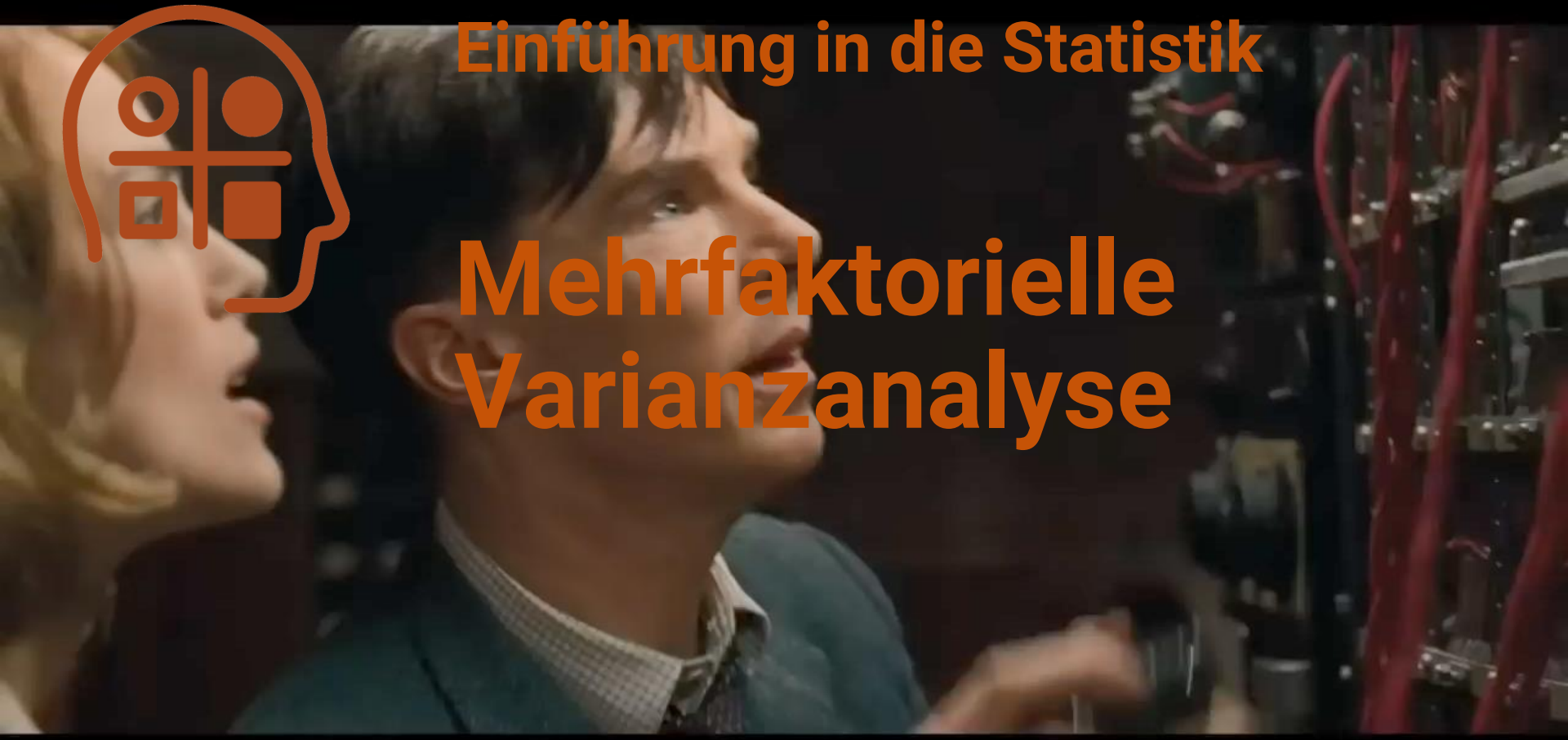
TECHNISCHE UNIVERSITÄT
IN DER KULTURHAUPTSTADT EUROPAS
CHEMNITZ

Professur Psychologie digitaler Lernmedien
Institut für Medienforschung
Philosophische Fakultät



Einführung in die Statistik

Mehrfaktorielle Varianzanalyse

The Imitation Game (2014). The Weinstein Company.

