G*Power verwendet als Effektstärkenmaß ϕ (bzw. ϕ_1, also die positive Quadratwurzel aus ϕ_1^2). Wir geben in dem Programm das unkorrigierte Effektstärkenmaß ϕ_1 an, da die Intraklassen-Korrelation gesondert eingegeben werden muss und das Programm die Umrechnung in ϕ_1' selbst vornimmt. Ein kleiner Effekt würde nach Cohen einem Wert von $\phi_1 = 0,1$ entsprechen. Die in G*Power einzugebenden Werte lauten also $\phi_1 = 0,1$; $\alpha = 0,05$; $1 - \beta = 0,95$; »number of groups« = 1; »repetitions« (das entspricht der Anzahl der Stufen des messwiederholten Faktors, also J) = 4; »corr among repeated measures« (das entspricht der Intraklassen Korrelation) $\hat{\rho} =$ 0,32 und »nonshericity correction« (das entspricht dem Korrekturfaktor ε) = 1 (da wir hier der Einfachheit halber von einer perfekt sphärischen Matrix ausgehen). G*Power ermittelt eine optimale Stichprobengröße von $n = 294$.

(2) Führen Sie mit dem Datenbeispiel aus der Datei »daten_kap13.txt« einen Friedman-Test durch. Zu welchem Ergebnis kommen Sie?

Die vier bedingungsspezifischen Rangsummen lauten $RS_{\cdot 1} = 78$; $RS_{\cdot 2} = 52,5$; $RS_{\cdot 3} = 46$ und $RS_{\cdot 4} = 23,5$. Daraus lässt sich der Wert der Prüfgröße S nach Formel F 14.111 ermitteln; er beträgt $s = 1508,8$. Rechnet man diesen Wert in den Wert der Prüfgröße Q um (nach Formel F 14.114), welche approximativ χ^2-verteilt ist, erhält man $Q = 46,18$. Der kritische χ^2-Wert beträgt für ein Signifikanzniveau von α = 5 % (zweiseitiger Test) bei $df = J - 1 = 3$ Freiheitsgraden $\chi^2_{(0,975;3)} = 9,35$. Unser empirischer Q-Wert ist größer als der kritische Wert, die Nullhypothese kann also verworfen werden.

