

Einführung in die Eyetracking-Methode 1

Amit Singh

06.11.2025

Inhaltsverzeichnis

1	Experimenteller Prozess	2
2	Schritt 2, 3, 4	2
2.1	Schritt 2: Entwicklung eines Experiment-Designs	2
2.2	Schritt 3: Durchführung des Experiments	2
2.3	Schritt 4: Inferenz und Interpretation der Ergebnisse	2
2.4	Beispiel aus der Psycholinguistik	3
3	Experiment-Design	4
3.1	Einfaktorielles Design	4
3.2	Mehrfaktorielle Designs	4
4	Manipulation der Faktoren	5
4.1	Between-Subjects-Manipulation	5
4.2	Within-Subjects-Manipulation	5
5	Latin-Square Design (Lateinisches Quadrat)	7
6	Haupteffekt (main effect)	7
6.1	Kinderaufmerksamkeit Beispiel	7
7	Interaktion (interaction effect)	8
8	Konkretes Beispiel mit Daten	9
9	Basisanalyse führen für Haupteffekts	10

1 Experimenteller Prozess

- Schritt 1. Hypothesenbildung
- Schritt 2. Entwicklung eines Experiment-Designs
- Schritt 3. Durchführung des Experiments
- Schritt 4. Inferenz und Interpretation der Daten

2 Schritt 2, 3, 4

2.1 Schritt 2: Entwicklung eines Experiment-Designs

- Forschungsdesign
- Definition abhängiger und unabhängiger Variablen

2.2 Schritt 3: Durchführung des Experiments

- Datenerhebung

2.3 Schritt 4: Inferenz und Interpretation der Ergebnisse

- Datenaufbereitung und -auswertung

Unabhängige Variable (UV) *engl.* independent variable (IV)

- Variable, die von Experimentleiter:in gezielt manipuliert bzw. kontrolliert wird (mit bestimmten Levels). z.B. Schwierigkeit der Aufgabe (leicht vs. schwer), Stimulus-Typ (Bild vs. Wort)

Abhängige Variable (AV) *engl.* dependent variable (DV)

- Variable, die gemessen wird, um den Effekt der UV zu erfassen. z.B. Reaktionszeit, Genauigkeit, Blickverhalten

Störvariable *engl.* confounding variable

- Variable, die unbeabsichtigt die AV beeinflussen kann. Störvariablen sollten kontrolliert oder randomisiert werden.

2.4 Beispiel aus der Psycholinguistik

Hypothese: Interessantes Object (z. B. glänzendes/buntes Objekt) wird von den Kinder lange betrachtet werden als uninteressantes (z. B. grau Objekt).

- UV = Art des Objekts (glänzend vs. grau)
- AV = Blickdauer auf das Objekt (in ms)

Beispiel

Unabhängige Variable (UV): Art des Objekts (glänzend vs. grau)

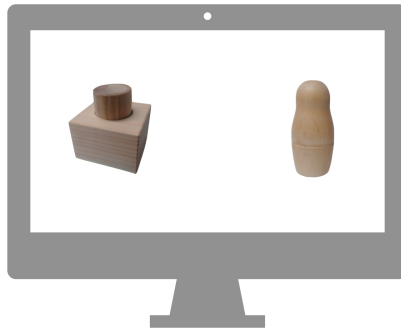


Abbildung 1: Grau (uninteressant)

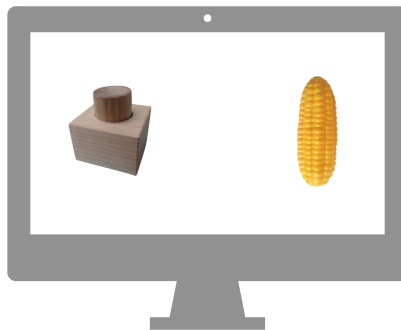


Abbildung 2: Glänzend (interessant)

Konkret könnte die Hypothese z. B. besagen, dass Kinder längere Blickzeiten auf bunte/glänzende Objekte haben als auf graue/uninteressante Objekte.