Prof. Dr. Günter Daniel Rey

Überblick

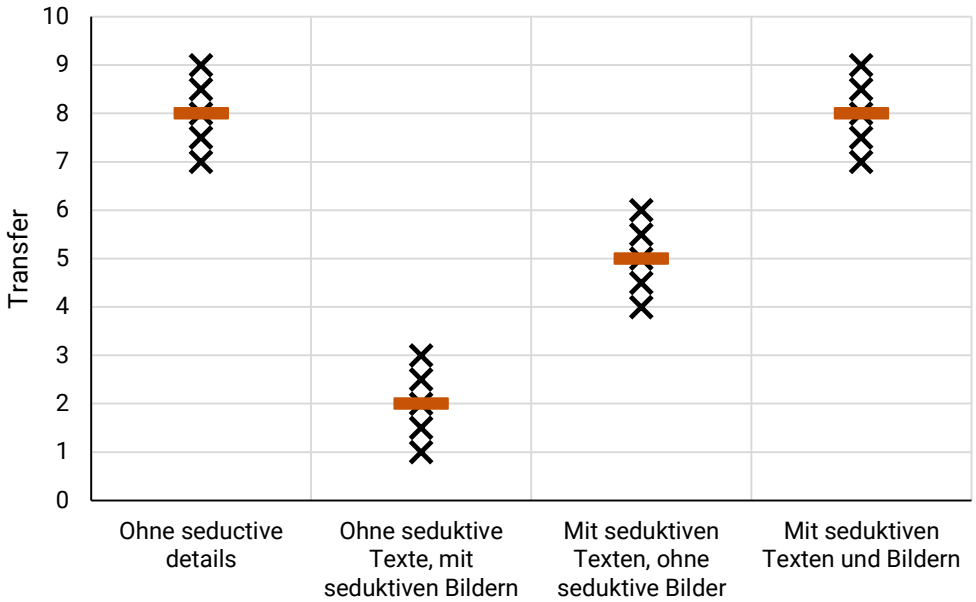
- Einführung
- Empirische *F*-Werte zu einer zweifaktoriellen Varianzanalyse
- Quadratsummen innerhalb der Zellen und für die Haupteffekte A und B sowie für die Wechselwirkung A x B
- Zähler- und Nennerfreiheitsgrade
- Inferenzstatistische Entscheidung und Ergebnisdarstellung
- Wechselwirkung
- Post-hoc-Analysen
- Inferenzstatistische Voraussetzungen

Einführung (z. B. Rasch, Friese, Hofmann & Naumann, 2021)

- Varianzanalyse (engl. analysis of variance, ANOVA): Statistisches Verfahren zum simultanen Vergleich mehrerer Mittelwerte
- Einfaktorielle Varianzanalyse: Varianzanalyse zu einem einfaktoriellen Versuchsdesign
- Zweifaktorielle Varianzanalyse: Varianzanalyse zu einem zweifaktoriellen Versuchsdesign mit drei Effekten
  - Haupteffekt A
  - Haupteffekt B
  - Wechselwirkung A x B

Beispiel zur Berechnung einer zweifaktoriellen Varianzanalyse

- Beispiel: Fiktive Ergebnisse einer Studie zum seductive detail Effekt mit einem 2 x 2 faktoriellen Design
  - UV_1: Seduktive Texte (mit vs. ohne)
  - UV_2: Seduktive Bilder (mit vs. ohne)
- Rohdaten als schwarze Kreuze, Mittelwerte als orangene Linien



Systematische Varianz hoch
Residualvarianz gering

Beispiel zur Berechnung einer zweifaktoriellen Varianzanalyse

- Beispiel: Fiktive Rohdaten zu einer Studie zum seductive detail Effekt (vgl. rechte Abbildung auf der vorherigen Seite)

Ohne seductive details		Ohne seduktive Texte, mit seduktiven Bildern		Mit seduktiven Texten, ohne seduktive Bilder		Mit seduktiven Texten und Bildern	
VPN	**Transfer**	**VPN**	**Transfer**	**VPN**	**Transfer**	**VPN**	**Transfer**
1	7.0	6	2.0	11	5.0	16	8.5
2	9.0	7	1.5	12	5.5	17	8.0
3	8.5	8	1.0	13	4.5	18	7.0
4	7.5	9	2.5	14	6.0	19	9.0
5	8.0	10	3.0	15	4.0	20	7.5
M	8.0	*M*	2.0	*M*	5.0	*M*	8.0

Berechnung der empirischen *F*-Werte

- Für die drei Effekte gilt grundsätzlich:

$$F = \frac{\frac{QS_{Zwischen}}{df_Z}}{\frac{QS_{Innerhalb}}{df_N}}$$

F	= Empirischer *F*-Wert
$QS_{Zwischen}$	= Quadratsumme zwischen den Zellen
$QS_{Innerhalb}$	= Quadratsumme innerhalb der Zellen
df_Z	= Zählerfreiheitsgrade
df_N	= Nennerfreiheitsgrade

