Aufgaben zu Kapitel 3

Aufgabe 1

a) Berechnen Sie einen t-Test für unabhängige Stichproben für den Vergleich der beiden Verarbeitungsgruppen „strukturell" und „emotional" für die abhängige Variable „neutrale Adjektive".

b) Berechnen Sie mit GPower den empirischen Effekt.

Aufgabe 2

Ein Forscher plant eine Untersuchung, bei der er nur einen kleinen Effekt zwischen den beiden unabhängigen Versuchsgruppen erwartet. Er setzt das α-Niveau auf 5% und strebt einen β-Fehler an, der nicht mehr als doppelt so groß sein soll. Wie viele Versuchspersonen braucht er, um seine gerichtete Annahme in jedem Fall interpretieren zu können? Wie viele wären es, wäre seine Hypothese ungerichtet gewesen?

Aufgabe 3

Ein anderer Forscher glaubt, dass es zwischen zwei unabhängigen Versuchsbedingungen keinen Unterschied gibt. Er möchte also die Nullhypothese interpretieren. Dafür muss er die Wahrscheinlichkeit klein halten, sich zu irren, also zu unrecht die H_0 anzunehmen, obwohl in Wirklichkeit die H_1 gilt. Statt dessen ist er etwas liberaler darin, sich evtl. zu Gunsten der Alternativhypothese zu irren. Er legt die entsprechende Fehlerwahrscheinlichkeit auf 10% fest. Wie viele Versuchspersonen braucht der Forscher, um hinreichend sicher sein zu können, dass nicht zumindest doch ein kleiner Effekt zwischen den Gruppen existiert?

Aufgabe 4

Eine Studie führt zu einem nicht signifikanten Ergebnis, obwohl in jeder der beiden unabhängigen Gruppen 100 Personen teilgenommen haben. Wie groß war die Power bei einer ungerichteten Fragestellung, einem α-Niveau von 5% und einem für inhaltlich relevant erachteten Effekt von $d = 0{,}50$?

Aufgabe 5

a) Prüfen Sie mit einem t-Test für abhängige Stichproben, ob sich im Experiment zu motorischen Fähigkeiten die Anzahl beendeter Sequenzen zwischen dem ersten und dem dritten Messzeitpunkt unterscheiden ($\alpha = 5\%$, Datei „Datensatz_Messwiederholung.sav").

b) Wie groß war der empirische Effekt d_z?

c) Wie groß war die Teststärke, einen Unterschied von $d_{unabhängig} = 0{,}5$ zu finden?

d) Wie viele Versuchspersonen wären erforderlich gewesen, um einen $\alpha = \beta$-Fehler von 5% zu erreichen bei einem angenommenen Effekt von $d_{unabhängig} = 0{,}5$, wenn wir vor der Untersuchung nur von einer Korrelation von $r = 0{,}30$ zwischen den Messwertreihen ausgegangen wären?

e) Auf welche Zahl hätte sich dieser Bedarf erhöht, wenn die Fragestellung mit unabhängigen Stichproben untersucht worden wäre?

Lösungen

a) „Analysieren“ → „Mittelwerte vergleichen“ → „t-Test bei unabhängigen Stichproben“. Testvariable ist „erinnerte neutrale Adjektive“, Gruppenvariable „Verarbeitungsbedingung“ und davon die beiden Stufen „1“ (strukturell) und „3“ (emotional).

Die Analyse liefert den folgenden Output:

Gruppenstatistiken

	Verarbeitungsbedingung	N	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes
erinnerte neutrale Adjektive	strukturell	50	1,96	1,428	,202
	emotional	50	4,06	2,064	,292

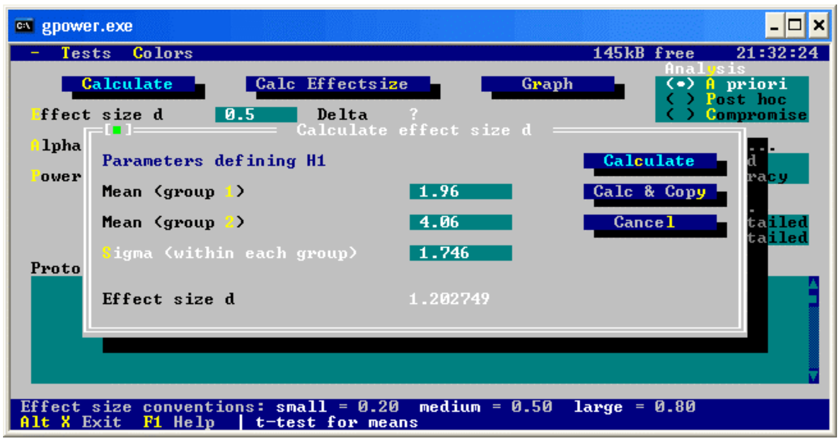
Test bei unabhängigen Stichproben

		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit						
									95% Konfidenzintervall der Differenz	
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	Untere	Obere
erinnerte neutrale Adjektive	Varianzen sind gleich	3,609	,060	-5,916	98	,000	-2,100	,355	-2,804	-1,396
	Varianzen sind nicht gleich			-5,916	87,156	,000	-2,100	,355	-2,806	-1,394

