

Beobachtung und Experiment

Literatur:

Myers, D.G. (1993). *Social Psychology, 4th Edition*.
New York: McGraw-Hill (pp. 12-23).

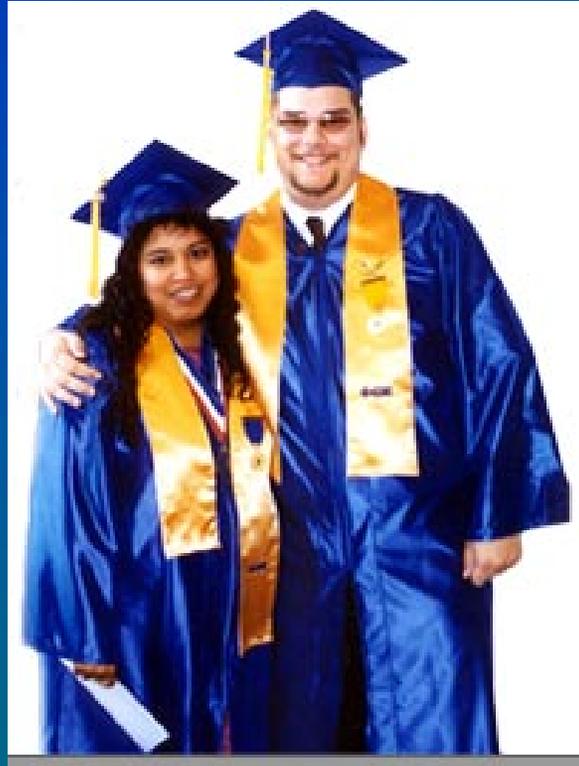
Maxwell, S.E. & Delaney, H.D. (1990). *Designing
Experiments and Analyzing Data: A Model Comparison
Approach*.

Ein Forschungsproblem:



"Wird man besser bezahlt, wenn man einen Collegeabschluß hat?"

Und wenn ja...:



"Ist der Collegeabschluß die *Ursache* für die bessere Bezahlung?"

Einige mögliche Strategien, diese Frage zu untersuchen:

- Umfragen und Fragebogen-Untersuchungen
- Feld- und Korrelationsstudien
- Experimente

Umfragen und Fragebögen



"Glauben Sie, daß die Qualität Ihrer Ausbildung Ihr Einkommen beeinflußt hat?"

JA

NEIN

"Glauben Sie an einen generellen Zusammenhang zwischen Ausbildung und Einkommen?"

JA

NEIN

Repräsentative Stichproben

- Da man nicht die gesamte Population der Menschen untersuchen kann, muß man aus dieser Population eine **Stichprobe** ziehen
- Stichproben in Fragebogenerhebungen müssen **sehr groß** sein (oft Tausende von Vpn)
- Stichprobe muß **repräsentativ** sein: muß die Grundgesamtheit widerspiegeln in Bezug auf Alter, Geschlecht, soziale Schicht, Ausbildungsgrad, ethnische Herkunft etc. etc.
- **Jede Stichprobe läßt nur Aussagen über diejenigen Populationen zu, für die sie repräsentativ ist!**

Häufiger Fehler: Sampling Bias

Literary Digest schickte eine Umfrage zur Präsidentschaftswahl von 1936 an 10 Millionen Haushalte.

Ergebnisse sprachen für einen klaren Wahlsieg von Alf Landon über Franklin D. Roosevelt, obwohl Roosevelt später klar gewann.

Problem: die Zeitschrift hatte den Fragebogen anhand von Telefonbüchern und KFZ-Zulassungen verschickt und damit alle Haushalte ausgelassen, die sich so etwas nicht leisten konnten!

Ein anderes Beispiel: Psychologische Experimente

Die meisten Experimente in der Psychologie werden mit Psychologie-Studierenden als Versuchspersonen durchgeführt

Die Ergebnisse sind also nicht generalisierbar auf andere Teilpopulationen: z.B. Menschen...

- ...mit höherem Alter;
- ...mit geringerer Intelligenz;
- ...mit niedrigerem Ausbildungsstand;
- ...mit geringerer Computererfahrung;
- ...mit höherem sozioökonomischem Status
- usw. usf.

Weitere Probleme:

Rücklaufquote

Der Anteil von Vpn, die nicht an der Umfrage teilnehmen möchten, kann systematisch von Merkmalen der Vpn abhängen

z.B. sehr alte, sehr junge, sehr kranke Menschen

z.B. Menschen mit sehr extremen Überzeugungen

Reihenfolge der Fragen

A: "Sollte die japanische Regierung den Import von amerikanischen Gütern limitieren?"

B: "Sollte die amerikanische Regierung den Import von japanischen Gütern limitieren?"

Erst A, dann B: die meisten Amerikaner lehnen A ab

Erst B, dann A: die meisten Amerikaner stimmen zu

(Schuman & Ludwig, 1983)

Effekt der Antwortoptionen

"Was ist zur Zeit das größte Problem in diesem Land:

- *Energieknappheit*
- *Qualität der öffentlichen Schulen*
- *Legalisierte Abtreibung*
- *Umweltverschmutzung?"*

32 % antworten "Qualität der Schulen"

"Was ist zur Zeit das größte Problem in diesem Land?"

nur 1% erwähnt die Schulen!

Schuman & Scott (1987)

Wortlaut der Fragen

"Sollte die Regierung öffentliche Reden gegen die Demokratie verbieten?"

54 % sagen "ja"

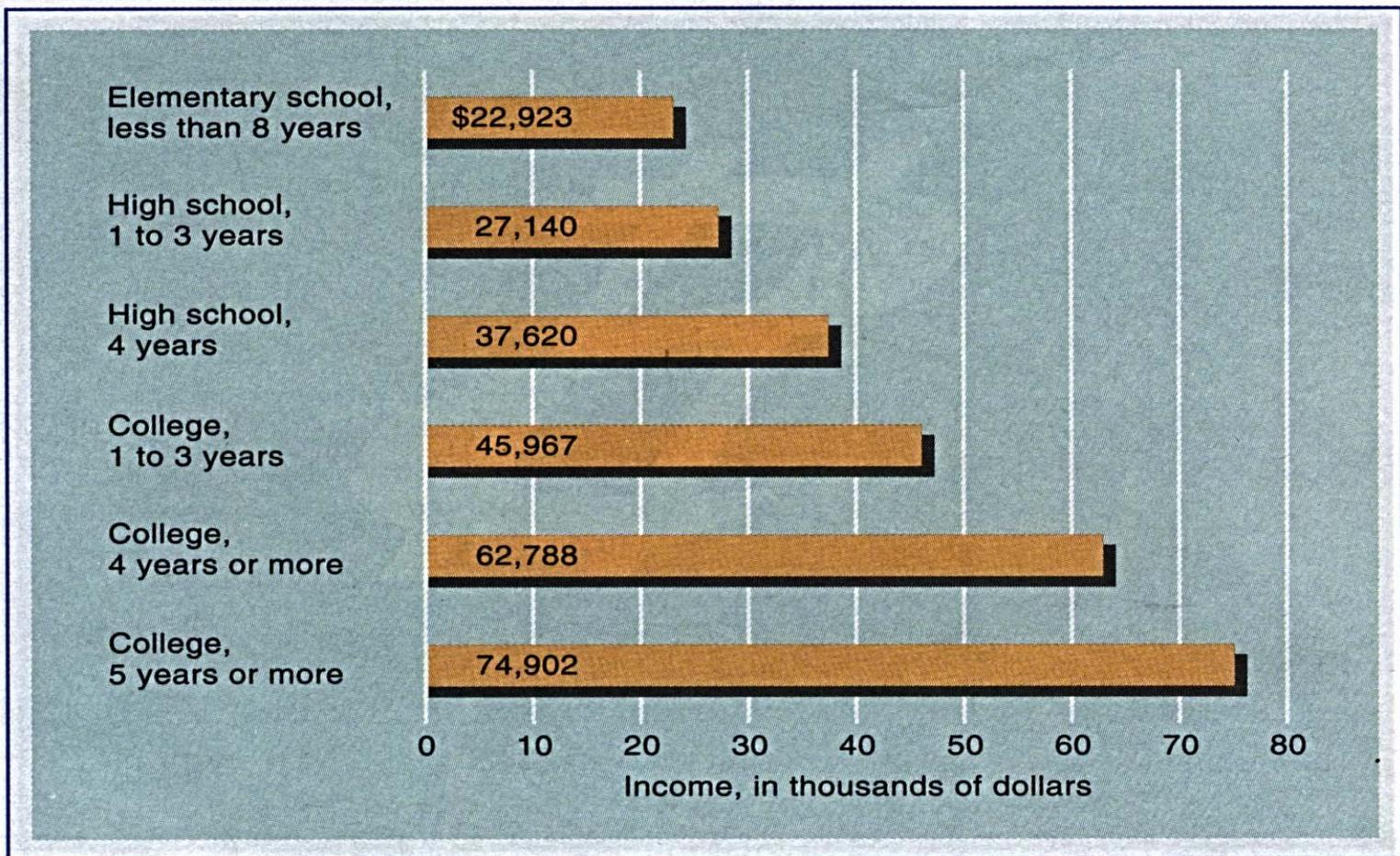
"Sollte die Regierung öffentliche Reden gegen die Demokratie erlauben?"