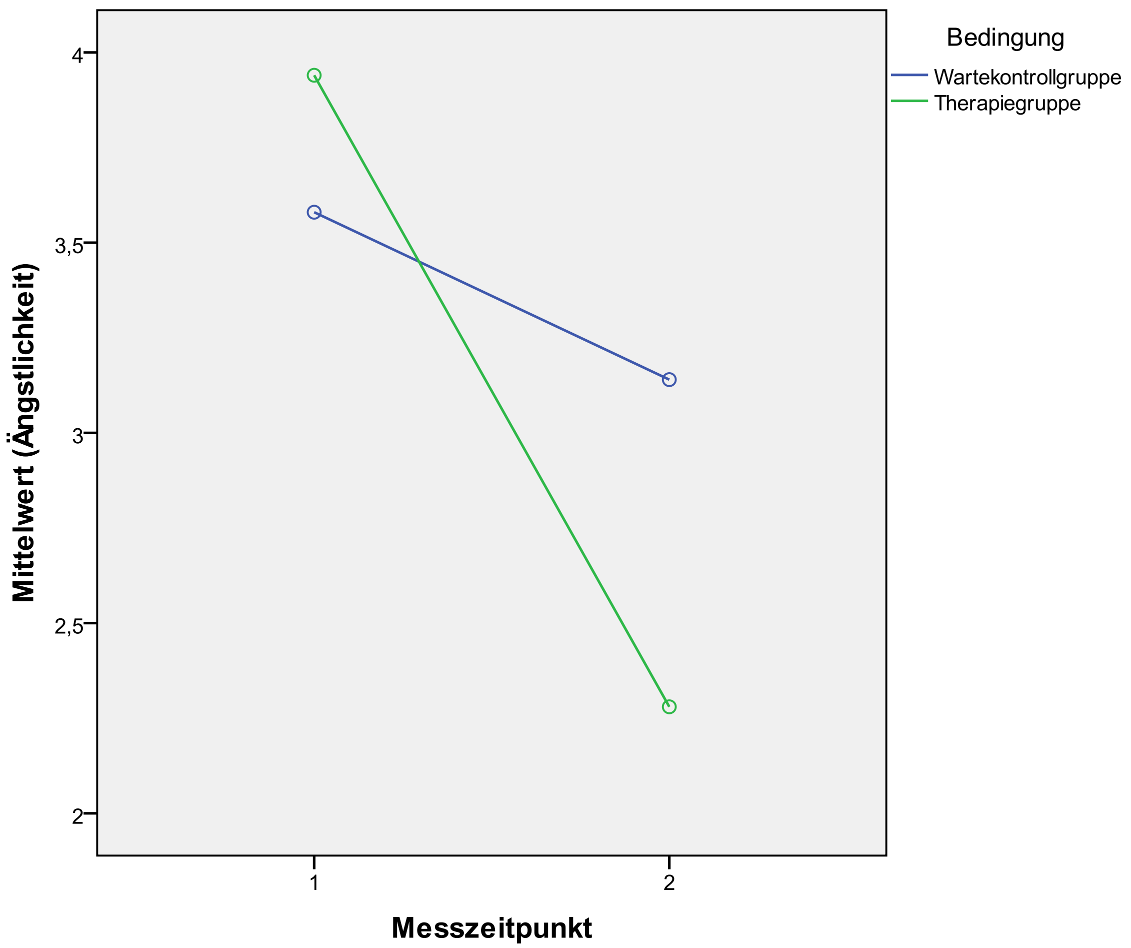
(3) Es soll die Wirksamkeit einer neuen therapeutischen Maßnahme zur Verringerung von Ängstlichkeit überprüft werden. Hierzu werden 100 Patientinnen und Patienten mit pathologischer Ängstlichkeit per Zufall einer Therapiebedingung oder einer Wartekontrollbedingung zugewiesen (Faktor *A*). Die Ängstlichkeit wird auf der Basis eines normierten Verfahrens (stetige Variable) zu zwei Messzeitpunkten erfasst: unmittelbar vor Beginn der therapeutischen Maßnahme und unmittelbar nach Beendigung der Maßnahme (Faktor *B*; messwiederholt). Die Rohdaten für diese Untersuchung finden Sie in der Datei »daten_kap14.txt«.

(a) Testen Sie mit Hilfe einer zweifaktoriellen Varianzanalyse mit Messwiederholung auf einem Faktor, ob es (1) Haupteffekte des Faktors *A*, (2) Haupteffekte des Faktors *B* und (3) Interaktionseffekte *A* × *B* gibt. Wie lassen sich die Ergebnisse der drei *F*-Tests inhaltlich interpretieren?

Die Ergebnisse sind in der folgenden Tafel der Varianzanalyse dargestellt.

Quelle der Variation	***QS***	***df***	***MQS***	***F***	***p***
Haupteffekte von *A* (Bedingung)	3,125	1	3,125	0,292	0,59
Person innerhalb *A*	1050,33	98	10,718		
Haupteffekte von *B* (Messzeitpunkt)	55,125	1	55,125	15,853	<0,001
Interaktionseffekte $A \times B$	18,605	1	18,605	5,351	0,02
Residuum	340,77	98	3,477		
Total	1467,955	199			

Die inhaltliche Interpretation lässt sich gut anhand des folgenden Liniendiagramms erläutern. Demnach nimmt die Ängstlichkeit in der Therapiegruppe zwischen den beiden Messzeitpunkten stärker ab als in der Kontrollgruppe. Diese Interaktion zwischen Messzeitpunkt und Bedingung ist hypothesenkonform und spricht für die Wirksamkeit der Therapie.