Der Mittelwertsvergleich mit bloßen Auge im Feld Gruppenstatistiken zeigt einen Unterschied zwischen den beiden Gruppen. Der Levene-Test auf Varianzhomogenität liefert ein marginal signifikantes Ergebnis. Dies drückt sich u.a. in den reduzierten Freiheitsgraden in der Zeile „Varianzen sind nicht gleich“ aus. Doch unabhängig von einer möglichen Korrektur der Freiheitsgrade wird der Unterschied zwischen den Gruppen hoch signifikant. Entsprechend der theoretischen Vorhersage erinnern Personen in der Verarbeitungsbedingung „emotional“ mehr Adjektive als Personen in der Bedingung „strukturell“.

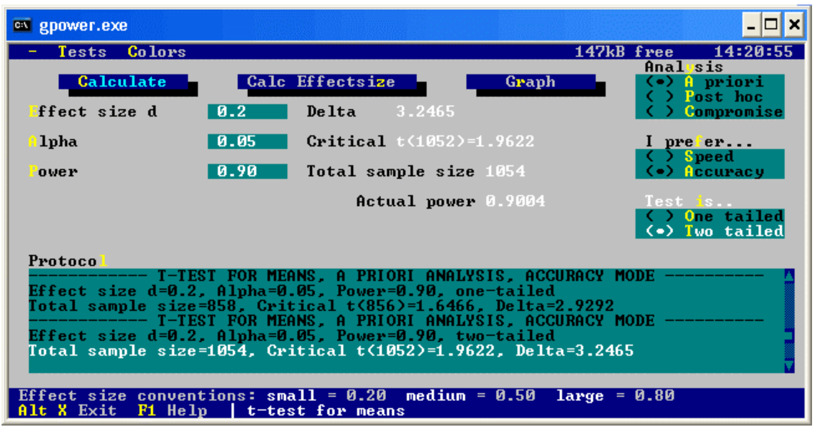
b) GPower startet mit den Optionen für einen t-Test für unabhängige Stichproben. Dort führt der Button „Calc Effectsize“ zur gesuchten Eingabemaske.

Für die Berechnung der Effektstärke benötigen wir gemäß der Definition von d (vgl. Kapitel 3.3.1) Angaben der zwei Stichprobenmittelwerte sowie der geschätzten Populationsstreuung. Die Mittelwerte entnehmen wir der Tabelle „Gruppenstatistiken“. Die Populationsstreuung schätzen wir durch den Mittelwert der beiden Stichprobenstreuungen, 1,746. Diese Werte liefern einen empirischen Effekt von d = 1,20. Dies ist nach den Konventionen von Cohen (1988) ein großer Effekt.



Aufgabe 2

Er braucht 858 Versuchspersonen, fast 430 Personen in jeder Zelle. Würde er eine ungerichtete Hypothese verfolgen, wären es sogar 1054 Personen insgesamt. Dieses Ergebnis macht deutlich, wie dramatisch die benötigte Anzahl Versuchspersonen bei kleinen angenommenen Effekten ansteigt, wenn eine hohe Teststärke erreicht werden soll.