- Für die Haupteffekte A und B sowie die Wechselwirkung A x B:

$$F_A = \frac{\frac{QS_A}{df_A}}{\frac{QS_{Innerhalb}}{df_N}}$$

$$F_B = \frac{\frac{QS_B}{df_B}}{\frac{QS_{Innerhalb}}{df_N}}$$

$$F_{AxB} = \frac{\frac{QS_{AxB}}{df_{AxB}}}{\frac{QS_{Innerhalb}}{df_N}}$$

Quadratsumme innerhalb der Zellen

- Formel zur Berechnung der Quadratsumme innerhalb der Zellen:

$$QS_{Innerhalb} = \sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{m=1}^{n}\left(x_{ijm} - \overline{A}\overline{B}_{ij}\right)^2$$

x_{ijm}	= Wert der Person *m* unter der Faktorstufenkombination *i* des Faktors A und *j* des Faktors B
$\bar{A}\bar{B}_{ij}$	= Mittelwert unter der Faktorstufenkombination *i* des Faktors A und *j* des Faktors B

- Für das Beispiel gilt:

$$QS_{Innerhalb} = (7.0-8.0)^2 + (9.0-8.0)^2 + (8.5-8.0)^2 + (7.5-8.0)^2 + (8.0-8.0)^2$$
$$+ \ldots + (7.5-8.0)^2 = (2.5+2.5+2.5+2.5) = 10.0$$

Quadratsumme für den Haupteffekt A

- Formel zur Berechnung der Quadratsumme für den Haupteffekt A:

$$QS_A = \sum_{i=1}^{p} n \cdot q \cdot \left(\overline{A}_i - \overline{G}\right)^2$$

p = Anzahl an Faktorstufen des Faktors A
n = Anzahl an Versuchspersonen in einer Gruppe
q = Anzahl an Faktorstufen des Faktors B
$\bar{A}_i$ = Mittelwert der Gruppe i
$\bar{G}$ = Gesamtmittelwert

- Für das Beispiel gilt:
  - Anzahl an Faktorstufen der Faktoren A und B: Jeweils 2
  - Anzahl an Versuchspersonen in einer Gruppe: 5
  - Gruppenmittelwerte: 5.0 (ohne seduktive Texte), 6.5 (mit seduktiven Texten); Gesamtmittelwert: 5.75
- Berechnung:
- $QS_A = [5 \cdot 2 \cdot (5.0 - 5.75)^2] + [5 \cdot 2 \cdot (6.5 - 5.75)^2] = [10 \cdot (-0.75)^2] + [10 \cdot (0.75)^2] = 11.25$

Quadratsumme für den Haupteffekt B

- Formel zur Berechnung der Quadratsumme für den Haupteffekt B:

$$QS_B = \sum_{j=1}^{q} n \cdot p \cdot \left(\overline{B}_j - \overline{G}\right)^2$$

q = Anzahl an Faktorstufen des Faktors B
n = Anzahl an Versuchspersonen in einer Gruppe
p = Anzahl an Faktorstufen des Faktors A
$\bar{B}_j$ = Mittelwert der Gruppe j
$\bar{G}$ = Gesamtmittelwert

- Für das Beispiel gilt:
  - Anzahl an Faktorstufen der Faktoren A und B: Jeweils 2
  - Anzahl an Versuchspersonen in einer Gruppe: 5
  - Gruppenmittelwerte: 6.5 (ohne seduktive Bilder), 5.0 (mit seduktiven Bildern); Gesamtmittelwert: 5.75
- Berechnung:
- $QS_B = [5 \cdot 2 \cdot (6.5 - 5.75)^2] + [5 \cdot 2 \cdot (5.0 - 5.75)^2] = [10 \cdot (0.75)^2] + [10 \cdot (-0.75)^2] = 11.25$

Quadratsumme für die Wechselwirkung A x B

- Formel zur Berechnung der Quadratsumme für die Wechselwirkung:

$$QS_{AxB} = \sum_{i=1}^{p} \sum_{j=1}^{q} n \cdot \left[\overline{A}\overline{B}_{ij} - (\overline{A}_i + \overline{B}_j - \overline{G})\right]^2$$

