



Einführung in die Pädagogische Psychologie

08: Diagnostik und Testverfahren

Prof. Dr. Christian Fischer

12. Januar 2021



Was bisher geschah...

Diagnostik in der Pädagogischen Psychologie sammelt unter anderem Informationen über Personen, um deren Verhaltensweisen zu erklären.

RICHTIG

97,1% hatten diese Frage richtig
2,9% hatten diese Frage falsch



Diagnostik: Definition - Bestimmungsstücke

Pädagogisch-psychologische Diagnostik...

- findet im Kontext pädagogischer Handlungsfelder statt
- bezieht sich auf Einzelpersonen - Individualdiagnostik
- nutzt Methoden der Psychologie
- sammelt Informationen über Personen, um deren Verhalten und Erleben zu beschreiben, zu erklären und vorherzusagen
- hilft bei Entscheidungsfindung: Optimierung wichtiger Entscheidungen unter Unsicherheit
- ist keine Alltagsdiagnostik



Was bisher geschah...

Diagnostik in der Pädagogischen Psychologie fußt insbesondere auf subjektiven Einschätzungen eines Beurteilers.

FALSCH

98,6% hatten diese Frage richtig
1,4% hatten diese Frage falsch



Diagnostik: Aufgabe

- In all diesen Fällen geht es darum, Informationen über Personen zu sammeln, die **über subjektive Einschätzungen eines Beurteilers hinausgehen**. Oft sind damit **schwerwiegende Entscheidungen** verbunden!
- Deswegen wollen wir **begründen** können, warum wir die Schlussfolgerungen, die wir aus den gesammelten Informationen für berechtigt halten! - (wir sagen, dass unsere Informationen „valide“ sind)

Warum ist das wichtig?

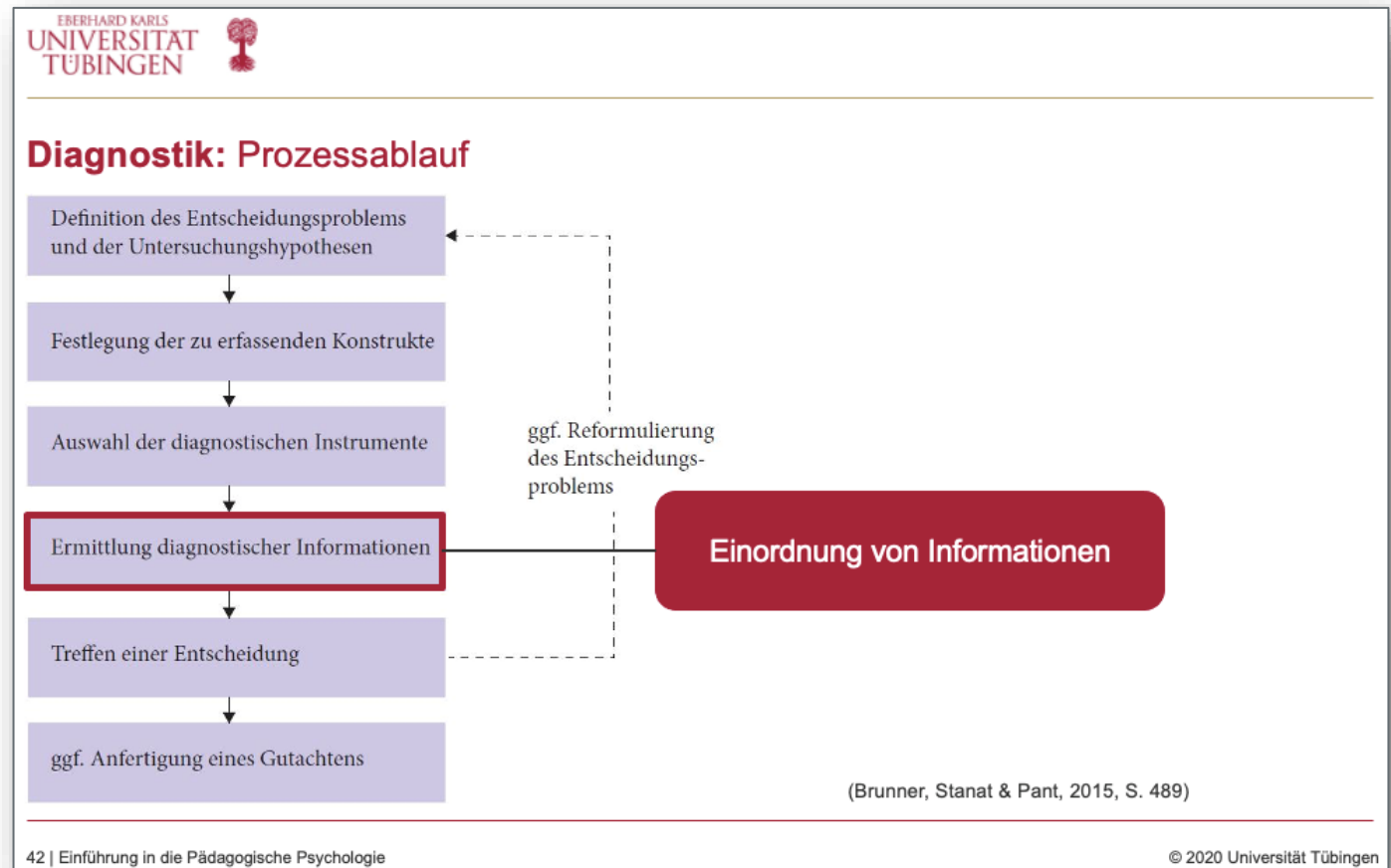


Was bisher geschah...

Die Ermittlung diagnostischer Informationen stellt ein Teil des Prozessablaufs der Diagnostik nach Brunner, Stanat und Pant (2015) dar.

RICHTIG

100% hatten diese Frage richtig
0% hatten diese Frage falsch





Was bisher geschah...

Soziale Phänomene und Konstrukte (z.B. Intelligenz, Wissen) können oftmals direkt beobachtet werden.