Abhängige Variable (AV): Blickdauer auf das Objekt (in ms)

Der Blickdauer auf jedes Objekt wird gemessen, um den Effekt der UV zu erfassen. Längere Blickzeiten auf bunte/glänzende Objekte im Vergleich zu grauen/uninteressanten Objekten würden die Hypothese unterstützen.

Hier muss man sich fragestellen

Frage 1: Wie viel Stufen (Levels) hat die UV (Faktor)?

Frage 2: Welche storvariablen könnten die AV beeinflussen?

3 Experiment-Design

3.1 Einfaktorielles Design

- In unsere Beispiel haben wir Faktor (UV): z.B., Art der Objekts
- Faktorstufen (factor levels): interessantes Objekt und uninteressantes Objekt)

d.h. wir haben 2×1 Design (2 Stufen der einzige UV).

3.2 Mehrfaktorielle Designs

Hier untersuchen wir Effekte von zwei oder mehreren Faktoren (UVs) auf die Unabhängige Variablen.

Fully crossed factorial design: Untersuchung jeder Kombination von Faktorstufen der jeweiligen Faktoren.

Beispiel: 2 X 2 fully-crossed factorial design

- zwei Faktoren mit jeweils zwei Stufen (Levels)
- Kombination der Faktorstufen: experimentelle Bedingungen

F1: Art des Objekts (interessant vs. uninteressant)

F2: Alter der Kinder (jung vs. alt)

- a) Interessantes Objekt (IO), junges Kind
- b) Interessantes Objekt (IO), altes Kind
- c) Uninteressantes Objekt (UO), junges Kind
- d) Uninteressantes Objekt (UO), altes Kind

a) IO jung	b) IO alt
c) UO jung	d) UO alt

Abbildung 3: 2 X 2 fully-crossed factorial design

4 Manipulation der Faktoren

4.1 Between-Subjects-Manipulation

Es gibt mehrere Versuchspersonen-Gruppen; jede Gruppe bekommt eine Bedingung (Kombination von Faktorstufen).

Beispiel: Vergleich eines Wirkstoffs und eines Placebos:

Ich möchte während des Experiments einem bestimmten Versuchssubjekt entweder nur den Wirkstoff oder nur das Placebo geben, aber logischerweise nicht beides.

Daher teile ich meine Versuchspersonen in zwei Gruppen auf: Gruppe A bekommt den Wirkstoff, Gruppe B das Placebo. Danach vergleiche ich die Ergebnisse der beiden Gruppen, um den Effekt des Wirkstoffs zu beurteilen.

4.2 Within-Subjects-Manipulation

Die gleichen Versuchspersonen durchlaufen alle Bedingungen (Kombinationen von Faktorstufen).

Beispiel: Kinder haben mehr Aufmerksamkeit für interessante Objekte (IO) als uninteressante Objekte (UO).

Hierfür zeige ich jedem Kind sowohl interessante als auch uninteressante Objekte während des Experiments. Ich messe die Blickzeiten auf beide Arten von Objekten für jedes Kind und vergleiche dann die Ergebnisse, um den Effekt der Objektart auf die Aufmerksamkeit zu beurteilen.

Note: Within-Subjects-Manipulationen haben in der Regel höhere **Power** und sind oft effizienter als Between-Subjects-Manipulationen.

```
## Diese code beantwortet unsere Frage.
## Wie viel Teststärke (Power) haben wir, wenn wir unsere Experiment
## mit insgesamt 35 kindern durchführen?
```

```
power.t.test(n=35, # Anzahl der Versuchspersonen
             delta=100, # Mindestunterschied in ms
             sd=200, # Standardabweichung in ms
             sig.level=0.05, # Signifikanzniveau
             type="paired", # gepaarter t-Test (within-subject)
             strict = TRUE)
```

```
##
##      Paired t test power calculation
##
##              n = 35
##              delta = 100
##              sd = 200
##              sig.level = 0.05
##              power = 0.8195358
##      alternative = two.sided
##
## NOTE: n is number of *pairs*, sd is std.dev. of *differences* within pairs
```