- Für das Beispiel gilt:
  - Anzahl an Versuchspersonen in einer Gruppe: 5
  - Gruppenmittelwerte: 5.0 und 6.5 sowie 8.0, 2.0, 5.0 und 8.0
  - Gesamtmittelwert: 5.75
- Berechnung:
- $QS_{AxB} = 5 \cdot [8.0 - (5.0 + 6.5 - 5.75)]^2 + 5 \cdot [2.0 - (5.0 + 5.0 - 5.75)]^2$

$$+ 5 \cdot [5.0 - (6.5 + 6.5 - 5.75)]^2 + 5 \cdot [8.0 - (6.5 + 5.0 - 5.75)]^2 = 101.25$$

Zählerfreiheitsgrade

- Bei einer zweifaktoriellen, univariaten Varianzanalyse: Folgende Formeln für die Zählerfreiheitsgrade:
  - Haupteffekt A: $df_A = p - 1$ | p = Anzahl an Faktorstufen des Faktors A
  - Haupteffekt B: $df_B = q - 1$ | q = Anzahl an Faktorstufen des Faktors B
  - Wechselwirkung A x B: $df_{A \times B} = (p - 1) \cdot (q - 1)$
- Beispiel für ein 2 x 2 faktorielles Design:
  - $df_A = 2 - 1 = 1$
  - $df_B = 2 - 1 = 1$
  - $df_{A \times B} = (2 - 1) \cdot (2 - 1) = 1$

Nennerfreiheitsgrade

- Formel Nennerfreiheitsgrade: $df_N = N - p \cdot q$ bzw. $p \cdot q \cdot (n - 1)$

N	= Gesamter Stichprobenumfang
p	= Anzahl an Faktorstufen des Faktors A
q	= Anzahl an Faktorstufen des Faktors B
n	= Anzahl an Versuchspersonen in einer Gruppe

- Beispiel für ein 2 x 2 faktorielles Design mit jeweils fünf Personen:
- $df_N = 20 - 4 = 16$ bzw. $2 \cdot 2 \cdot (5 - 1) = 16$

Zähler- und Nennerfreiheitsgrade

In einem Experiment mit einem 2 x 3 faktoriellen, univariaten Versuchsdesign wurden pro Versuchsbedingung 12 Probanden erhoben. Wie hoch sind die Zähler- und Nennerfreiheitsgrade für diese Studie?

- ☐ A: $df_A = 1$ und $df_N = 66$
- ☐ B: $df_B = 2$ und $df_N = 72$
- ☐ C: $df_{A \times B} = 1$ und $df_N = 72$
- ☐ D: $df_{A \times B} = 2$ und $df_N = 66$
- ☐ E: $df_{A \times B} = 6$ und $df_N = 66$

Beispiel zur Ermittlung der empirischen *F*-Werte und des kritischen *F*-Wertes

- Berechnung der empirischen *F*-Werte:

$$F_A = \frac{\frac{QS_A}{df_A}}{\frac{QS_{Innerhalb}}{df_N}} \qquad F_B = \frac{\frac{QS_B}{df_B}}{\frac{QS_{Innerhalb}}{df_N}} \qquad F_{AxB} = \frac{\frac{QS_{AxB}}{df_{AxB}}}{\frac{QS_{Innerhalb}}{df_N}}$$

$$F_A = \frac{\frac{11.25}{1}}{\frac{10}{16}} = 18 \qquad F_B = \frac{\frac{11.25}{1}}{\frac{10}{16}} = 18 \qquad F_{AxB} = \frac{\frac{101.25}{1}}{\frac{10}{16}} = 162$$

- Ermittlung des kritischen *F*-Wertes für $df_z = 1$ und $df_N = 16$ sowie $\alpha = .05$: $F_{krit} \approx 4.49$

Inferenzstatistische Entscheidung

- Zwei Möglichkeiten für jeden einzelnen Effekt (Haupteffekt A, Haupteffekt B und Wechselwirkung A x B)
  - Nicht signifikant ($F_{emp} < F_{krit}$): H_0 wird vorläufig beibehalten
  - Signifikant ($F_{emp} \geq F_{krit}$): H_0 wird zugunsten der H_1 verworfen
- Beispiel
  - Haupteffekt A: $F_{emp} = 18 \geq F_{krit} \approx 4.49$ → Signifikant
  - Haupteffekt B: $F_{emp} = 18 \geq F_{krit} \approx 4.49$ → Signifikant
  - Wechselwirkung A x B: $F_{emp} = 162 \geq F_{krit} \approx 4.49$ → Signifikant