FALSCH

99,3% hatten diese Frage richtig
0,7% hatten diese Frage falsch



Operationalisierung: Merkmale (Konstrukte)

Problem: soziale Phänomene bzw. theoretische Größen, die nicht direkt beobachtbar sind (theoretische Konstrukte, z.B. Intelligenz, Bildungswissenschaftliches Wissen), sollen erfasst werden

- für quantitative Erhebungen müssen theoretische Konstrukte „**messbar**“ gemacht werden
- Übersetzung des Konstrukts in direkt **beobachtbare Indikatoren**, d.h. in Fragen, Aussagen, Testaufgaben, etc.
- Analyse der Antworten anhand **mathematischer Rechnungen**

Operationalisierung
=
Übersetzung theoretischer Konstrukte in messbare Indikatoren



Semesterplan

Woche	Datum	Thema
01	10.11.2020	Einführung
02	17.11.2020	Entwicklung, Sozialisation und Lernen
03	24.11.2020	Gedächtnismodelle und kognitive Basisfunktionen
04	01.12.2020	Intelligenz
05	08.12.2020	Selbstkonzept
06	15.12.2020	Motivation
07	22.12.2020	Diagnostik und Evaluation
08	12.01.2021	Diagnostik und Testverfahren
09	19.01.2021	Selbstregulation
10	26.01.2021	Lernstrategien
11	02.02.2021	Unterrichtsmodelle
12	09.02.2021	Unterrichtsqualität
13	16.02.2021	Digitale Technologien
14	23.02.2021	Wiederholung
15	02.03.2021	Klausur



Evaluation

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN
Universität Tübingen

Kontrastmodus aktivieren

Liebe Studierende,
der Fragebogen soll dazu beitragen, die Qualität der Lehre zu beurteilen und zu verbessern. Wir bitten Sie daher, diesen Fragebogen sorgfältig auszufüllen. Lesen Sie die folgenden Aussagen und geben Sie an, wie stark Sie persönlich zustimmen. Sie können jeweils zwischen fünf Abstufungen wählen, die von "trifft voll und ganz zu" bis "trifft gar nicht zu", reichen.
Die Teilnahme ist freiwillig. Die Befragung wie auch die Auswertung unterliegen datenschutzrechtlichen Bestimmungen.
Vielen Dank für Ihre Mitwirkung!

Fragen zum Studiengang

In welchem der folgenden Studiengänge studieren Sie?
(Mehrfachnennungen möglich; HF=Hauptfach, NF=Nebenfach)

- B.A. Empirische Kulturwissenschaft, HF
- B.A. Empirische Kulturwissenschaft, NF
- B.A. Erziehungswissenschaft und Soziale Arbeit/Erwachsenenbildung
- B.A. Erziehungswissenschaft, HF
- B.A. Erziehungswissenschaft, NF
- B.Ed. LA Berufl. Schulen Sozialpädagogik/Pädagogik und allgemeinb. Fach
- B.Sc. Empirische Bildungsforschung und Pädagogische Psychologie
- B.A. Politikwissenschaft, HF
- B.A. Politikwissenschaft, NF
- B.Ed. LA Gymnasium Politikwissenschaft
- B.A. Soziologie, HF
- B.A. Soziologie, NF
- B.Sc. Sportwissenschaft Gesundheitsförderung
- B.Sc. Sportwissenschaft Sportmanagement
- B.Sc. Sportwissenschaft Medien & Kommunikation
- B.Ed. LA Gymnasium Sport



Evaluation – Breakout Rooms (15 Minuten)

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN
Universität Tübingen

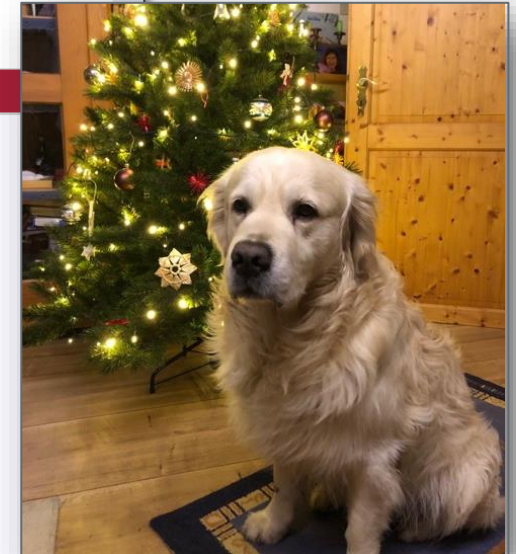
Kontrastmodus aktivieren

Liebe Studierende,
der Fragebogen soll dazu beitragen, die Qualität der Lehre zu beurteilen und zu verbessern. Wir bitten Sie daher, diesen Fragebogen sorgfältig auszufüllen. Lesen Sie die folgenden Aussagen und geben Sie an, wie stark Sie persönlich zustimmen. Sie können jeweils zwischen fünf Abstufungen wählen, die von "trifft voll und ganz zu" bis "trifft gar nicht zu", reichen.
Die Teilnahme ist freiwillig. Die Befragung wie auch die Auswertung unterliegen datenschutzrechtlichen Bestimmungen.
Dank für Ihre Mitwirkung!

Wählen Sie den Studiengang

In den folgenden Studiengängen studieren Sie?
(Nennungen möglich; HF=Hauptfach, NF=Nebenfach)

- B.A. Empirische Kulturwissenschaft, HF
- B.A. Empirische Kulturwissenschaft, NF
- B.A. Erziehungswissenschaft und Soziale Arbeit/Erwachsenenbildung
- B.A. Erziehungswissenschaft, HF
- B.A. Erziehungswissenschaft, NF
- B.Ed. LA Berufl. Schulen Sozialpädagogik/Pädagogik und allgemeinb. Fach
- B.Sc. Empirische Bildungsforschung und Pädagogische Psychologie
- B.A. Politikwissenschaft, HF
- B.A. Politikwissenschaft, NF
- B.Ed. LA Gymnasium Politikwissenschaft
- B.A. Soziologie, HF
- B.A. Soziologie, NF
- B.Sc. Sportwissenschaft Gesundheitsförderung
- B.Sc. Sportwissenschaft Sportmanagement
- B.Sc. Sportwissenschaft Medien & Kommunikation
- B.Ed. LA Gymnasium Sport





Übersicht

- **Prozessablauf (inkl. Konstrukt, Operationalisierung)**
- **Deskriptive Statistik**
- **Inferenzstatistik**
- **Anforderungen an Testverfahren**