75 % sagen "nein"

nach Schuman & Kalton (1985)

Korrelationsstudien

"Gibt es überhaupt eine Beziehung zwischen Ausbildung und Bezahlung?"



Korrelation und Kausalität: der Korrelations-Koeffizient

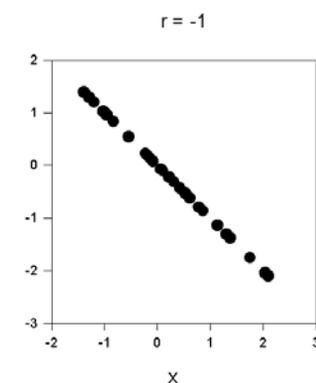
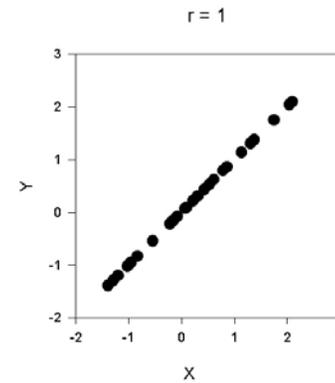
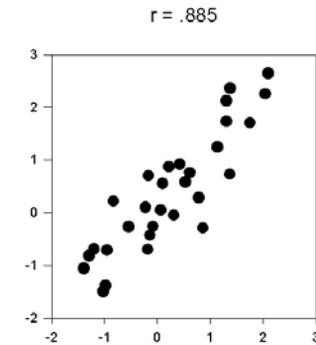
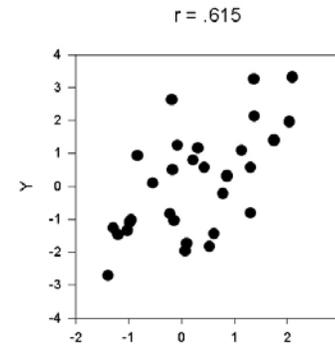
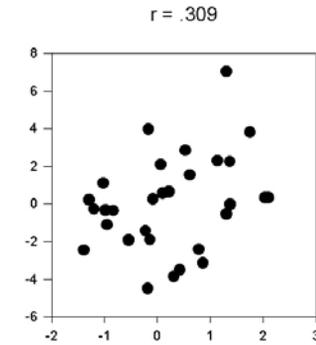
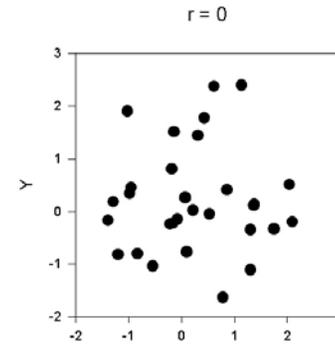
$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})}{n \cdot s_x \cdot s_y}$$

$r = 1$ bei perfekter positiver Korrelation

$r = -1$ bei perfekter negativer Korrelation

$r = 0$, wenn gar keine Korrelation vorliegt

Einige Beispiele für Korrelationen:



Problem: es gibt verschiedene kausale Wege, auf denen eine Korrelation zustande kommen kann !

1) X verursacht Y:



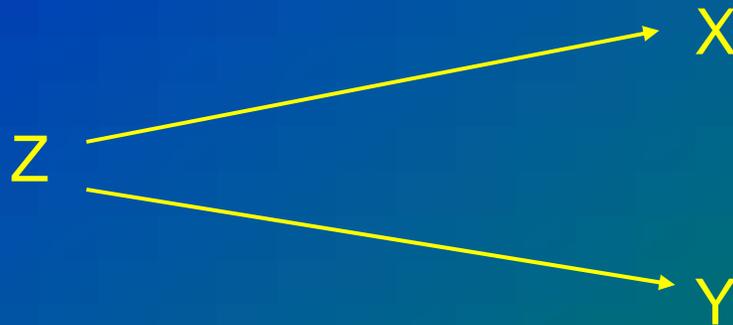
"Höhere Schulbildung verursacht ein höheres Einkommen."

2) Y verursacht X:



z.B. "Schüler aus Familien mit höherem Einkommen bekommen eine bessere Schulbildung."

3) Z verursacht sowohl X als auch Y:



z.B. "Schüler mit höherer Intelligenz haben sowohl eine bessere Schulbildung als auch ein höheres Einkommen."

Landers (1988):

Es gibt eine positive Korrelation zwischen der Vorliebe für **Heavy-Metal-Musik** und positiven Einstellungen zu:

Pornographie...

Drogen- und Alkoholgebrauch...

Satanismus...

...und vorehelichem Sex!

Wie kann man das erklären?





Unser erstes Experiment!

„Gibt es Geschlechterunterschiede bei der Wahrnehmung von Farben?“

Unabhängige Variable:
Geschlecht

Abhängige Variable:
Geschwindigkeit bei der Erkennung von Farben

ROT

GELB

BLAU

GELB

GRÜN

ROT

GRÜN

BLAU

BLAU

BLAU

GELB

GELB

GELB

ROT

ROT

GRÜN

BLAU

GRÜN

GRÜN

BLAU

GRÜN

GELB

ROT

ROT

BLAU

ROT

GELB

GELB

GRÜN

GRÜN

ROT

BLAU

GELB

BLAU

BLAU

GELB

ROT

GELB

ROT

GRÜN

GRÜN

BLAU

GRÜN

BLAU

ROT

GRÜN

GELB

ROT

GELB

ROT

GELB

BLAU

BLAU

GRÜN

ROT

GRÜN

GELB

BLAU

BLAU

GELB

GRÜN

GELB

ROT

ROT

BLAU

BLAU

GRÜN

GRÜN

ROT

GRÜN

GELB

ROT

ROT

GELB

BLAU

GELB

GRÜN

ROT

GRÜN

BLAU

BLAU

BLAU

GELB

GELB

GELB

ROT

ROT

GRÜN

BLAU

GRÜN

GRÜN

BLAU

GRÜN

GELB

ROT

ROT

BLAU

GELB

GELB

ROT

GRÜN

BLAU

ROT

GRÜN

GELB

GELB

BLAU

BLAU

ROT

GRÜN

ROT

GELB

GRÜN

BLAU

GRÜN

BLAU

ROT

ROT

GELB

GRÜN

BLAU

ROT

GELB

GELB

GRÜN

GRÜN

ROT

BLAU

GELB

BLAU

BLAU

GELB

ROT

GELB

ROT

GRÜN

GRÜN

BLAU

GRÜN

BLAU

ROT

GRÜN

GELB

ROT

nature neuroscience

volume 5 no 5 may 2002

<http://neurosci.nature.com>

Special Issue: Complete colorblindness in human males

Behavioral scientists solve the mystery
why men fail in color perception tasks

ROT

GELB

BLAU

GELB

GRÜN

ROT

GRÜN

BLAU

BLAU

BLAU

GELB

GELB

GELB

ROT

ROT

GRÜN

BLAU

GRÜN

GRÜN

BLAU

GRÜN

GELB

ROT

ROT

Farbwörter und Druckfarbe stimmen überein!