Diese code lässt sich unsere Frage zu Antworten.
##Wie viel Versuchspersonen (Kinder) brauchen wir, um unsere Hypothese zu testen?

```
power.t.test(delta=100,
             sd=200,
             power=0.8, # gewünschte Teststärke
             sig.level=0.05,
             type="paired",
             strict = TRUE)
```

```
##
##      Paired t test power calculation
##
##              n = 33.36713
##              delta = 100
##              sd = 200
##              sig.level = 0.05
##              power = 0.8
##      alternative = two.sided
##
## NOTE: n is number of *pairs*, sd is std.dev. of *differences* within pairs
```

Tabelle 1: Lateinisches Quadrat für 4 Bedingungen (A–D)

	Item 1	Item 2	Item 3	Item 4
Liste 1	A	B	C	D
Liste 2	B	C	D	A
Liste 3	C	D	A	B
Liste 4	D	A	B	C

5 Latin-Square Design (Lateinisches Quadrat)

Ein Lateinisches Quadrat ist ein Gegenbalancierungs-Design, in dem jede Bedingung in jeder Positionsstelle gleich oft erscheint. Es kontrolliert Positions-/Reihenfolgeeffekte, ohne dass jede VP alle Versionen eines Items sehen muss.

Einsatz: Within-Subjects mit mehreren Bedingungen/Versionen pro Item.

Beispiel 4 Bedingungen (A–D)

6 Haupteffekt (main effect)

Effekt einer einzelnen UV auf die AV, unabhängig von anderen UVs.

- einfaktorielle Designs: ein Haupteffekt
 - Ist der Wirkung von Stufen einer UV auf die AV signifikant unterschiedlich?
- mehrfaktorielle Designs: mehrere Haupteffekte (einer pro UV)
 - z. B. in einem 2×2 Design: Haupteffekt von UV1 und Haupteffekt von UV2

6.1 Kinderaufmerksamkeit Beispiel

Haupteffekt 1 : Beeinflusst die Art des Objekts (interessant vs. uninteressant) die Blickdauer der Kinder?

Haupteffekt 2 : Beeinflusst das Alter der Kinder (jung vs. alt) die Blickdauer der Kinder?

Faktor 1: Art des Objekts (IO vs. UO)

Faktor 2: Alter der Kinder (jung vs. alt)

Haupteffekt 1 (Art des Objekts)

$$\underbrace{\frac{\mu_a + \mu_b}{2}}_{\text{Mittelwert IO_Bedingungen}} - \underbrace{\frac{\mu_c + \mu_d}{2}}_{\text{Mittelwert UO_Bedingungen}}$$

a) IO jung	b) IO alt
c) UO jung	d) UO alt

Abbildung 4: 2 X 2 fully-crossed factorial design

Haupteffekt 2 (Alter der Kinder)

$$\underbrace{\frac{\mu_a + \mu_c}{2}}_{\text{Mittelwert junge Kinder}} - \underbrace{\frac{\mu_b + \mu_d}{2}}_{\text{Mittelwert alte Kinder}}$$

7 Interaktion (interaction effect)

Interaktionseffekte (Wechselwirkungseffekte) in Untersuchung mit mehr als einer UV.

- resultieren aus unseren zwei UVs, die kombiniert ein anderes Ergebnis erzeugen als das Ergebnis, welches von jeder UV einzeln erwartet wird. -> Effekt der UV1 ist abhängig von der Stufe der UV2 (und umgekehrt).
- Anders gesagt : Differenz zwischen dem Effekt des Faktor 1 innerhalb einer Stufe des Faktor 2 und dem Effekt des Faktor 1 innerhalb der anderen Stufe des Faktor 2.

Interagieren die Faktoren **Alter der Kinder X Art des Objekts**, um die Blickdauer zu beeinflussen?