Abb.1



Diagnostik: Prozessablauf

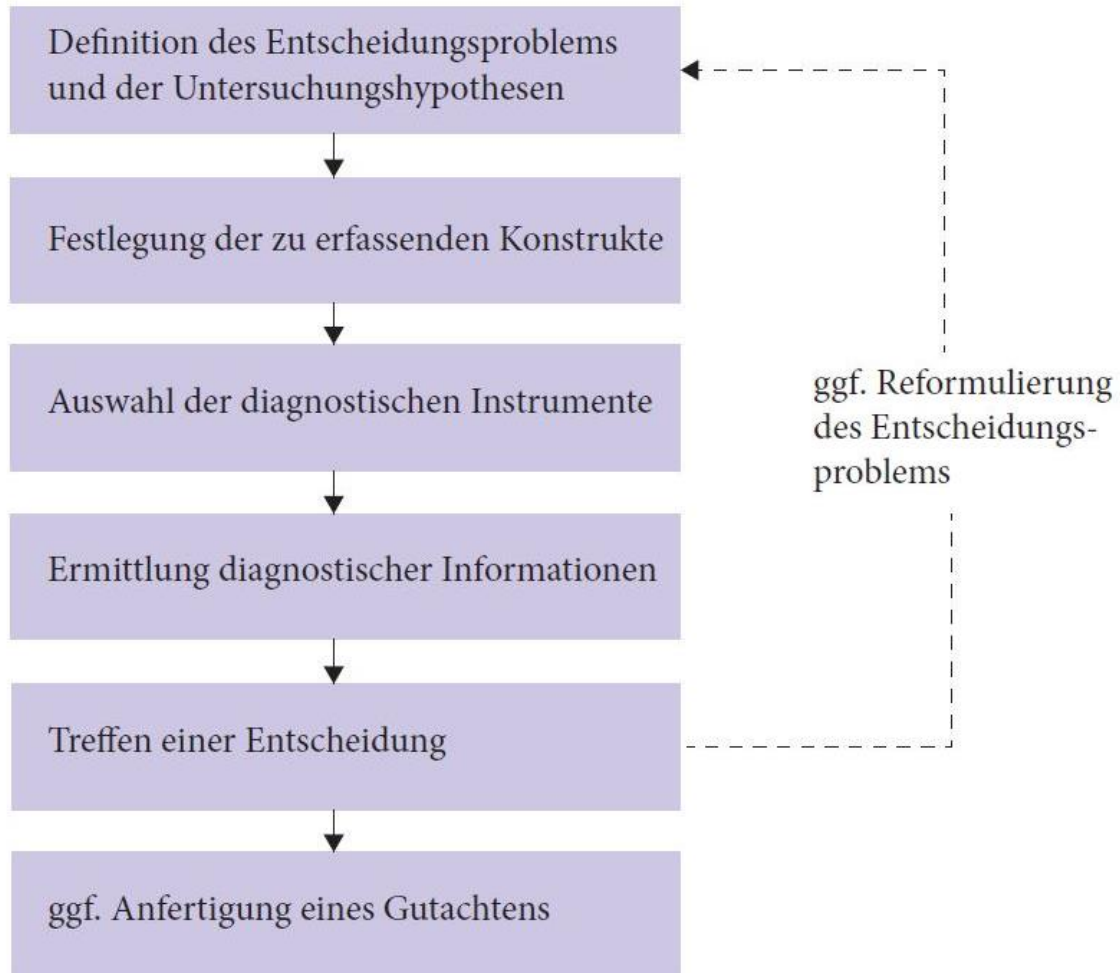


Abb.2: (Brunner, Stanat & Pant, 2014, S. 489)



Diagnostik: Prozessablauf

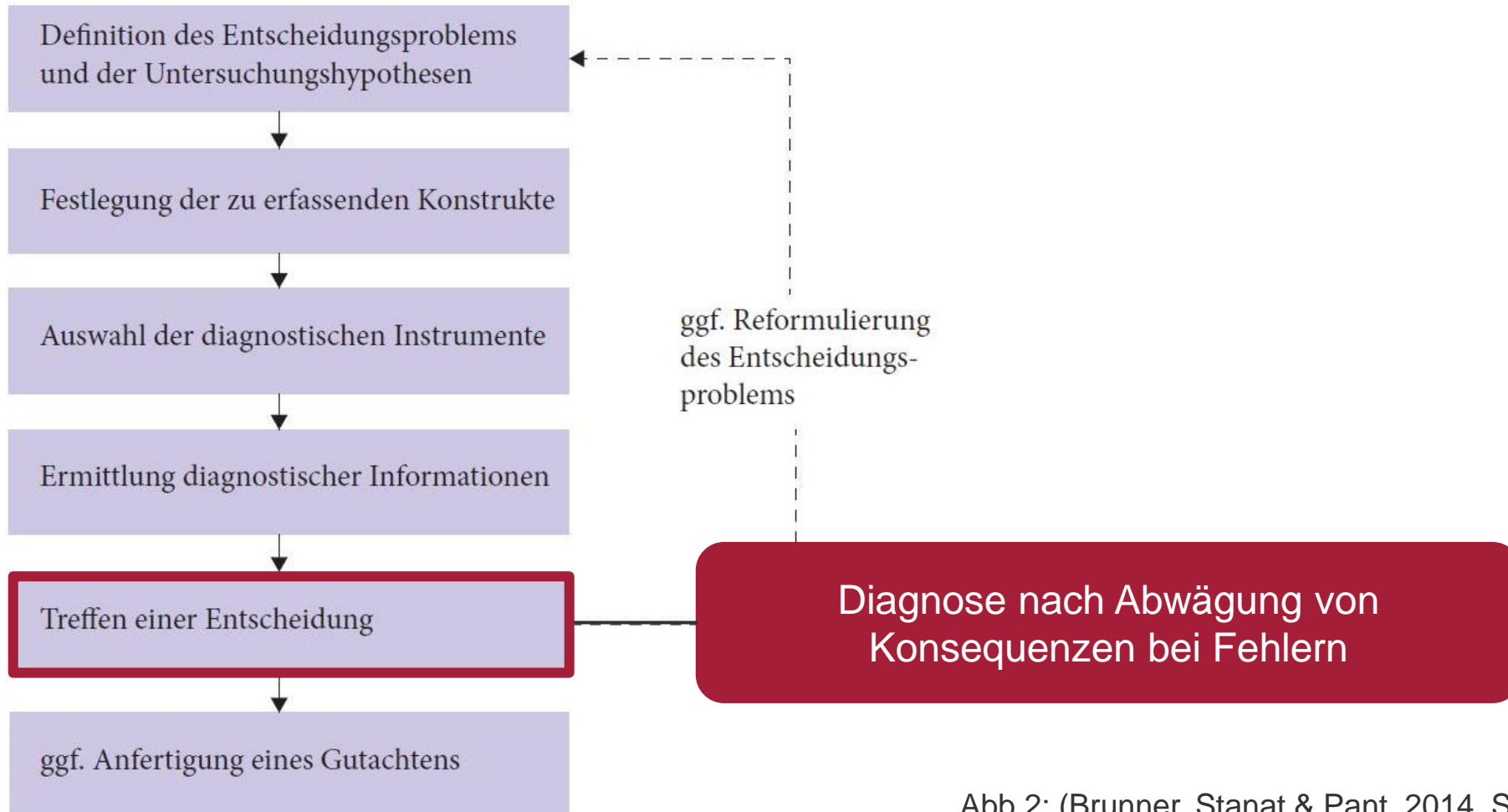


Abb.2: (Brunner, Stanat & Pant, 2014, S. 489)