BLAU

ROT

GELB

GELB

GRÜN

GRÜN

ROT

BLAU

GELB

BLAU

BLAU

GELB

ROT

GELB

ROT

GRÜN

GRÜN

BLAU

GRÜN

BLAU

ROT

GRÜN

GELB

ROT

Farbwörter und Druckfarbe stimmen nicht überein!

Was ist ein Experiment?

- kontrollierte Variation der Meßbedingungen:
Manipulation der **unabhängigen Variablen**
- geeignet für Aussagen über **kausale Einflüsse** von unabhängigen Variablen auf abhängige Variablen
- **zufällige Zuordnung** von Versuchspersonen zu Meßbedingungen (sonst „Quasi-Experiment“)

Grundstruktur eines Experiments

- unabhängige Variablen $X_1 \dots X_n$ werden systematisch verändert
- abhängige Variablen $Y_1 \dots Y_m$: werden gemessen

Idee:

Wenn die Manipulation der unabhängigen Variablen **zuverlässig** zu einer Veränderung in den abhängigen Variablen führt, **dann** kann man auf einen kausalen Einfluß der unabhängigen Variablen schließen!

Back to College:

Treatment:

abhängige Variable:

College
(Experimental-
gruppe)

\$

kein College
(Kontrollgruppe)

\$

Vpn

Zufällige
Zuweisung

Artefakte

Unser College-Experiment führt zu irreführenden Ergebnissen, wenn sich die beiden Gruppen schon vor der experimentellen Manipulation unterschieden haben:

- War die College-Gruppe vielleicht schon vorher reicher?
- Waren die Eltern reicher?
- Hatten die Eltern eine höhere Schulbildung?

Solche Fehlerquellen nennt man **Störvariablen**.

Es gibt drei Techniken zur Kontrolle solcher Störeinflüsse:

- Kontrollieren
- Konstanthalten
- Randomisieren

Kontrollieren

Statt die Störvariable so hinzunehmen, wie sie ist, variiert man sie absichtlich.

Beispiel:

Wir unterteilen das Bildungsniveau der Eltern in drei Stufen und untersuchen systematisch Kinder von niedrig-, mittel- und hochgebildeten Eltern.

Nachteil:

Dadurch brauchen wir auch dreimal so viele Versuchspersonen!

Verwandte Technik: geschichtete Stichproben

Wir achten darauf, daß sich in allen Versuchsbedingungen vergleichbare Gruppen befinden – so entstehen „repräsentative“ Stichproben, die wir selbst erzeugt haben.

Beispiel:

Für jede Versuchsperson in der einen Gruppe suchen wir eine Versuchsperson mit vergleichbaren Werten der Störvariable und weisen sie der anderen Gruppe zu.

Nachteil:

Auch dafür brauchen wir sehr viele Versuchspersonen, sonst werden die Stichproben nicht repräsentativ.

Konstanthalten

Wir sorgen dafür, daß die Störvariable immer denselben Wert annimmt.

Beispiel:

Wir untersuchen nur Kinder von Eltern mit mittlerem Bildungsgrad.

Nachteil:

Damit schränken wir die Generalisierbarkeit unserer Ergebnisse ein!

Randomisieren

Wir sorgen dafür, daß die Werte der Störvariablen nur durch reinen Zufall bestimmt werden – die Versuchsbedingungen sollten dann auf lange Sicht ungefähr vergleichbar sein.

Beispiel:

Wir bestimmen zufällig, welche Eltern in die Studie aufgenommen werden, und verteilen dann deren Kinder zufällig auf die beiden Versuchsgruppen.

Nachteil: Das funktioniert nur bei großen Stichproben!

Ethische Probleme beim Experimentieren

- einige Experimente können Vpn möglicherweise **Schaden zufügen** (z.B. Zuordnung zu schlechten Ausbildungsbedingungen; Zuordnung zu einer Kontrollgruppe, die keine Therapie erhält)
- einige Experimente erfordern die **Täuschung** von Vpn (z.B. Forschung zu Hilfeleistung bei Unfällen)
- einige Experimente sind **unangenehm, anstrengend oder langweilig**



Ethische Richtlinien der American Psychological Association (APA 1981, 1992):

- Vpn müssen genug Informationen erhalten, um entscheiden zu können, ob sie überhaupt teilnehmen wollen („informed consent“)!
- Täuschung sollte nur dann benutzt werden, wenn es wirklich keine andere Alternative gibt!
- Vpn müssen vor Schaden oder unangenehmen Begleiterscheinungen geschützt werden!
- Informationen über Vpn müssen vertraulich behandelt werden!
- Vpn müssen hinterher vollständig aufgeklärt werden!

Experimente finden meist im Labor statt und basieren auf einem vereinfachten Modell der Wirklichkeit.

Problem: Kann eine Laborsituation die Wirklichkeit überhaupt angemessen abbilden?

⇒ Frage nach der "**ökologischen Validität**" von Experimenten

Ökologisch valide Experimente?

Vorteile:

- leichtere Generalisierung auf Situationen außerhalb des Labors

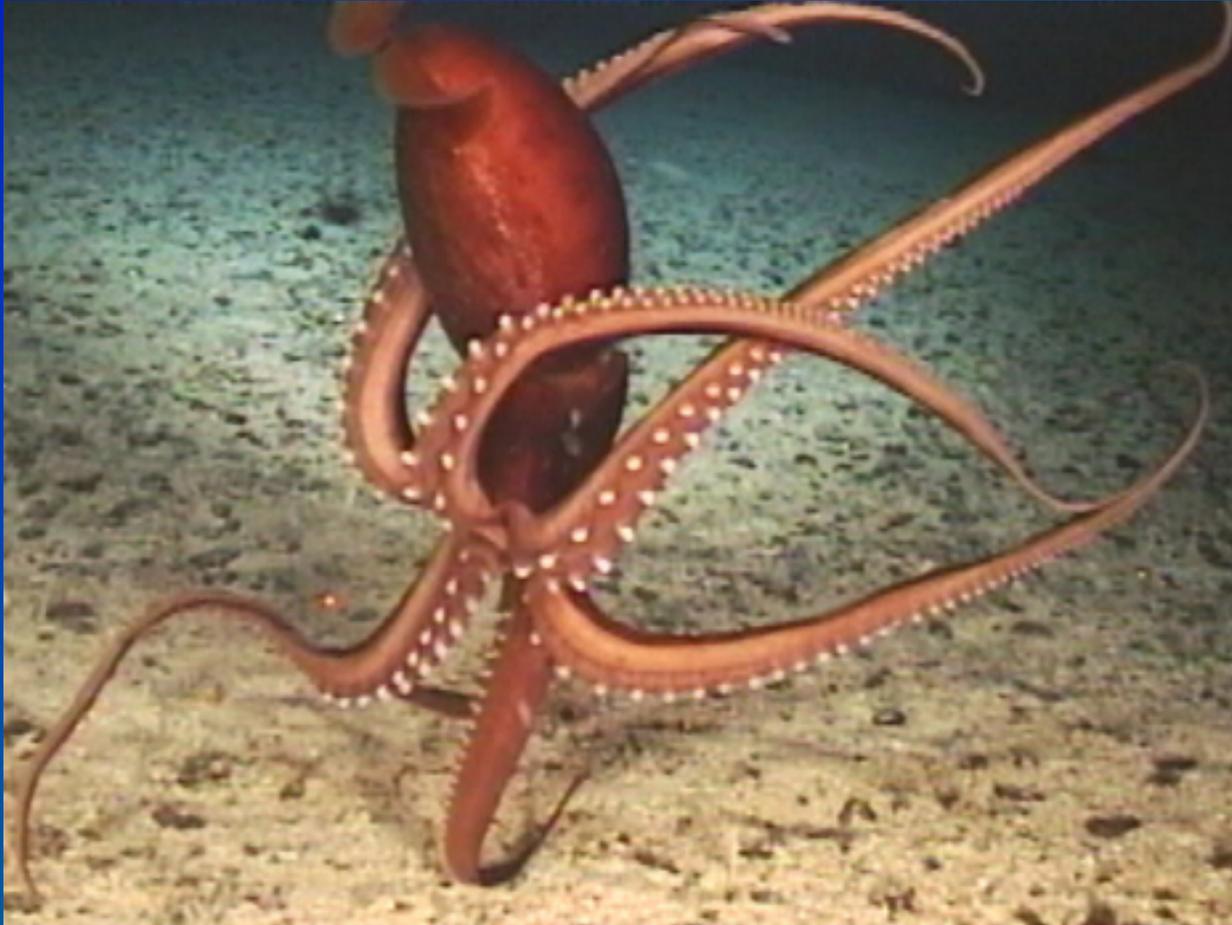
Nachteile:

- meist Zunahme von Störeinflüssen
- meist Zunahme von Alternativerklärungen

also:

Das Experiment hat zwar eine höhere "intuitive Gültigkeit", ist aber logisch weniger zwingend!