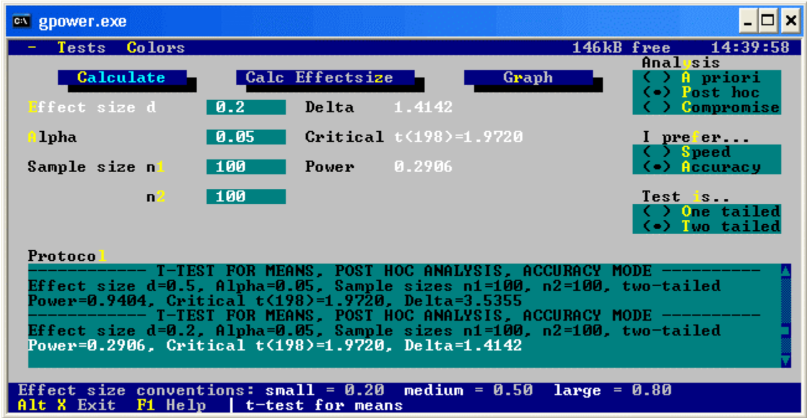


Aufgabe 3

Der Forscher möchte den β-Fehler minimieren, während er beim α-Fehler etwas liberaler ist. Entsprechend legt er den β-Fehler auf 5%, den α-Fehler auf 10% fest. Da ein Unterschied zwischen den Gruppen unabhängig vom Vorzeichen gegen seine Annahme sprechen würde, muss er einen zweiseitigen Test durchführen. Bei einem existierenden kleinen Effekt von d = 0,20 müsste das Ergebnis also bei 1302 Versuchspersonen signifikant werden. Wird es dies nicht, kann der Forscher mit 95%iger Sicherheit davon ausgehen, dass es keinen Effekt von d = 0,20 oder größer gibt.

Aufgabe 4

Das Experiment war gut geplant. Die Power lag bei 94%. Da die Daten trotzdem kein signifikantes Ergebnis zeigten, liegt die Vermutung nahe, dass der angenommene Effekt zu hoch war. Wäre er in Wirklichkeit nur von der Größe d = 0,20, hätte das Experiment nur eine Power von 29% gehabt! Prüfen Sie es nach!



Aufgabe 5

a) „Analysieren“ → „Mittelwerte vergleichen“ → „t-Test bei gepaarten Stichproben“. Gepaarte Variablen sind „Messung 1“ und „Messung 3“. Die Analyse liefert folgenden Output:

Statistik bei gepaarten Stichproben

		Mittelwert	N	Standardabweichung	Standardfehler des Mittelwertes
Paaren 1	Messung1	16,56	36	4,931	,822
	Messung3	18,25	36	4,789	,798

Korrelationen bei gepaarten Stichproben

		N	Korrelation	Signifikanz
Paaren 1	Messung1 & Messung3	36	,649	,000

Test bei gepaarten Stichproben

		Gepaarte Differenzen							
		Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	95% Konfidenzintervall der Differenz: Untere	95% Konfidenzintervall der Differenz: Obere	T	df	Sig. (2-seitig)
Paaren 1	Messung1 - Messung3	-1,694	4,077	,679	-3,074	-,315	-2,494	35	,018

Der Unterschied von durchschnittlich 1,694 mehr geschafften Sequenzen am dritten Messzeitpunkt ist signifikant.

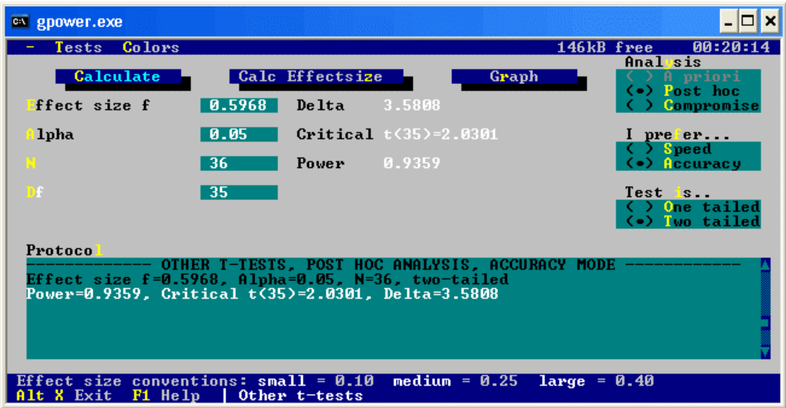
b) In den Erläuterungen zu GPower für Kapitel 3 ist beschrieben, wie sich die Effektstärke d_z im Fall von abhängigen Stichproben per Hand berechnen lässt. Mit GPower ist dies leider nicht möglich.

$$d_z = \frac{|\bar{x}_d|}{\sigma_d} = \frac{1{,}694}{4{,}077} = 0{,}416$$

c) In GPower können wir die Teststärke berechnen, die vorlag, um einen mittleren Effekt von $d_{unabhängig} = 0{,}5$ aufdecken zu können. Da es sich um abhängige Stichproben handelt, müssen wir in das Feld für die Effektstärke d_z’ eingeben, die die empirische Korrelation zwischen den Messwertreihen mit einbezieht.

$$d_z' = f' = \sqrt{\frac{2}{1-r}} \cdot f_{\text{unabhängig}} = \sqrt{\frac{2}{1-r}} \cdot \frac{d_{\text{unabhängig}}}{2} = \sqrt{\frac{2}{1-0{,}649}} \cdot \frac{0{,}5}{2} = 0{,}5968$$