Diagnostik: Prozessablauf

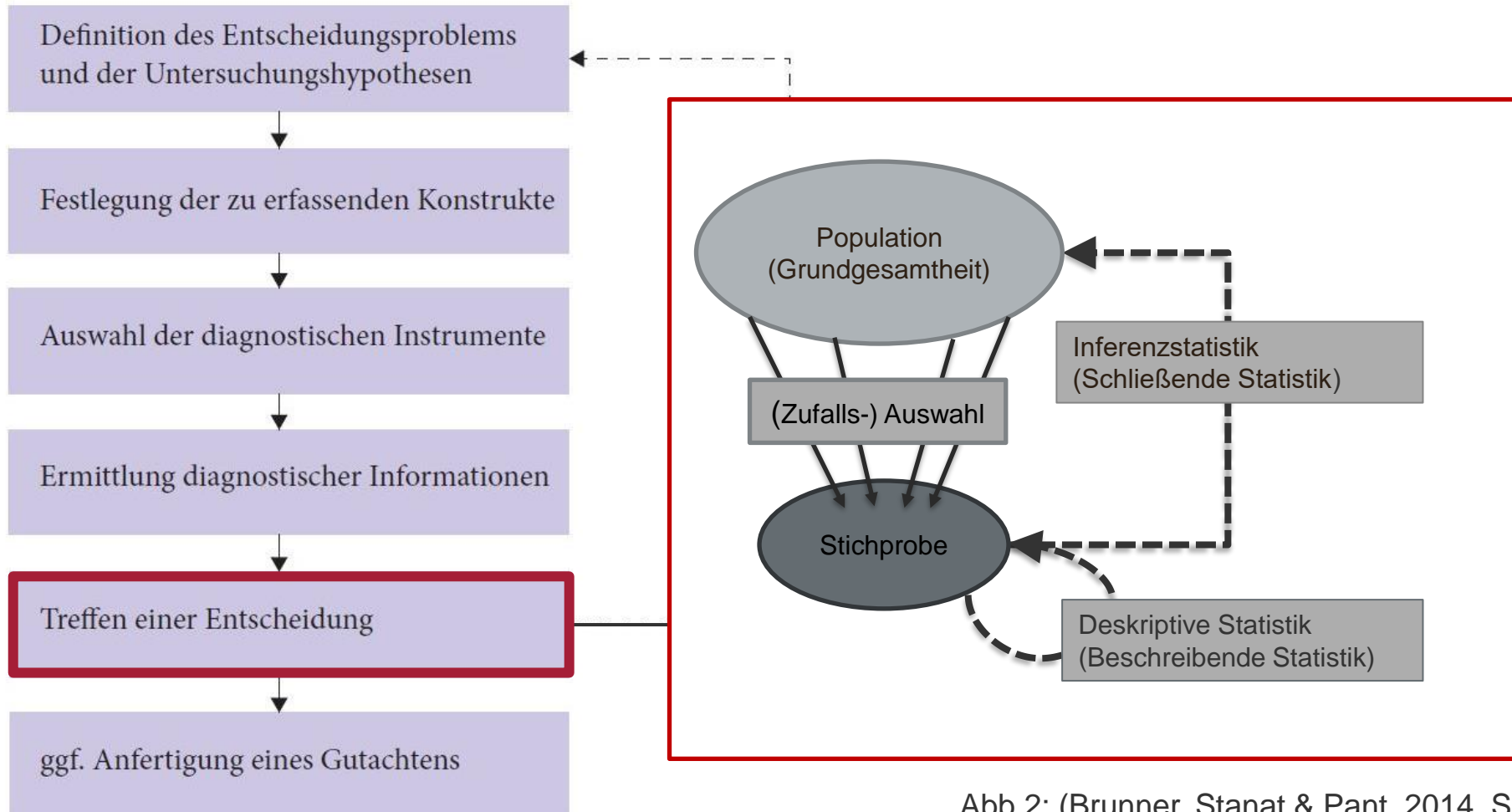


Abb.2: (Brunner, Stanat & Pant, 2014, S. 489)



Übersicht

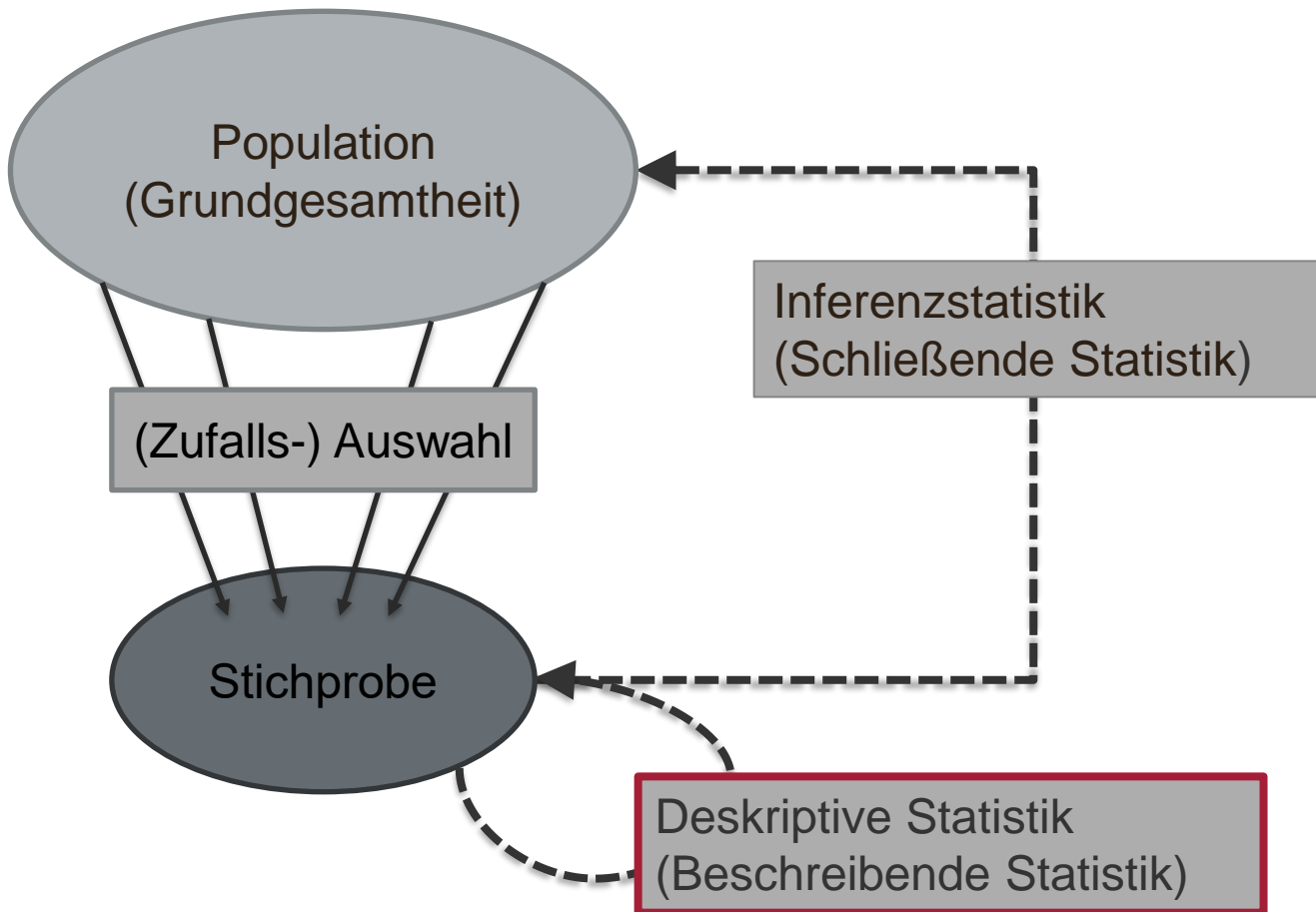
- Prozessablauf (inkl. Konstrukt, Operationalisierung)
- **Deskriptive Statistik**
- Inferenzstatistik
- Anforderungen an Testverfahren