„A Study in Squid“



Wer hat mehr Erfolg bei den Weibchen:
dunkelbraune Männchen oder hellgraue Männchen?

Theorie:

„Wir wissen bereits, daß Krakenmännchen im Laufe ihres Lebens die Farbe wechseln – von dunkelbraun zu hellgrau. Wir glauben außerdem, daß Krakenweibchen jüngere Männchen bevorzugen. Außerdem glauben wir, daß sie diese Unterscheidung aufgrund der Färbung des Männchens treffen.“

Hypothese:

„Krakenweibchen sollten dunkelbraune Männchen gegenüber hellgrauen Männchen bevorzugen.“

Jedes Experiment hat einen **Versuchsplan**
(„Design“):

Was wird gemessen?

⇒ Abhängige Variable(n) (AV)

Was wird variiert?

⇒ Unabhängige Variable(n) (UV, „Faktoren“)

Welche Versuchspersonen sehen welche Bedingungen?

⇒ Meßwiederholungsstruktur des Experiments

Unser erstes Versuchsdesign:

AV: Zahl der Kontakte zwischen Männchen und Weibchen

UV: Farbe des Männchens (dunkelbraun oder hellgrau)

Meßwiederholungsstruktur: es soll zwei Gruppen von Männchen geben, eine für jede Bedingung

Gruppe 1:
dunkelbraun

Gruppe 2:
hellgrau

Wie bilden wir die Gruppen?

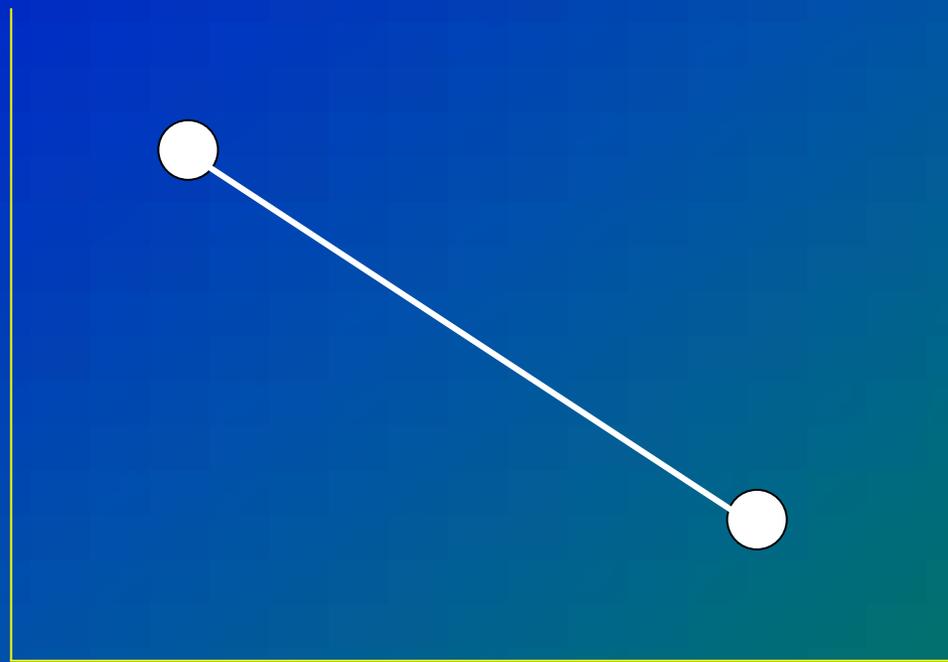
Eine Möglichkeit: Wir sortieren Krakenmännchen nach Farbe und stecken die „dunkelbraunsten“ in die eine Gruppe und die „hellgrauesten“ in die andere.

Problem: Dann haben wir die Männchen nicht zufällig den Versuchsbedingungen zugewiesen – es könnten sich systematische Unterschiede zwischen den Gruppen ergeben.

Besser: Zufällige Zuweisung zu den Versuchsbedingungen

- jedes Männchen wird zufällig einer Versuchsgruppe zugewiesen;
- jedes Männchen wird entsprechend eingefärbt...!

AV:
Paarungs-
erfolg



dunkelbraun

hellgrau

UV: Farbe des Männchens

Hypothese bestätigt: dunkelbraune Männchen sind erfolgreicher!

Meßwiederholung

Oft kann man dieselbe Gruppe von Versuchspersonen oder Versuchstieren in mehreren Bedingungen des Experiments testen!

Vorteile:

- Unterschiede zwischen Bedingungen sind nicht mit möglichen Unterschieden zwischen den Stichproben konfundiert
- höhere Präzision der Messung, da man Unterschiede zwischen Versuchspersonen aus den Daten „herausrechnen“ kann

Nachteile:

- mögliche Ermüdungs-, Trainings, Carry-Over-Effekte

Wir replizieren unser erstes Experiment:

AV: Zahl der Kontakte zwischen Männchen und Weibchen

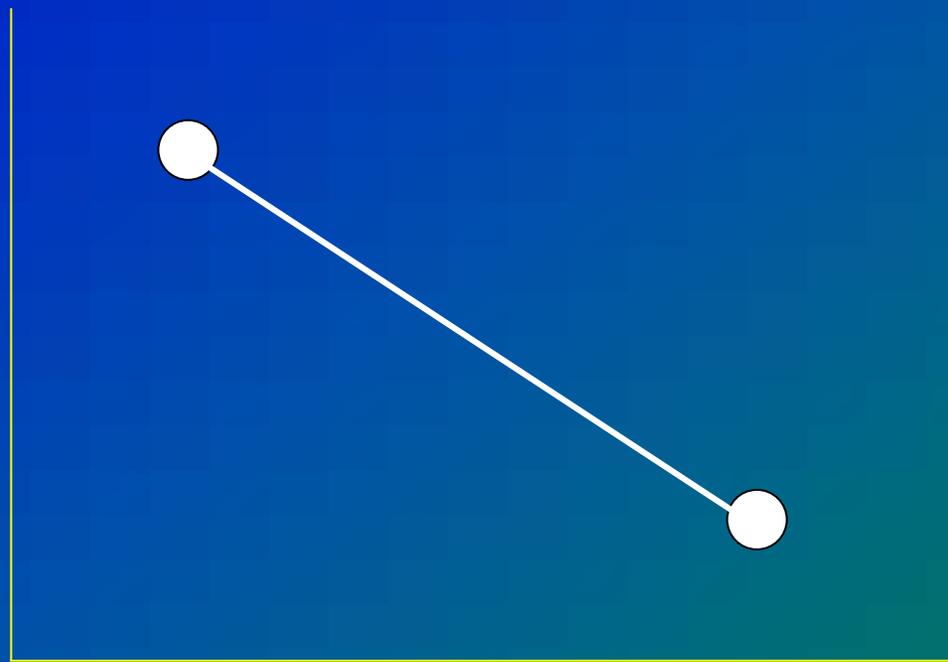
UV: Farbe des Männchens (dunkelbraun oder hellgrau)

Meßwiederholungsstruktur: diesmal soll jedes Versuchstier in beiden Versuchsbedingungen mitmachen.