Inferenzstatistische Ergebnisdarstellung

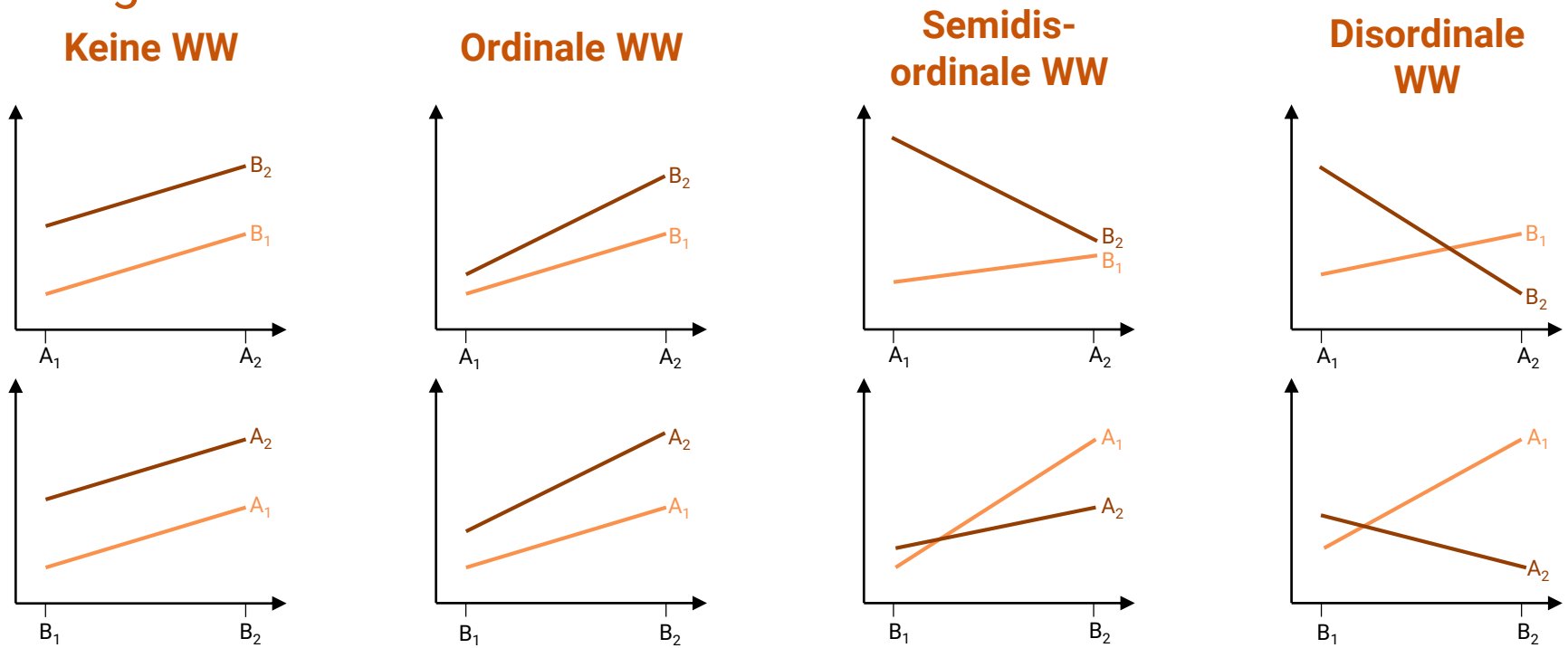
- Bei signifikanten Ergebnissen: Angabe des *F*-Wertes (inklusive Zähler- und Nennerfreiheitsgrade) und *p*-Wertes sowie der Effektstärke
- Bei nicht signifikanten Ergebnissen: Zusätzlich Angabe der Teststärke
- Beispiele
  - $F_A(1,16) = 18$, $p < .001$, $\eta_p^2 = .53$
  - $F_B(1,16) = 18$, $p < .001$, $\eta_p^2 = .53$
  - $F_{A \times B}(1,16) = 162$, $p < .001$, $\eta_p^2 = .91$
- Fiktives Beispiel für ein nicht signifikantes Ergebnis
  - $F_A(4,163) = 1.73$, $p = .63$, $\eta_p^2 = .01$. Die Teststärke (1-β) für einen Effekt von $f^2 = .15$ und $\alpha = .05$ beträgt .63.