Abb.1



Deskriptive Statistik und Inferenzstatistik





Deskriptive Statistik und Inferenzstatistik

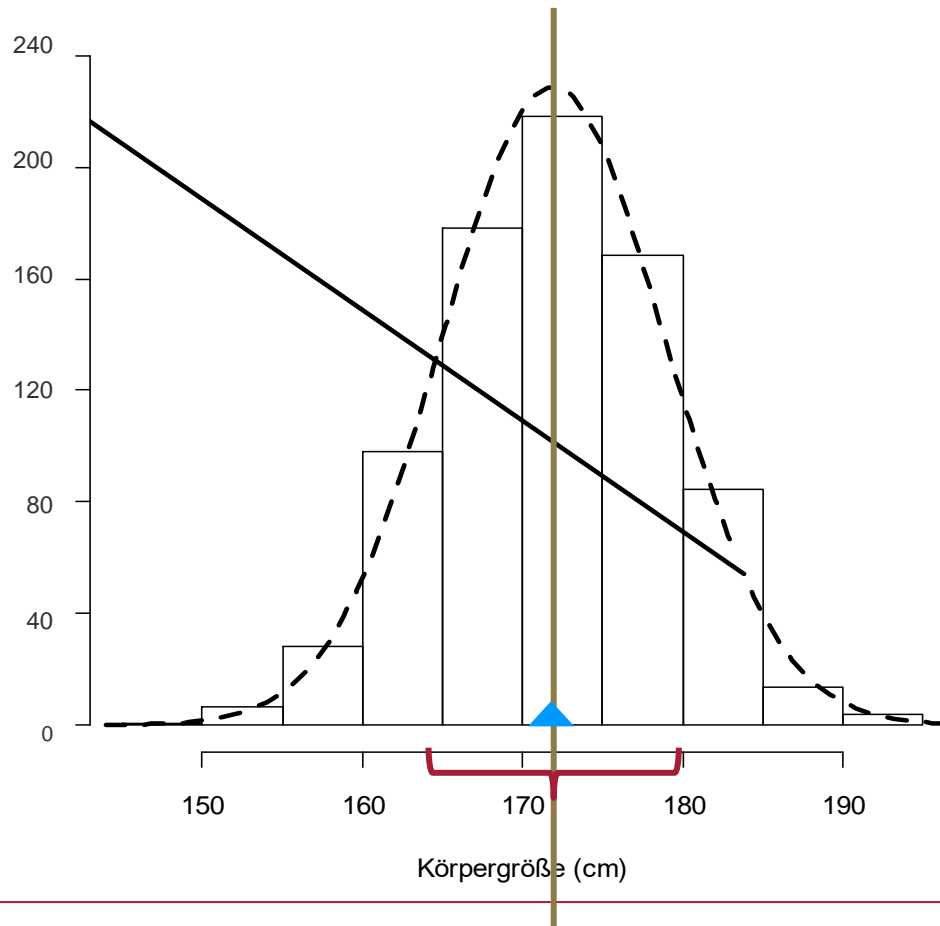
Die deskriptive (auch: beschreibende) Statistik hat zum Ziel, **empirische Daten** durch Tabellen, Kennzahlen (auch: Maßzahlen oder Parameter) und Grafiken **übersichtlich darzustellen und zu ordnen**. Dies ist vor allem bei umfangreichem Datenmaterial sinnvoll, da dieses nicht leicht überblickt werden kann.

Statistiken beziehen sich auf die Stichprobe, Verallgemeinerungen bezogen auf die zugrundeliegende Population werden nicht vorgenommen!



Statistiken: Mittelwert und Standardabweichung

Verteilung der Körpergröße von 1078 Männer, Histogramm der beobachteten Körpergrößen



Mittelwert

Information zur durchschnittlichen Ausprägung der Variable über alle Personen hinweg

Standardabweichung

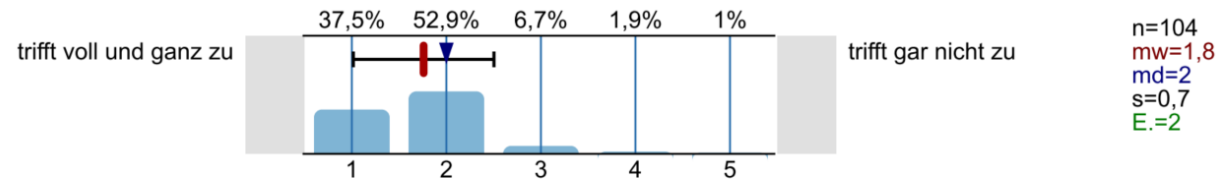
Information zur Streuung der individuellen Ausprägungen in der Variable (Unterschiedlichkeit der Personen)

Median

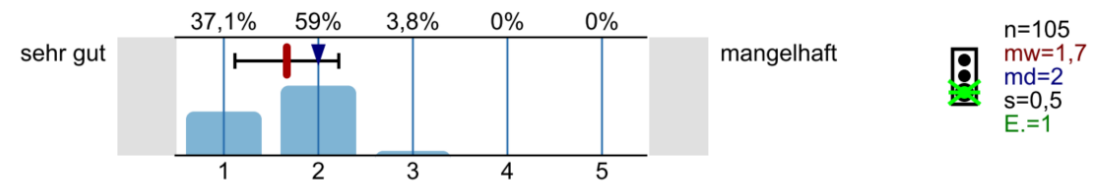
Information zum Zentralwert der individuellen Ausprägungen in der Variable

Deskriptive Statistiken

5.1) Der Besuch der Veranstaltung hat sich für mich gelohnt.



5.2) Ich gebe der Lehrveranstaltung die Gesamtnote:



5.3) Ich gebe dem Dozenten / der Dozentin die Gesamtnote:

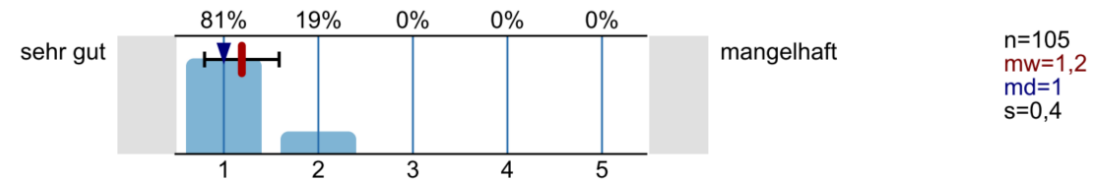
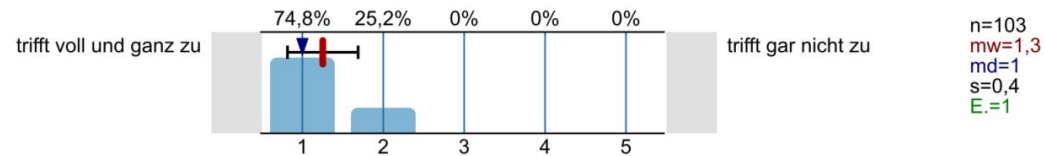