Dazu verwenden wir die Technik der **gekreuzten Listen:**

- Es gibt nur noch eine Gruppe von Kraken, aber die Kraken werden zwei verschiedenen **Listen** (a und b) zugewiesen.
- Wir führen den Versuch zweimal durch, mit ein paar Tagen Abstand, damit sich die Kraken von den Strapazen erholen können.
- Am ersten Meßzeitpunkt färben wir Liste a) hellgrau und Liste b) dunkelbraun. Am zweiten Meßzeitpunkt machen wir es umgekehrt.

AV:
Paarungs-
erfolg



dunkelbraun

hellgrau

UV: Farbe des Männchens

Hypothese bestätigt: dunkelbraune Männchen sind erfolgreicher!

Unser Experiment war ein Erfolg!

Jetzt können wir einen Artikel beim renommierten „Journal of Deep Sea Monsters“ einreichen – ein weiterer Schritt hin zu einer der begehrten Professuren im Fach „Experimentelle Krakologie“.

Gerade noch rechtzeitig macht uns eine Kollegin darauf aufmerksam, daß Krakenweibchen vielleicht gar keine Farbe sehen können!

Können die Weibchen also tatsächlich die Farbe der Männchen unterscheiden, oder richten sie sich ausschließlich nach deren Helligkeit?

Wie sieht ein Design mit zwei Faktoren aus?

- Faktor A: Farbe (braun, grau)
- Faktor B: Helligkeit (hell, dunkel)
- Eine Gruppe von Versuchstieren pro Bedingung

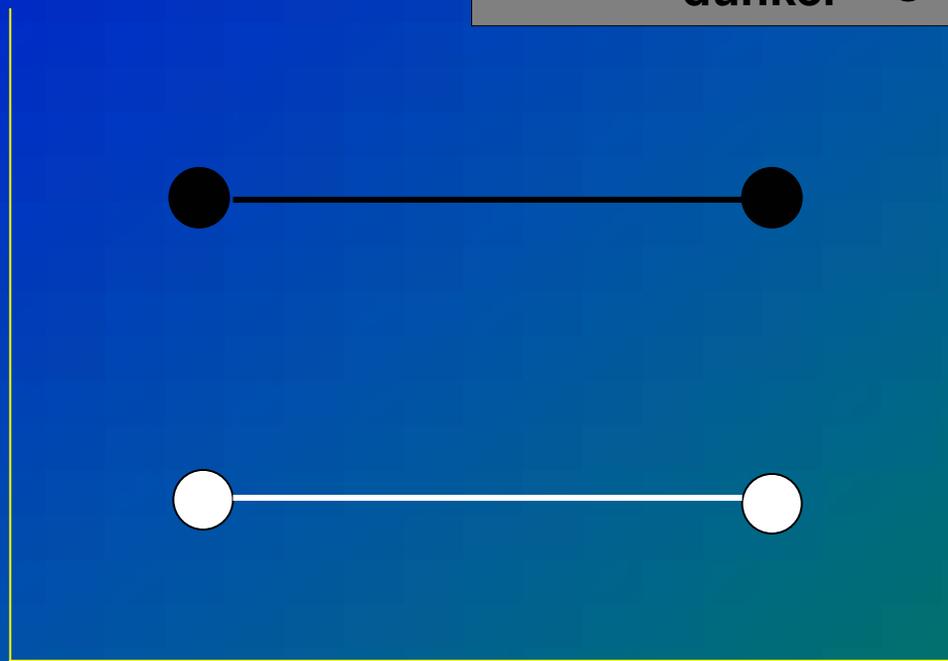
	A ₁	A ₂
B ₁	Gruppe 1: braun + dunkel	Gruppe 2: grau + dunkel
B ₂	Gruppe 3: braun + hell	Gruppe 4: grau + hell

AV:
Paarungs-
erfolg

UV 2: Helligkeit des Männchens:

hell ○

dunkel ●



braun

grau

UV 1: Farbe des Männchens

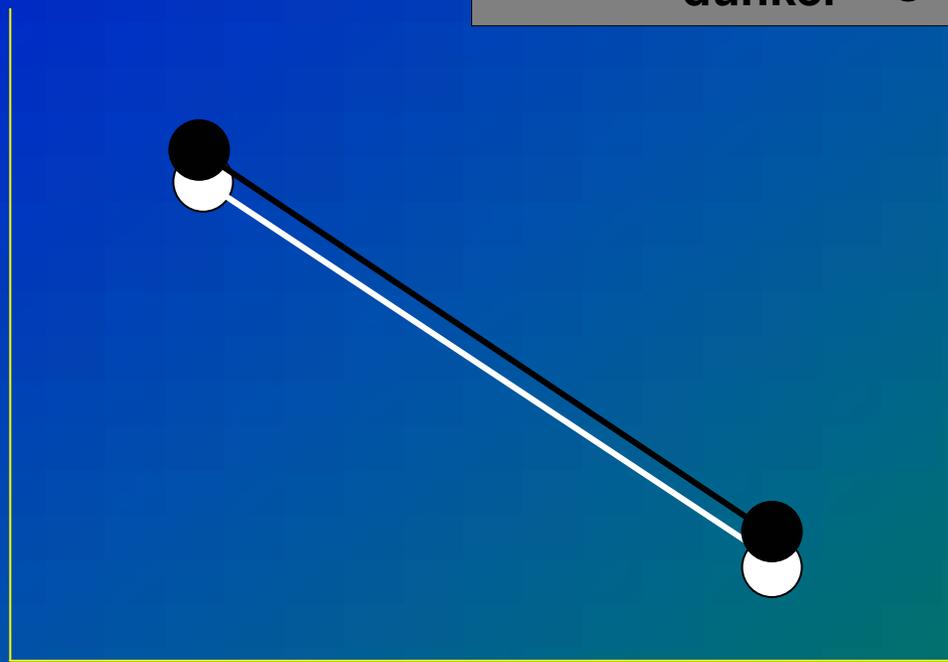
Farbe egal, aber dunkle Männchen sind erfolgreicher.

AV:
Paarungs-
erfolg

UV 2: Helligkeit des Männchens:

hell ○

dunkel ●



braun

grau

UV 1: Farbe des Männchens

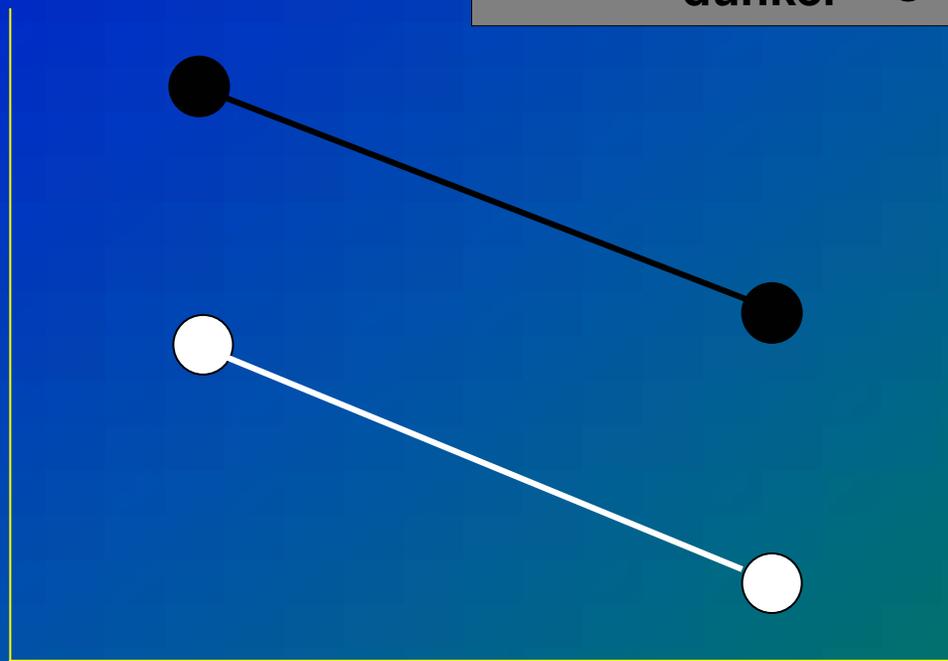
Helligkeit egal, aber braune Männchen sind erfolgreicher.