Wechselwirkung (z. B. Rasch, Friese, Hofmann & Naumann, 2021)

- Wechselwirkung (auch Interaktionseffekt oder Moderatoreffekt): Über die Haupteffekte hinausgehender Effekt, der durch die Kombination einzelner Faktorstufen zustande kommt
- Drei Arten von Wechselwirkungen (WW)
  - Ordinale Wechselwirkung: Effekt der WW ist kleiner als jeder der beiden Haupteffekte
  - Semidisordinale Wechselwirkung: Effekt der WW ist kleiner als einer der beiden Haupteffekte und größer als der andere der beiden Haupteffekte
  - Disordinale Wechselwirkung: Effekt der WW ist größer als jeder Haupteffekt

Wechselwirkung (z. B. Rasch, Friese, Hofmann & Naumann, 2021)

- Verschiedene Arten von Wechselwirkungen (WW) für ein 2 x 2 Design:



- x-Achse: Faktorstufen eines Faktors; Farben: Faktorstufe 1 und Faktorstufe 2 des anderen Faktors; y-Achse: Abhängige Variable

Post-hoc-Analysen (z. B. Rasch, Friese, Hofmann & Naumann, 2021)

- Tukey HSD-Test: Post-hoc-Verfahren für den Vergleich von (mehr als zwei) Gruppenmittelwerten
- HSD: Berechnung über die kleinste noch signifikante Differenz zwischen zwei Gruppenmittelwerten (engl. honest significant difference, HSD)
- Für einfaktorielle Varianzanalysen gilt:

$$\mathrm{HSD} = q_{krit(\alpha;r;df_{Innerhalb})} \cdot \sqrt{\frac{\hat{\sigma}^2_{Innerhalb}}{n}}$$

q = Kennwert abhängig von der Zahl der Mittelwerte *r* und den Freiheitsgraden innerhalb
α = Alpha-Niveau des *F*-Tests
r = Anzahl der betrachteten Zellmittelwerte
n = Anzahl an Versuchspersonen pro Zelle

Post-hoc-Analysen (z. B. Rasch, Friese, Hofmann & Naumann, 2021)

- Auch bei mehrfaktoriellen Varianzanalysen: Post-hoc-Verfahren wie der Tukey HSD-Test dienen zum Vergleich von Mittelwerten
- Berechnung der HSD für jeden Effekt mit einem anderen n_{HSD}:
  - Für den Haupteffekt A: $n_{HSD(A)} = q \cdot n$
  - Für den Haupteffekt B: $n_{HSD(B)} = p \cdot n$
  - Für die Wechselwirkung: $n_{HSD(A \times B)} = n$

q = Anzahl an Stufen des Faktors B
p = Anzahl an Stufen des Faktors A
n = Anzahl an Personen pro Zelle

- Berechnung der HSD beispielsweise für den Haupteffekt A:

$$\mathrm{HSD}_{(A)} = q_{krit(\alpha;r;df_{Innerhalb})} \cdot \sqrt{\frac{\hat{\sigma}^2_{Innerhalb}}{n_{HSD(A)}}}$$

α = Alpha-Niveau des *F*-Tests
r = Anzahl der betrachteten Zellmittelwerte

Post-hoc-Analysen

- Berechnung der HSD für die Wechselwirkung A x B aus dem vorherigen Beispiel:
- *q*-Wert laut Tabelle: 4.05 (für α = .05, *r* = 4 und df_N = 16)
- Berechnung von $\hat{\sigma}^2_{Innerhalb}$ mittels der Formel:
- $\hat{\sigma}^2_{Innerhalb} = \frac{QS_{Innerhalb}}{df_{Innerhalb}}$ ergibt: $\hat{\sigma}^2_{Innerhalb} = \frac{10}{16} = 0.625$
- Einsetzen ergibt: $\mathrm{HSD}_{(\mathrm{AxB})} = 4.05 \cdot \sqrt{\frac{0.625}{5}} \approx 4.05 \cdot 0.35 \approx 1.43$
- HSD ≈ 1.43 für die Wechselwirkung A x B

Tukey HSD-Test

In einem Experiment mit einem einfaktoriellen, dreifachgestuften Design werden pro Bedingung drei Probanden erhoben. Wie hoch ist die HSD bei $\hat{\sigma}^2_{Innerhalb}$ = 12 und α = .05 (q_{krit} kann aus der Tabelle rechts entnommen werden)?