Abb.3

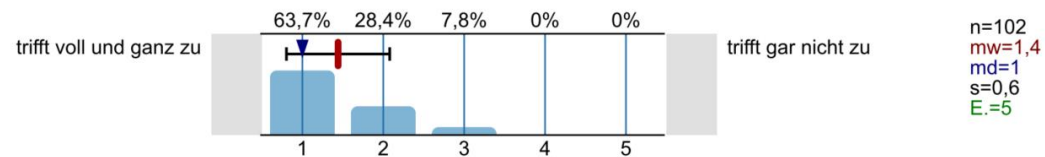
Deskriptive Statistiken: Häufigkeitsverteilung Skala

2. Fragen zur Lehrveranstaltung

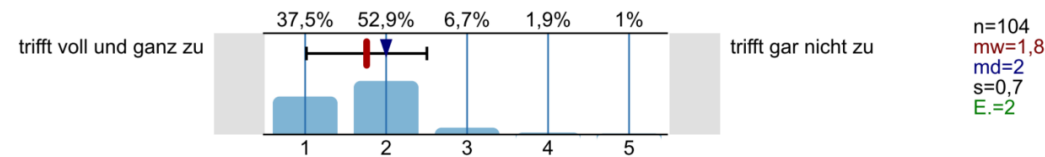
2.1) Die Veranstaltung ist inhaltlich klar strukturiert.



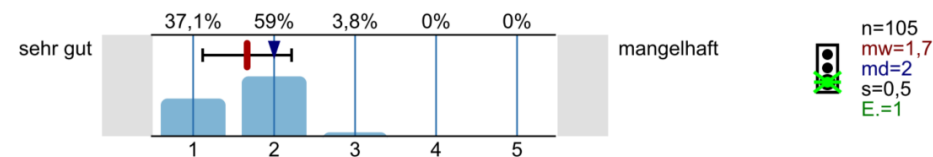
2.2) Die Unterlagen zur Vorbereitung sind hilfreich.



5.1) Der Besuch der Veranstaltung hat sich für mich gelohnt.



5.2) Ich gebe der Lehrveranstaltung die Gesamtnote:



5.3) Ich gebe dem Dozenten / der Dozentin die Gesamtnote:

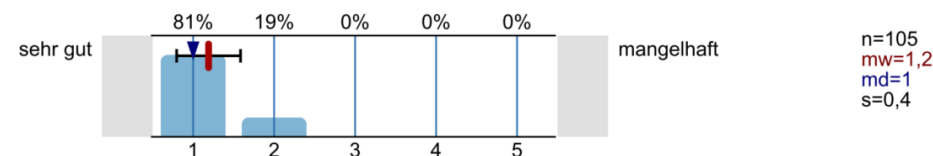


Abb.4



Übersicht

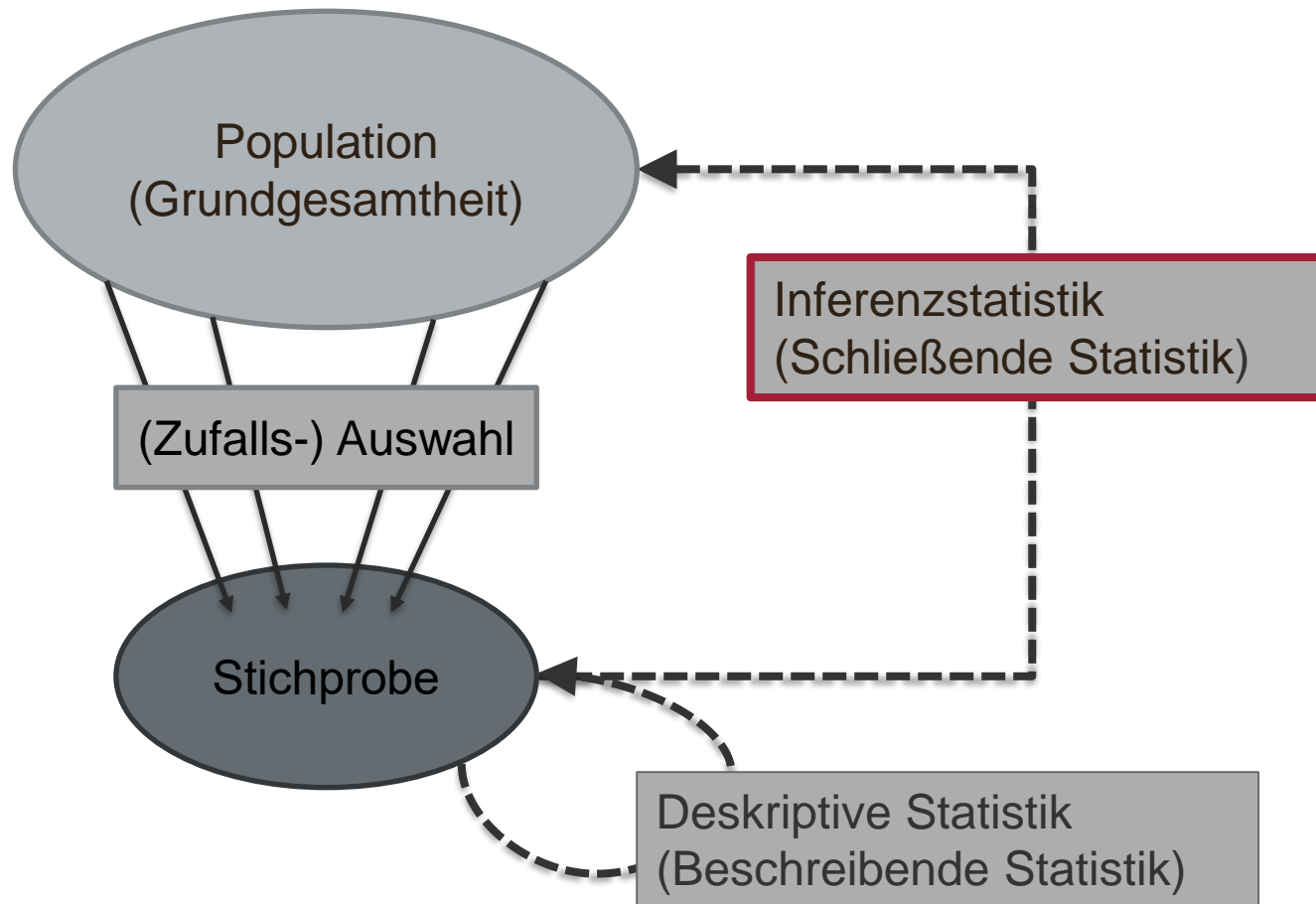
- Prozessablauf (inkl. Konstrukt, Operationalisierung)
- Deskriptive Statistik
- **Inferenzstatistik**
- Anforderungen an Testverfahren



Abb.1



Deskriptive Statistik und Inferenzstatistik





Inferenzstatistik

Die Inferenzstatistik oder auch schließende Statistik der Psychologie beschäftigt sich mit der Prüfung von **Hypothesen** (Nullhypothese gegen Alternativhypothese), wobei auf Grund einer Prüfverteilung von den Kennwerten einer Stichprobenverteilung auf die Populationsparameter geschlossen wird.