(b) Schätzen Sie die Stärke des Effekts der Interaktion in der Metrik der Effektgrößenmaße $\eta^2_{A\times B}$ (nicht-partiell) und $\eta^2_{p\_A\times B}$ (partiell). Berechnen Sie zusätzlich die Effektgrößenmaße δ_{TK} und δ_V. Interpretieren Sie alle ermittelten Effektgrößenschätzer.

Nicht-partielles Effektstärkenmaß:

$$\eta^2_{A\times B} = QS_{A\times B} / (QS_{\text{tot}} - QS_A - QS_B) = 18{,}605/(1467{,}955 - 3{,}125 - 55{,}125) = 0{,}013$$

Partielles Effektstärkenmaß:

$$\eta^2_{\text{p\_}A\times B} = QS_{A\times B} \Big/ \left(QS_{A\times B} + QS_{\text{Res}}\right) = 18{,}605 \Big/ \left(18{,}605 + 340{,}77\right) = 0{,}052$$

Um das Effektstärkenmaß δ_{TK} zu berechnen, müssen wir zunächst die gepoolte Innerhalb-Standardabweichung zum ersten Messzeitpunkt nach Formel F 14.102 berechnen. Zuvor ermitteln wir die Innerhalb-Varianz zum ersten Messzeitpunkt in den beiden Bedingungen $\hat{\sigma}^2_{\text{T1}}$ = 6,792 und $\hat{\sigma}^2_{\text{K1}}$ = 8,698. Diese Werte setzen wir nun in Formel F 14.102 ein und erhalten

$$\hat{\sigma}_1 = \sqrt{\frac{\hat{\sigma}^2_{\text{Tg1}} \cdot \left(n_{\text{Tg1}} - 1\right) + \hat{\sigma}^2_{\text{Kg1}} \cdot \left(n_{\text{Kg1}} - 1\right)}{\left(n_{\text{Tg1}} - 1\right) + \left(n_{\text{Kg1}} - 1\right)}} = \sqrt{\frac{6{,}792 \cdot 49 + 8{,}698 \cdot 49}{49 + 49}} = 2{,}783$$

Setzen wir nun diesen Wert in Formel F 14.101 ein, erhalten wir

$$\hat{\delta}_{\text{TK}} = \frac{\bar{x}_{\text{Tg1}} - \bar{x}_{\text{Tg2}}}{\hat{\sigma}_1} - \frac{\bar{x}_{\text{Kg1}} - \bar{x}_{\text{Kg2}}}{\hat{\sigma}_1} = -0{,}60 - (-0{,}16) = -0{,}44$$

Dieser Wert bedeutet, dass die standardisierte Differenz in der Therapiegruppe um 0,44 Standardabweichungseinheiten extremer ausfiel als die standardisierte Differenz in der Kontrollgruppe.

Um das Effektstärkenmaß δ_{V} zu berechnen, müssen wir zunächst die gepoolte Standardabweichung der Differenzvariablen nach Formel F 14.106 berechnen. Zuvor ermitteln wir die Varianz der Differenzen in den beiden Bedingungen wir nun in Formel F 14.106 ein und erhalten $\hat{\sigma}^2_{\text{TD}}$ = 6,025 und $\hat{\sigma}^2_{\text{KD}}$ = 7,884.