- ☐ A: 6.92
- ☐ B: 12.66
- ☐ C: 8.68

Error *df*		Anzahl Mittelwerte *r*		
	α	2	3	4
2	0,05	6,08	8,33	9,80
	0,01	14,00	19,00	22,30
3	0,05	4,50	5,91	6,82
	0,01	8,26	10,60	12,20
4	0,05	3,93	5,04	5,76
	0,01	6,51	8,12	9,17
5	0,05	3,64	4,60	5,22
	0,01	5,70	6,98	7,80
6	0,05	3,46	4,34	4,90
	0,01	5,24	6,33	7,03
7	0,05	3,34	4,16	4,68
	0,01	4,95	5,92	6,54
8	0,05	3,26	4,04	4,53
	0,01	4,75	5,64	6,20
9	0,05	3,20	3,95	4,41
	0,01	4,60	5,43	5,96
10	0,05	3,15	3,88	4,33
	0,01	4,48	5,27	5,77

Quelle: Rasch, Friese, Hofmann und Naumann (2021)

Mehrfaktorielle Varianzanalysen mit mehr als zwei Faktoren

- Berechnung: Mehrfaktorielle Varianzanalysen mit mehr als zwei Faktoren analog zur Berechnung zweifaktorieller Varianzanalysen
- Mehrere Wechselwirkungen: z. B. A x B, A x C und B x C sowie Wechselwirkungen höherer Ordnung (z. B. A x B x C)
- Vorteil: Höhere Varianzaufklärung durch zusätzliche Faktoren und deren Wechselwirkungen
- Nachteile
  - Verringerung der Nennerfreiheitsgrade df_N
  - Wechselwirkungen höherer Ordnung häufig schwierig bis kaum noch interpretierbar
- Empfehlung: Zusätzliche Faktoren nur dann aufnehmen, wenn Hypothesen zu diesen formuliert worden sind

Inferenzstatistische Voraussetzungen (z. B. Rasch, Friese, Hofmann & Naumann, 2021)

- Inferenzstatistische Voraussetzungen der mehrfaktoriellen Varianzanalyse: Entsprechen denen der einfaktoriellen Varianzanalyse
  - Unabhängigkeit der Messwerte in den einzelnen Bedingungen (bei nicht messwiederholten Versuchsdesigns)
  - Intervallskalenniveau der abhängigen Variable
  - Normalverteilung der abhängigen Variable in der Population (getrennt für jede Versuchsbedingung oder auf Basis der Residuen)
  - Varianzhomogenität als Gleichheit der Populationsvarianzen, aus denen die Stichproben stammen
- Robustheit: *F*-Test reagiert bei ungefähr gleichgroßen und hinreichend großen Stichproben der Gruppen relativ robust gegenüber Verletzungen der Voraussetzungen

Post-hoc-Analysen bei mehrfaktoriellen Varianzanalysen mit mehr als zwei Faktoren

Sind Post-hoc-Analysen wie der Tukey HSD-Test bei Versuchsplänen mit ausschließlich zweifachgestuften Faktoren (z. B. bei einem 2 x 2 x 2 faktoriellen Versuchsdesign) sinnvoll?

- ☐ A: Ja, in bestimmten Situationen können diese sinnvoll sein.
- ☐ B: Nein!

Beispiele für mehrfaktorielle Varianzanalysen bzw. *F*-Tests in Fachzeitschriften

Descriptive statistics for pretest and posttest scores and mental effort invested in learning from the animations are shown in Table 2. On the tests, missing answers were scored as errors. The significance level for the comparisons was set at 0.05, and eta-squared is reported as a measure of effect size, with 0.01 indicating a small, 0.06 a moderate and 0.14 a large effect. The conditions did not differ in pretest scores, $F(3,157) = 1.21$, MSE $= 1.02$, $p = 0.31$, $\eta_p^2 = 0.02$.

A 2-by-2 ANOVA revealed a main effect of Pausing on posttest scores: participants studying animations with pauses ($M = 6.46$, sd $= 1.79$) performed significantly better on the posttest than participants studying animations without pauses ($M = 5.80$, sd $= 2.16$), $F(1,157) = 4.39$, $p = 0.04$, MSE $= 3.92$, $\eta_p^2 = 0.03$. There was no significant main effect of Cueing, $F(1,157) = 0.49$, $p = 0.48$, $\eta_p^2 = 0.003$ nor an interaction effect between Cueing and Pausing, $F(1,157) = 2.85$, $p = 0.09$, $\eta_p^2 = 0.02$.