Statistiken beziehen sich auf die Grundgesamtheit (Population). Es werden Verallgemeinerungen bezogen auf die zugrundeliegende Population vorgenommen!

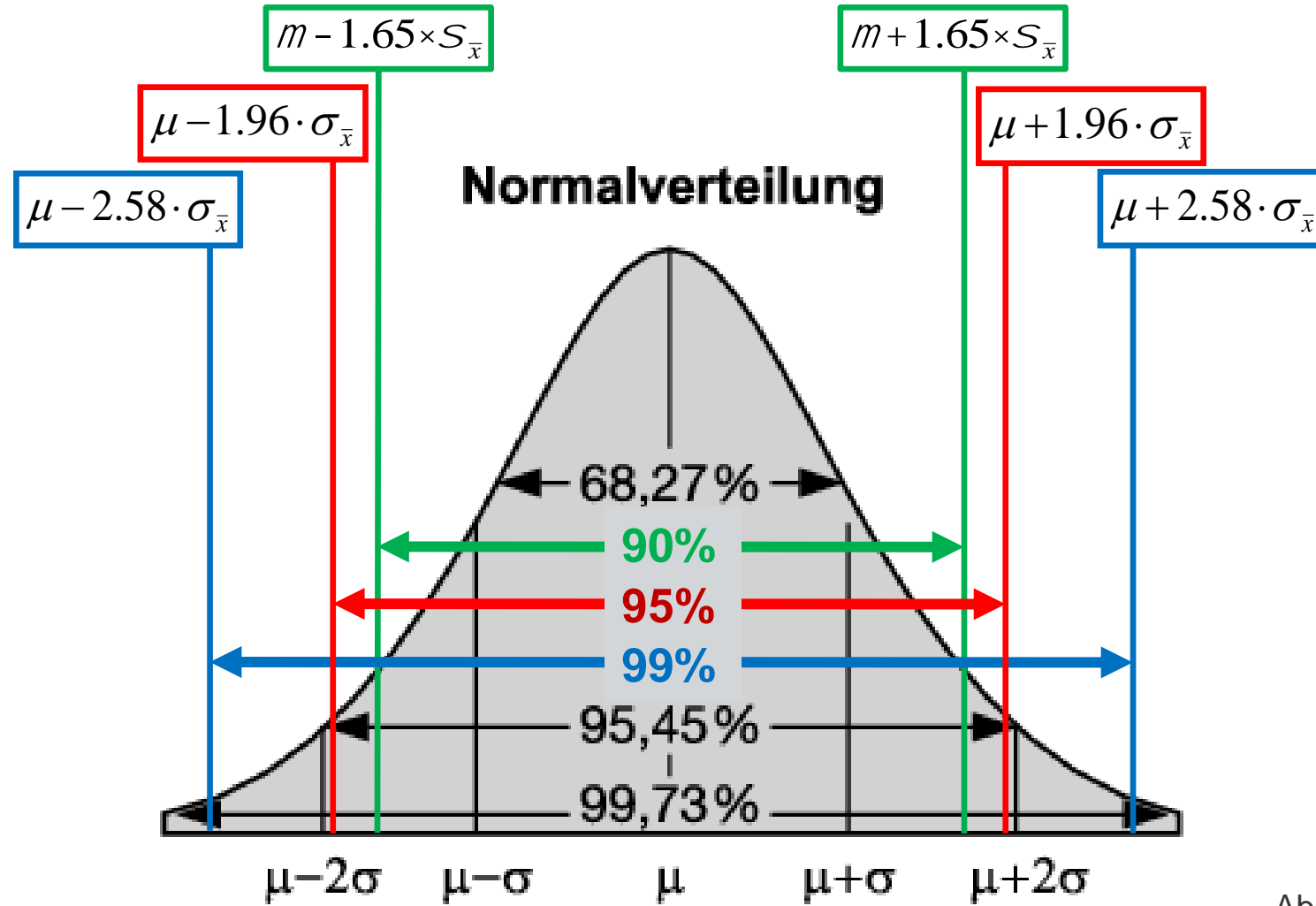


Abb.5



Achtung, bei Hypothesentests!

Im Falle der Ablehnung der Nullhypothese (da $p \leq \alpha$) ist es dennoch nicht ausgeschlossen, dass die Nullhypothese trotzdem gilt und wir mit der Rückweisung einen Fehler begehen:

- Wir haben einfach per Zufall ein sehr unwahrscheinliches Stichprobenergebnis bekommen.
- Dies ist die Irrtumswahrscheinlichkeit α .



Entscheidung aufgrund der Stichprobe	In Population gilt:	
	H_0	H_1
H_0	richtige Entscheidung	β -Fehler
H_1	α -Fehler	richtige Entscheidung

α -Fehler (Fehler 1. Art)

- Wahrscheinlichkeit dieses Fehlers wird vor der Auswertung festgelegt (α -Niveau)

β -Fehler (Fehler 2. Art)

- Wahrscheinlichkeit dieses Fehlers ist nicht einfach zu bestimmen, ist aber bei kleinen Stichproben höher (da Unterschiede größer ausfallen müssen, um signifikant zu werden)