Diese Werte setzen wir nun in Formel F 14.106 ein und erhalten

$$\hat{\sigma}_D = \sqrt{\frac{\hat{\sigma}^2_{\text{Tg}D} \cdot \left(n_{\text{Tg}} - 1\right) + \hat{\sigma}^2_{\text{Kg}D} \cdot \left(n_{\text{Kg}} - 1\right)}{\left(n_{\text{Tg}} - 1\right) + \left(n_{\text{Kg}} - 1\right)}} = \sqrt{\frac{6{,}025 \cdot 49 + 7{,}884 \cdot 49}{49 + 49}} = 2{,}637\,.$$

Setzen wir nun diesen Wert in Formel F 14.105 ein, erhalten wir

$$\hat{\delta}_{\text{V}} = \frac{\bar{x}_{\text{Tg1}} - \bar{x}_{\text{Tg2}}}{\hat{\sigma}_D} - \frac{\bar{x}_{\text{Kg1}} - \bar{x}_{\text{Kg2}}}{\hat{\sigma}_D} = -0{,}63 - \left(-0{,}17\right) = -0{,}46\,.$$

Dieser Wert bedeutet, dass der Veränderungseffekt in der Therapiegruppe um 0,46 Standardabweichungseinheiten extremer ausfiel als der Veränderungseffekt in der Kontrollgruppe.

(c) Wo liegen die Grenzen des zweiseitigen 90 %-Konfidenzintervalls für die Effektgröße $\eta'^2_{A\times B}$?

Mit Hilfe des Programms NDC bestimmen wir auf der Basis des empirischen *F*-Wertes (F = 5,351; df_1 = 1; df_2 = 98) folgende Unter- bzw. Obergrenze des zweiseitigen 90 %-Konfidenzintervalls für den Nonzentralitätsparameter: λ'_{u} = 0,386 und λ'_{o} = 15,797. Diese beiden Werte können mit Hilfe von Formel F 14.107 in die Metrik des Effektgrößenmaßes ϕ'^2 zurücktransformiert werden. Hierzu müssen wir diese Formel nach ϕ'^2 auflösen:

$$\phi_1'^2 = \frac{\lambda'}{n_{\text{Zelle}} \cdot J \cdot K}$$

In unserem Beispiel ergibt sich für die Untergrenze ein Wert von $\phi_u'^2 = 0{,}386/200 = 0{,}002$ und für die Obergrenze ein Wert von $\phi_o'^2 = 15{,}797/200 = 0{,}002 \; = 0{,}08$. Rechnet man nun mit Hilfe von Formel F 14.48 diese Werte in die Metrik des Effektstärkenmaßes η'^2 um, erhält man als Untergrenze den Wert $\eta_u'^2 = \phi_u'^2 / \left(1 + \phi_u'^2\right) = 0{,}002/1{,}002 = 0{,}002$ und als Obergrenze den Wert $\eta_o'^2 = \phi_o'^2/(1 + \phi_o'^2) = 0{,}08/1{,}08 = 0{,}074$. Der wahre Populationseffekt wird also mit einer Wahrscheinlichkeit von 90 % von einem Intervall mit den Grenzen [0,002; 0,074] überdeckt.

(d) Wie viele Personen wären nötig gewesen, um die Nullhypothese für die Interaktion *A × B* mit einer Wahrscheinlichkeit von 90 % abzulehnen, wenn der Test auf einem Signifikanzniveau von α = 5 % durchgeführt wird und der Effekt in der Population η_1^2 = 0,1 beträgt? Verwenden Sie für die Intraklassen-Korrelation ρ den Wert 0,70.

Zur Berechnung des optimalen Stichprobenumfangs nehmen wir das Programm G*Power zu Hilfe. G*Power verwendet als Effektstärkenmaß ϕ. Umgerechnet ergibt sich also ein Wert von $\phi_1 = \sqrt{\eta_1^2 / \left(1 - \eta_1^2\right)} = 0{,}333$. Die in G*Power einzugebenden Werte lauten also ϕ_1 = 0,333, α = 0,05; 1 – β = 0,90; »number of groups« = 2; »repetitions« = 2; »corr among repeated measures« = 0,70 und »nonshericity correction« = 1 (da wir hier der Einfachheit halber von einer perfekt sphärischen Matrix ausgehen). G*Power ermittelt eine optimale Gesamtstichprobengröße von n = 18.