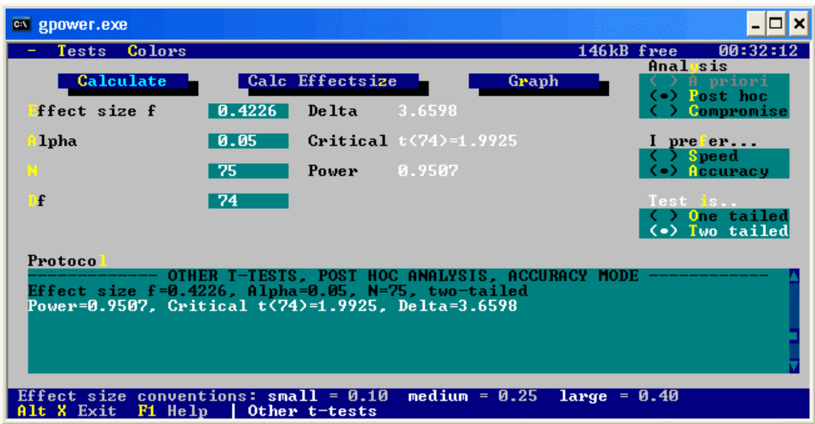
Außerdem handelt es sich um eine ungerichtete Fragestellung. Deshalb müssen wir den entsprechenden Punkt markieren.



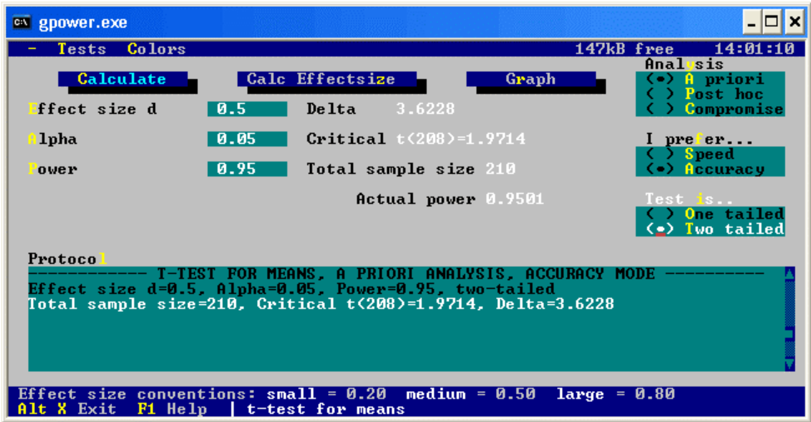
Dank der hohen Korrelation zwischen den Messwertreihen hatte die Untersuchung eine hervorragende Teststärke, um einen angenommenen Effekt von $d_{\text{unabhängig}} = 0{,}5$ aufzudecken.

d) Um einen Effekt von $d_{\text{unabhängig}} = 0{,}5$ mit einer Power von 95% bei $\alpha = 0{,}05$ und einer angenommenen Korrelation zwischen den abhängigen Messwertpaaren von $r = 0{,}30$ zu finden, wären 75 Personen nötig gewesen.

$$d_z' = f' = \sqrt{\frac{2}{1-r}} \cdot f_{\text{unabhängig}} = \sqrt{\frac{2}{1-r}} \cdot \frac{d_{\text{unabhängig}}}{2} = \sqrt{\frac{2}{1-0{,}3}} \cdot \frac{0{,}5}{2} = 0{,}4226$$



e) Im Menü für einen t-Test für unabhängige Stichproben steht uns die A priori Analyse zur Verfügung.



Wäre diese Fragestellung mit unabhängigen Stichproben untersucht worden, hätte sich der Bedarf an Versuchspersonen auf 210 Personen im Gegensatz zu 75 Personen in Aufgabe d) mehr als verdoppelt. Dieses Beispiel macht die große Überlegenheit des t-Tests für abhängige Stichproben in Fragen der Teststärke deutlich. Ohne den Einfluss der Korrelation (also bei r = 0) wären im t-Test für abhängige Stichproben ungefähr 105 Versuchspersonen notwendig gewesen (bitte nachrechnen). Dies entspricht der Hälfte der hier ermittelten Anzahl von 210 Personen. Bei der Messwiederholung gibt jede Person zwei Messwerte ab. Im Fall einer Nullkorrelation zwischen abhängigen Stichproben führt eine identische Zahl von Messwerten also zu identischen Ergebnissen bei t-Tests für unabhängige und abhängige Stichproben.