A 2-by-2 ANOVA on mental effort invested during animation study showed a significant main effect of Cueing: learning from animations with cues ($M = 2.19$, sd $= 1.05$) required significantly less investment of mental effort than learning from animations without cues ($M = 2.73$, sd $= 1.70$), $F(1,157) = 5.67$, $p = 0.02$, MSE $= 2.01$, $\eta_p^2 = 0.03$. There was no significant main effect of Pausing $F(1,157) = 2.33$, $p = 0.13$, $\eta_p^2 = 0.01$, nor a significant interaction effect $F(1,157) = 0.33$, $p = 0.56$, $\eta_p^2 = 0.002$.

Quelle: Spanjers, van Gog, Wouters und van Merriënboer (2012)

The first-year students' data (test task performance, time spent on the learning tasks, mental effort invested in the learning tasks, and mental effort invested in the test task) were analyzed using 2×2 ANOVAs with factors 'Worked Examples' and 'Process Steps'. On test performance, there was a significant main effect of 'Worked Examples', $F(1,68) = 105.221$, $MSE = 426.037$, $p < .001$, $f = 1.24$, indicating that students who studied worked examples during the learning phase, performed significantly better on the test task ($M = 58.63$, $SD = 27.74$) than students who engaged in problem solving ($M = 8.59$, $SD = 4.06$). There was no significant main effect of 'Process Steps', $F(1,68) < 1$, nor a significant interaction effect $F(1,68) < 1$. Because the assumption of homogeneity of variance

Quelle: Nievelstein et al. (2013)

In addition, the sexy female pedagogical agents were perceived as more attractive than the sexy male agents. Table 1 shows that both male ($F(1, 38) = 7.18$, $p = .011$) and female ($F(1, 55) = 19.41$, $p < .001$) students rated the seductive female pedagogical agent as more attractive than the seductive male agent, and both male ($F(1, 45) = 48.04$, $p < .001$) and female ($F(1, 63) = 26.336$, $p < .001$) students rated the overtly sexual female pedagogical agent as more attractive than the overtly sexual male agent. These results show that if a pedagogical agent with a sexy appearance is to be used in learning materials, a female agent is to be perceived as more attractive than a male agent.

Quelle: Wang und Yeh (2013)

Zusammenfassung

- Varianzanalyse: Statistisches Verfahren zum simultanen Vergleich mehrerer Mittelwerte
- Grundprinzip: Zerlegung der Gesamtvarianz in systematische Varianz und Residualvarianz
- Inferenzstatistische Überprüfung der Varianzverhältnisse mit dem kritischen und empirischen *F*-Wert

$$F = \frac{\frac{QS_{Zwischen}}{df_Z}}{\frac{QS_{Innerhalb}}{df_N}}$$

- Ordinale, semidisordinale und disordinale Wechselwirkung
- Post-hoc-Verfahren wie der Tukey HSD-Test zur spezifischen Testung ausgewählter Mittelwerte
- Voraussetzung von Varianzanalysen: Intervallskalenniveau, Normalverteilung, Varianzhomogenität und ggf. Unabhängigkeit der Messwerte

Prüfungsliteratur

- Rasch, B., Friese, M., Hofmann, W., & Naumann, E. (2021). *Quantitative Methoden 2: Einführung in die Statistik für Psychologie, Sozial- & Erziehungswissenschaften* (5. Aufl.). Heidelberg: Springer.
  - Zweifaktorielle Varianzanalyse (S. 41–69)

Weiterführende Literatur

- Bortz, J., & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler* (7. Aufl.). Berlin: Springer.
  - Zweifaktorielle Versuchspläne (S. 237–256)
  - Kontraste für zweifaktorielle Versuchspläne (S. 257–264)
  - Drei- und mehrfaktorielle Versuchspläne (S. 265–274)
- Eid, M., Gollwitzer, M., & Schmitt, M. (2017). *Statistik und Forschungsmethoden* (5. Aufl.). Weinheim: Beltz.
  - Zweifaktorielle Varianzanalyse (S. 430–454)
- Leonhart, R. (2017). *Lehrbuch Statistik. Einstieg und Vertiefung* (4. Auflage). Bern: Huber.
  - Zweifaktorielle Varianzanalyse mit festen Effekten (S. 433–465)
- Sedlmeier, P., & Renkewitz, F. (2018). *Forschungsmethoden und Statistik: Ein Lehrbuch für Psychologen und Sozialwissenschaftler* (3. Aufl.). München: Pearson.
  - Mehrfaktorielle Varianzanalyse (S. 465–484)