AV:
Paarungs-
erfolg

UV 2: Helligkeit des Männchens:

hell ○

dunkel ●



braun grau

UV 1: Farbe des Männchens

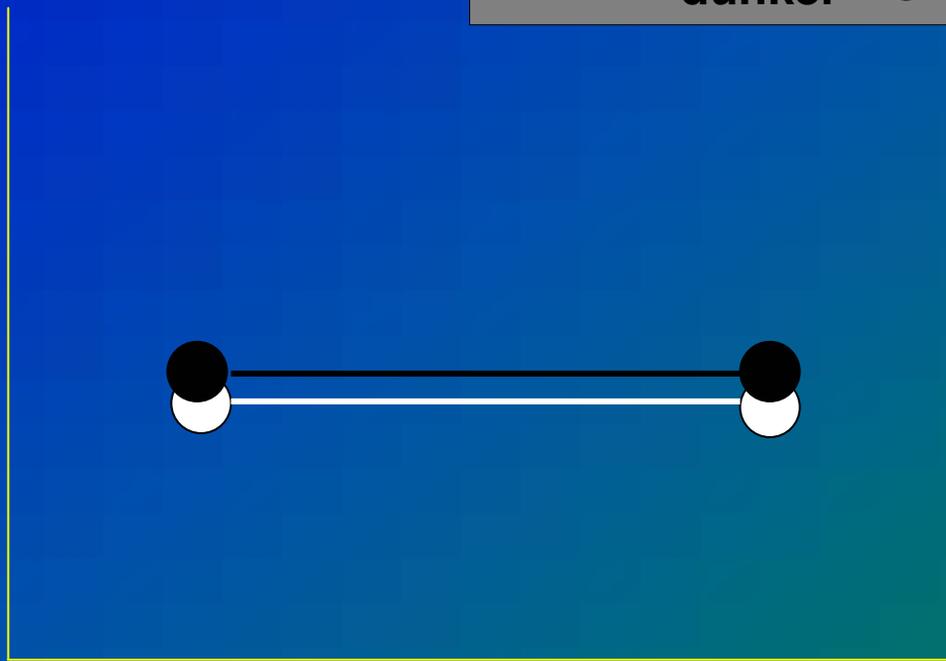
Sowohl Helligkeit als auch Farbe spielen eine Rolle.

AV:
Paarungs-
erfolg

UV 2: Helligkeit des Männchens:

hell ○

dunkel ●



braun

grau

UV 1: Farbe des Männchens

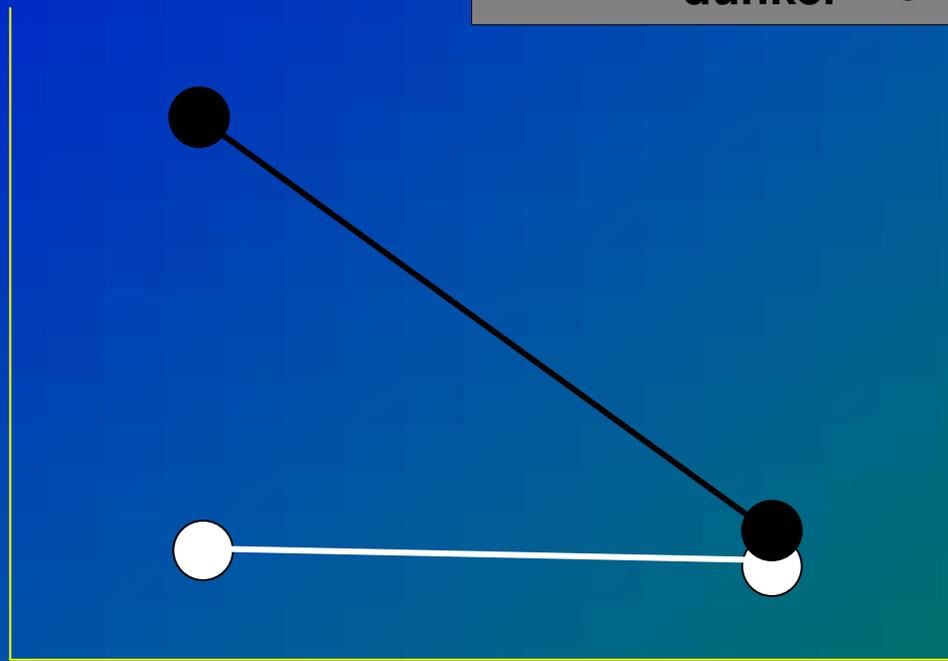
Kein Effekt: Farbe egal, Helligkeit egal.

AV:
Paarungs-
erfolg

UV 2: Helligkeit des Männchens:

hell ○

dunkel ●



UV 1: Farbe des Männchens

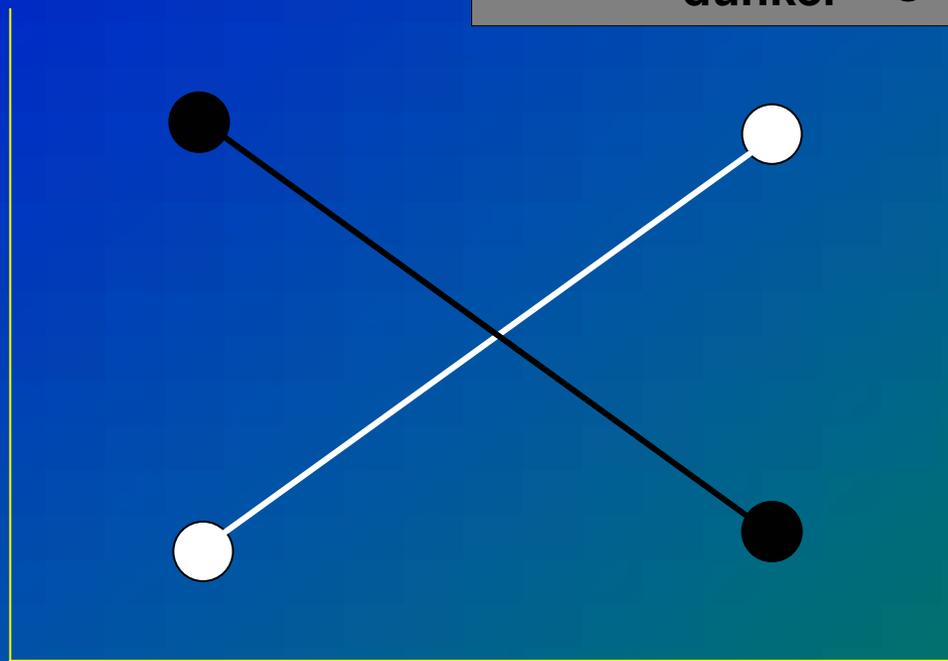
Nur die **Kombination** von Braun und Dunkel hat mehr Erfolg!

AV:
Paarungs-
erfolg

UV 2: Helligkeit des Männchens:

hell ○

dunkel ●



UV 1: Farbe des Männchens

Nur dunkelbraune und hellgraue Männchen haben Vorteile!