*** zwischen a und c**

a) IO jung	b) IO alt
c) UO jung	d) UO alt

$$\underbrace{(\mu_a + \mu_c)}_{\text{Effekt ObjArt innerhalb junge Kinder}} - \underbrace{(\mu_b + \mu_d)}_{\text{Effekt ObjArt innerhalb alte Kinder}}$$

8 Konkretes Beispiel mit Daten

```
#zuerst müssen wir die Daten laden  
data<-read.csv("/Users/singh/Documents/Kognition/Lehre/blickdauer.csv")
```

```
#erste 6 Zeilen des Datensatzes anzeigen  
head(data)
```

```
##      ParticipantName      AOI GazeEventDuration  
## 1      VP02 Interessantes Objekt           83  
## 2      VP02 Interessantes Objekt           83  
## 3      VP02 Interessantes Objekt           83  
## 4      VP02 Interessantes Objekt           83  
## 5      VP02 Interessantes Objekt           83  
## 6      VP02 Interessantes Objekt           83
```

```
#diese paket brauchen wir für die Visualisierung und Datenmanipulation  
library(ggplot2)  
library(dplyr)
```

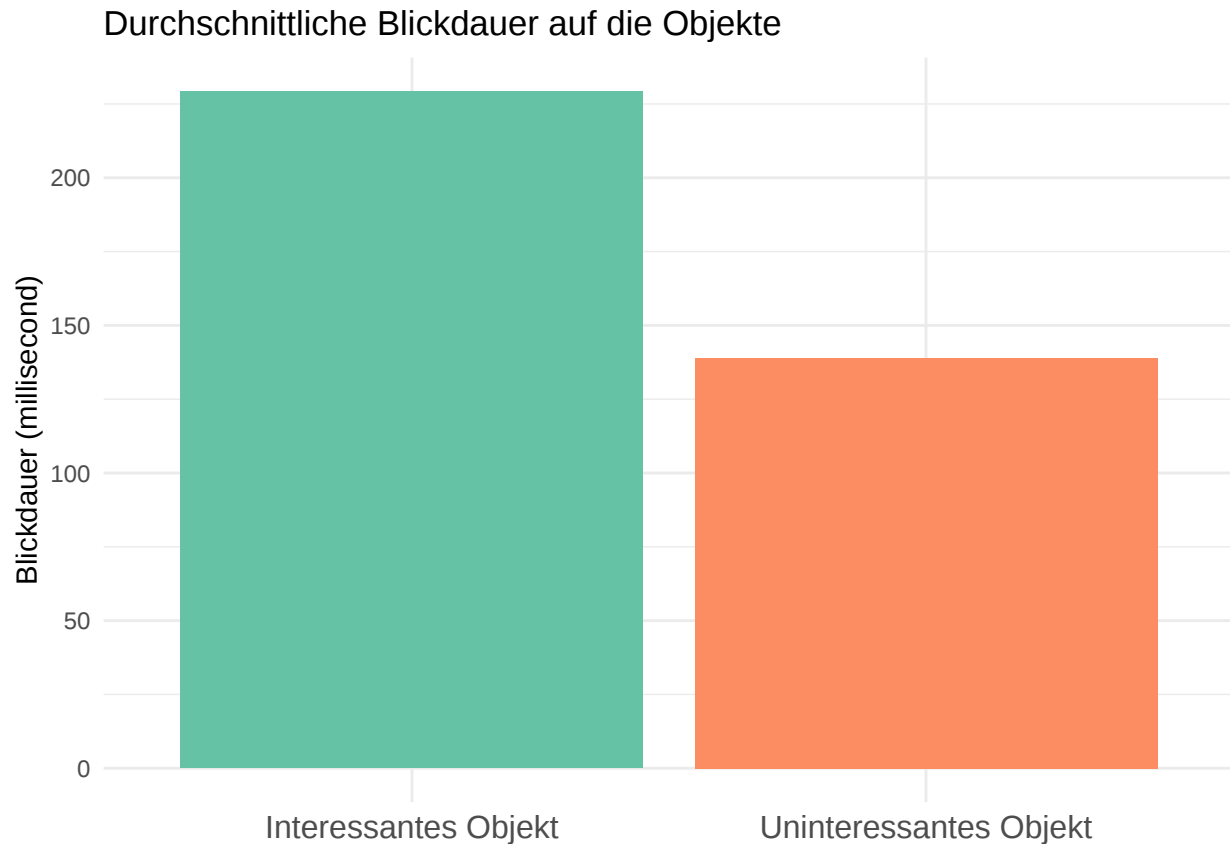
```
# Mittelwerte berechnen (die brauchen wir um die Effkte zu berechnen)  
summary_data <- data %>%  
  group_by(AOI) %>%  
  summarize(Mittelwerte = mean(GazeEventDuration, na.rm = TRUE))
```

```
# Mittelwerte anzeigen  
print(summary_data)
```

```
## # A tibble: 2 x 2  
##   AOI      Mittelwerte  
##   <chr>      <dbl>  
## 1 Interessantes Objekt      229.  
## 2 Uninteressantes Objekt    139.
```

```
# Plot erstellen  
ggplot(summary_data, aes(x = AOI, y = Mittelwerte, fill = AOI)) +  
  geom_bar(stat = "identity", position = "dodge") +  
  labs(title = "Durchschnittliche Blickdauer auf die Objekte",  
        x = "AOI",  
        y = "Blickdauer (millisecond)") +
```

```
theme_minimal() +
scale_fill_brewer(palette = "Set2")+
theme(legend.position = "none",
      axis.title.x = element_blank(),
      axis.text.x = element_text(size=12),face="bold")
```