Haupteffekte und Wechselwirkungen

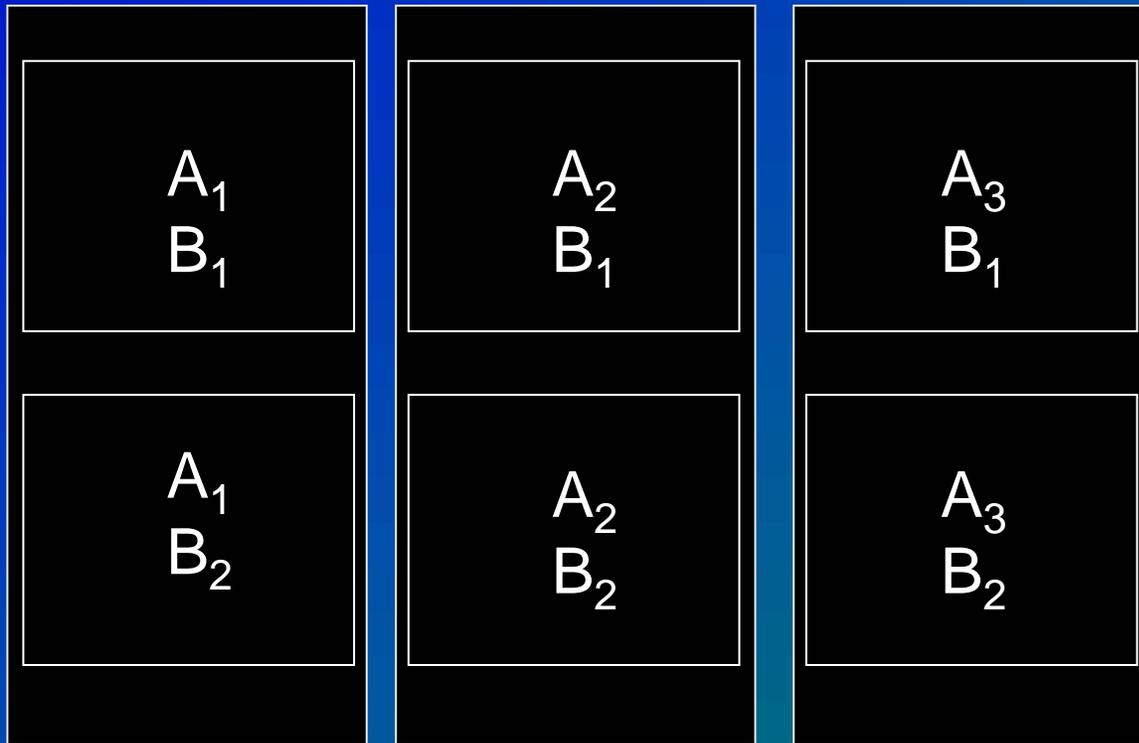
In Designs mit mehr als zwei Faktoren können zwei Arten von Effekten auftreten:

Haupteffekte

"Sind alle Stufen eines Faktors gleich, wenn man über alle anderen Faktoren mittelt?"

Wechselwirkungen (Interaktionen)

"Hängen die Werte auf den Stufen des einen Faktors von den Stufen der anderen Faktoren ab?"

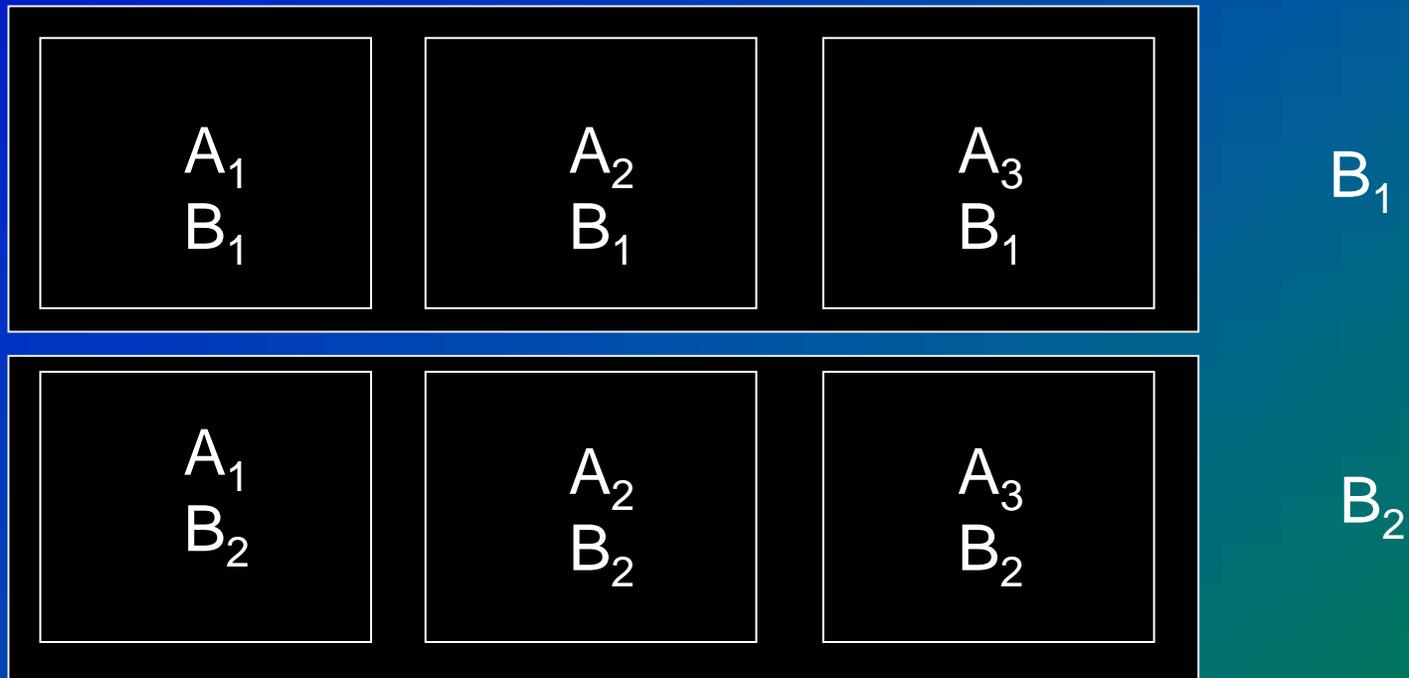


A_1

A_2

A_3

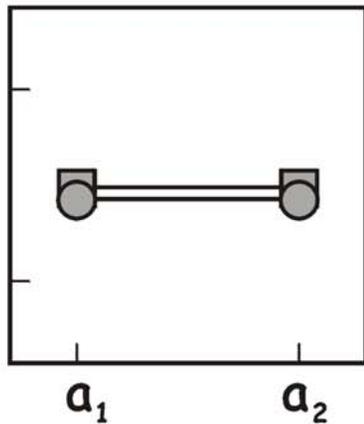
Haupteffekt des Faktors A



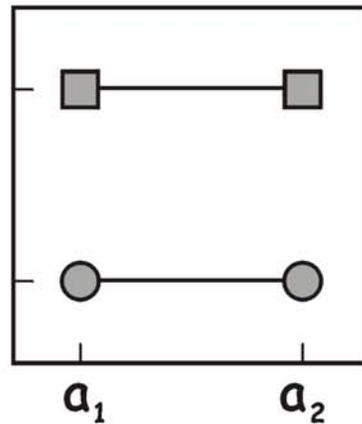
Haupteffekt des Faktors B

2x2-Designs ohne Wechselwirkungen:

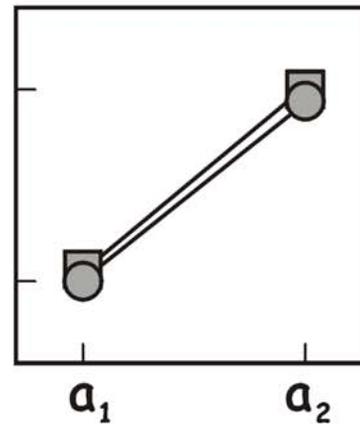
■ b_1 ● b_2



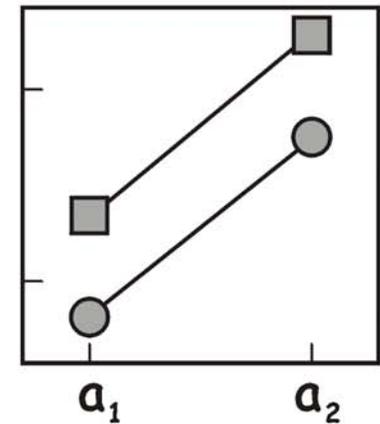
keine
Haupteffekte:



nur Haupteffekt
B:



nur Haupteffekt
A:

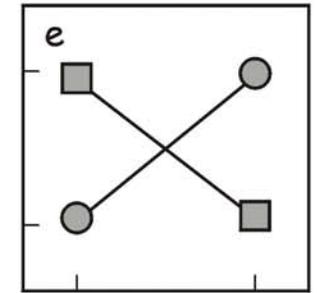
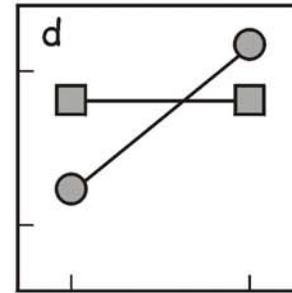
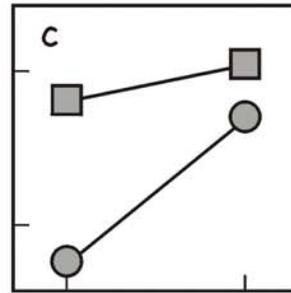
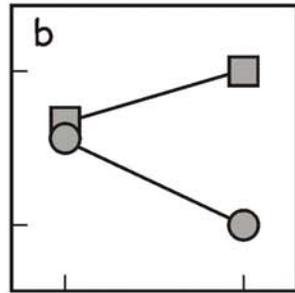
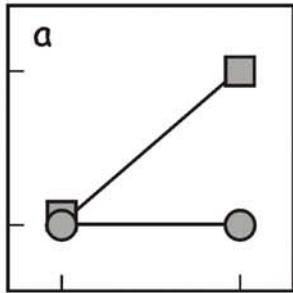


beide
Haupteffekte

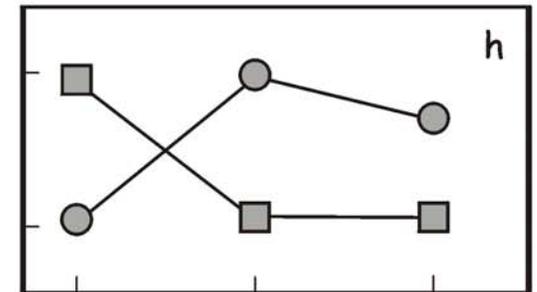
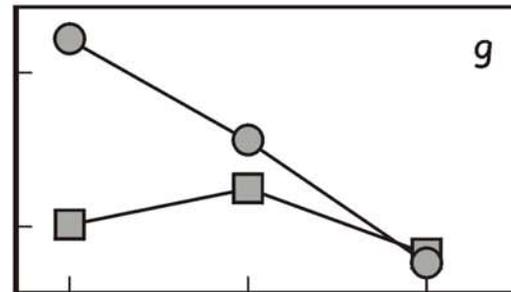
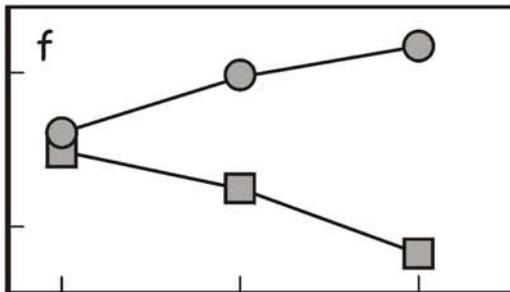
In einem Liniendiagramm erkennt man Wechselwirkungen daran, daß die Linien nicht parallel verlaufen!

Parallele Linien bedeuten, daß die Haupteffekte **unabhängig** voneinander wirken, sich also einfach addieren.

Wechselwirkungen in 2x2-Designs:



Wechselwirkungen in 2x3-Designs:



Nicht jedes Experiment besteht aus einer kleinen Anzahl von Versuchsbedingungen, die einzeln miteinander verglichen werden können.

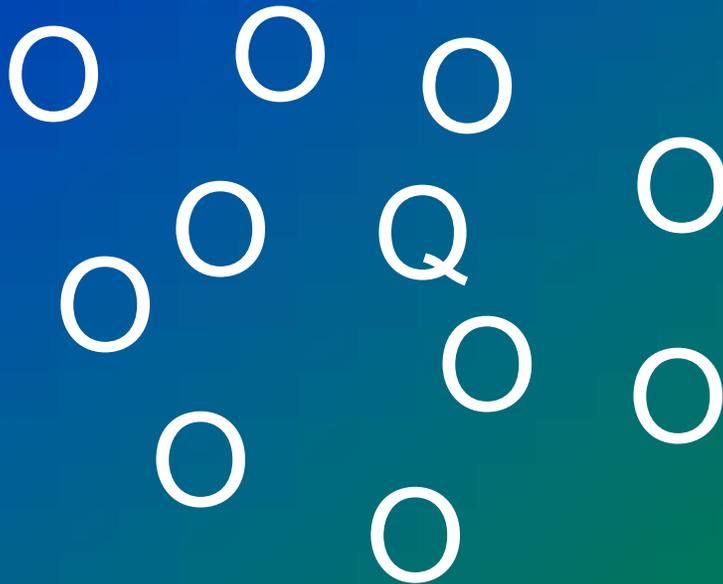
Besonders interessant sind Experimente, bei denen die Unabhängige Variable kontinuierliche Werte annehmen kann.

Mit geeigneten Modellen kann man dann quantitative Vorhersagen über das Verhalten der AV in Abhängigkeit von der UV machen.

Visuelle Suche

(z.B. Treisman & Gelade, 1980)

Target definiert durch ein einzelnes Feature:
Suche ist einfach!

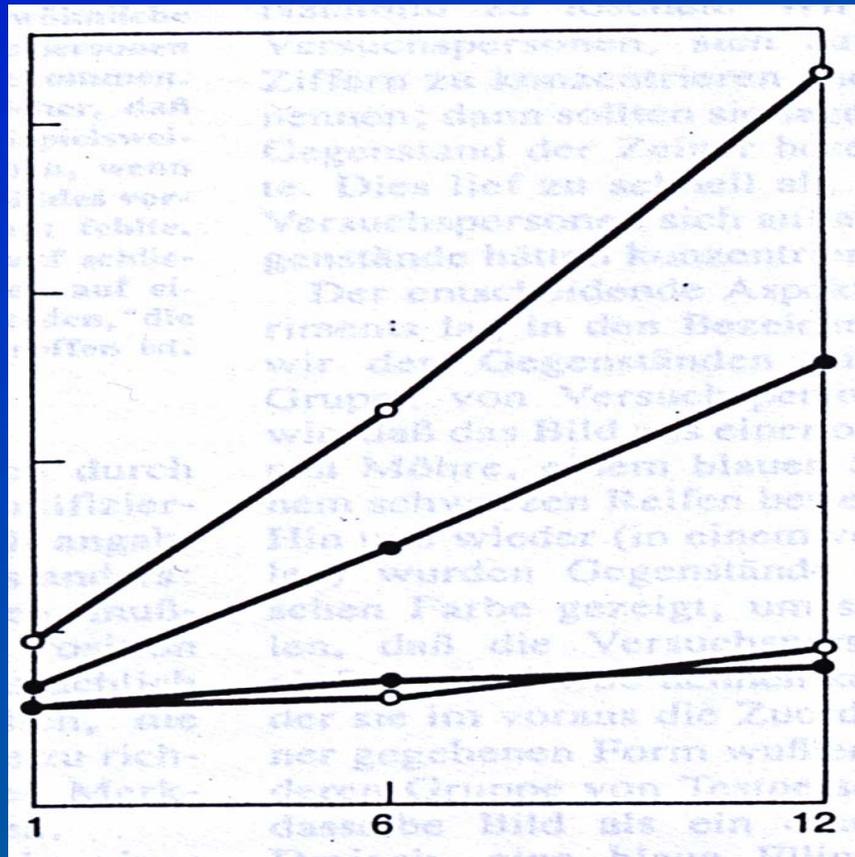


Visuelle Suche

(z.B. Treisman & Gelade, 1980)

Target definiert durch Konjunktion von Features:
Suche ist schwierig!





Target absent

Target present

Pop-Out

Treisman:

- bei Pop-Out: Suchzeit unabhängig von Displaygröße
- bei serieller Suche: bei Displays mit Target sollte die Entscheidung doppelt so schnell getroffen werden wie bei Displays ohne Target - also sollten die Suchfunktionen nur die halbe Steigung haben