9 Basisanalyse führen für Haupteffekts

- Hier möchten wir wissen, ob es einen signifikanten Unterschied in der Blickdauer auf interessante (AOI = IO) und uninteressante Objekte (AOI = UO) gibt.

```
# Deshalb müssen wir daten auf UV aggregieren
agg_data<-aggregate(GazeEventDuration ~ AOI , data = data, FUN = mean)

# Haupteffekt zeigen (Mittelwert IO - Mittelwert UO)
Effekt<-agg_data$GazeEventDuration[agg_data$AOI=="Interessantes Objekt"] -
agg_data$GazeEventDuration[agg_data$AOI=="Uninteressantes Objekt"]
```

```
print(Effekt)
```

```
## [1] 90.24863
```

```
# Signifikanztest durchführen
```

```
# T-Test
```

```
t_test_result <- t.test(GazeEventDuration ~ AOI, data = data)
```

```
# T-Test-Ergebnisse anzeigen
```

```
print(t_test_result)
```

```
##
```

```
## Welch Two Sample t-test
```

```
##
```

```
## data: GazeEventDuration by AOI
```

```
## t = 33.688, df = 2254.1, p-value < 2.2e-16
```

```
## alternative hypothesis: true difference in means between group Interessantes Objekt a
```

```
## 95 percent confidence interval:
```

```
## 84.99511 95.50214
```

```
## sample estimates:
```

```
## mean in group Interessantes Objekt mean in group Uninteressantes Objekt
```

```
## 229.2437 138.9951
```

Wir können sehen, dass **Effekt** ist eine Differenz in der Blickdauer zwischen interessanten und uninteressanten Objekten. Man kann diese Werte entweder manuell berechnen (siehe oben) oder aus dem t-Test-Ergebnis entnehmen.

Der t-Test zeigt, dass dieser Unterschied statistisch signifikant ist ($p < 0.001$). Das bedeutet, dass Kinder signifikant länger auf interessante Objekte schauen als auf uninteressante Objekte, was unsere Hypothese unterstützt.

Fragen

- Was macht die Funktion `power.t.test()` in R? Welche Argumente sind wichtig und was bedeuten sie?
- Was macht die Funktion `t.test()` in R? Welche Argumente sind wichtig ein t-test zu führen?
- Was zeigt der Differenzwert zwischen zwei Stufen des Faktors?

Literatur

- Ito, A., Knoeferle, P. Analysing data from the psycholinguistic visual-world paradigm: Comparison of different analysis methods. Behav Res 55, 3461–3493 (2023).