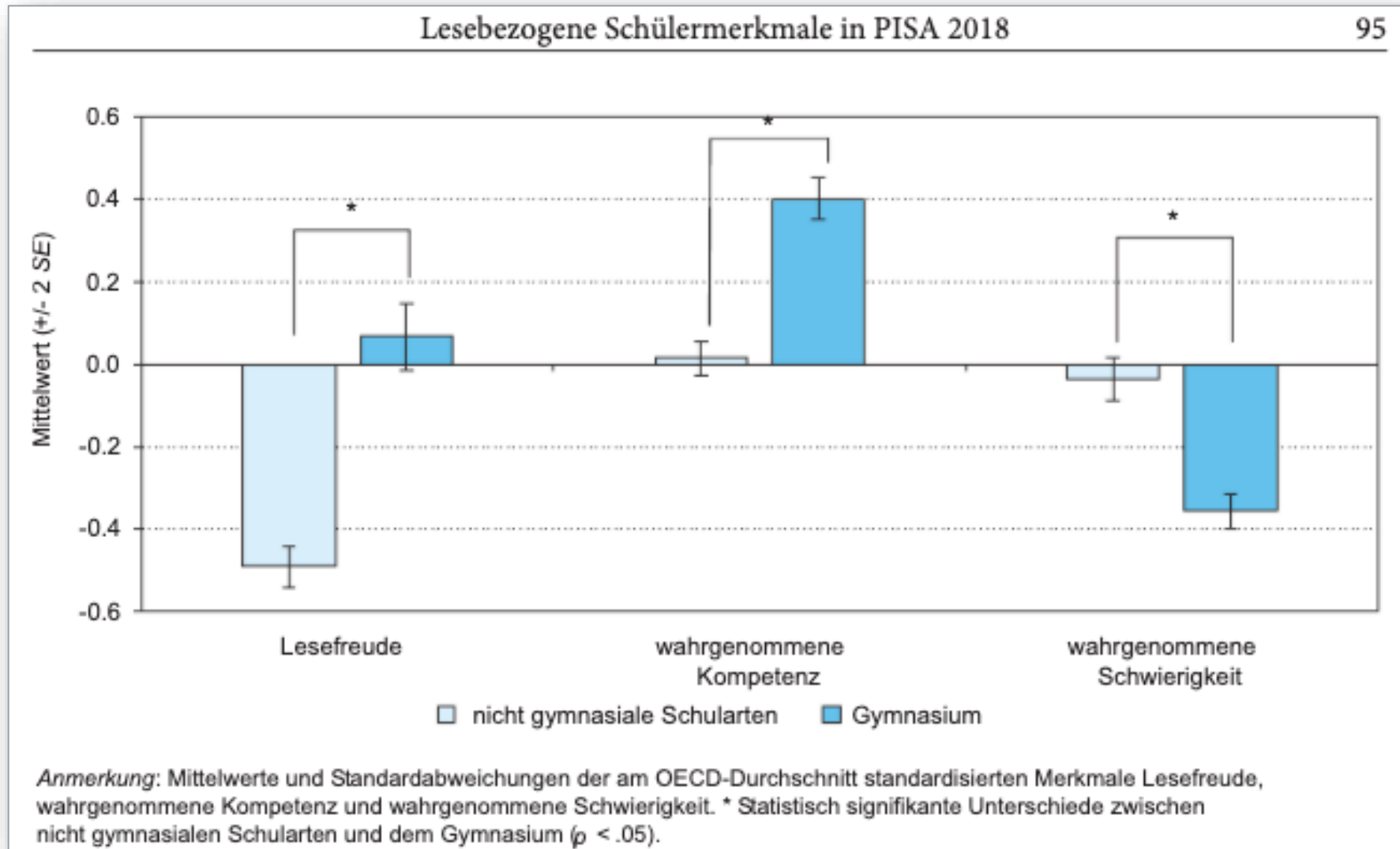


Abb.6



Inferenzstatistik: Standardverfahren

Wahl der Analysemethode anhand der Fragestellung u. d. Datenmaterials

- **Korrelation**

Wie eng ist der Zusammenhang zwischen zwei Merkmalen?

- **Regression**

Kann ein Merkmal durch ein anderes mittels Regressionsgleichung vorhergesagt werden?



Inferenzstatistik: Standardverfahren

Wahl der Analysemethode anhand der Fragestellung u. d. Datenmaterials

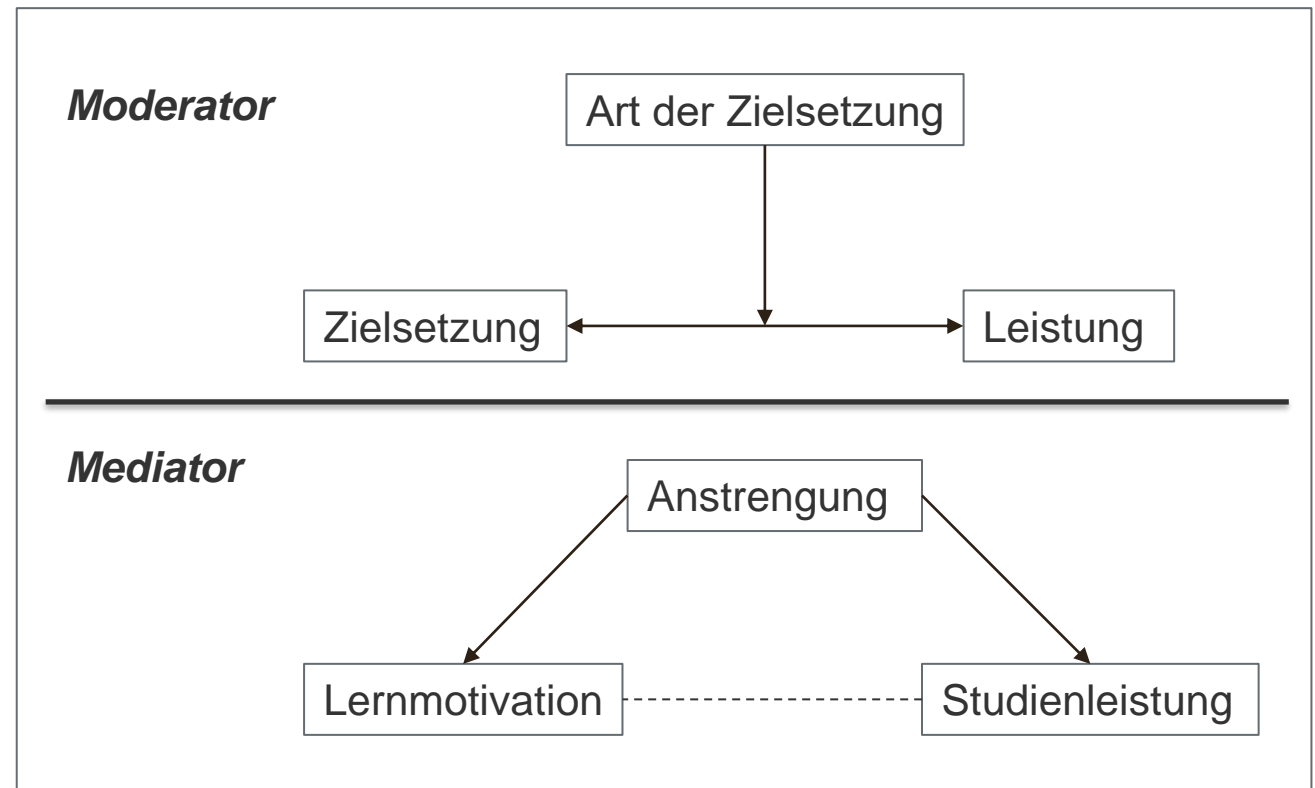
- **Korrelation**

- **Regression**

Einfluss dritter Variablen

- Moderation

- Mediation





Analyseverfahren: Korrelation

- Der Korrelationskoeffizient (meist mit „ r “ abgekürzt) ist ein Maß für den Grad des **linearen Zusammenhangs** zwischen zwei Merkmalen.
- kann Werte zwischen **-1 und +1** annehmen
- Bei einem Wert von +1 (bzw. -1) besteht ein perfekt positiver (bzw. negativer) linearer Zusammenhang zwischen den betrachteten Merkmalen.
- Wenn der Korrelationskoeffizient den **Wert 0** aufweist, hängen die beiden Merkmale überhaupt nicht linear miteinander zusammen.
- Klassifikation von Cohen (1988):
 $r = 0.1$ -> gering, $r = 0.3$ -> mittel, $r = 0.5$ -> hoch

Veranschaulichung: Korrelationskoeffizienten

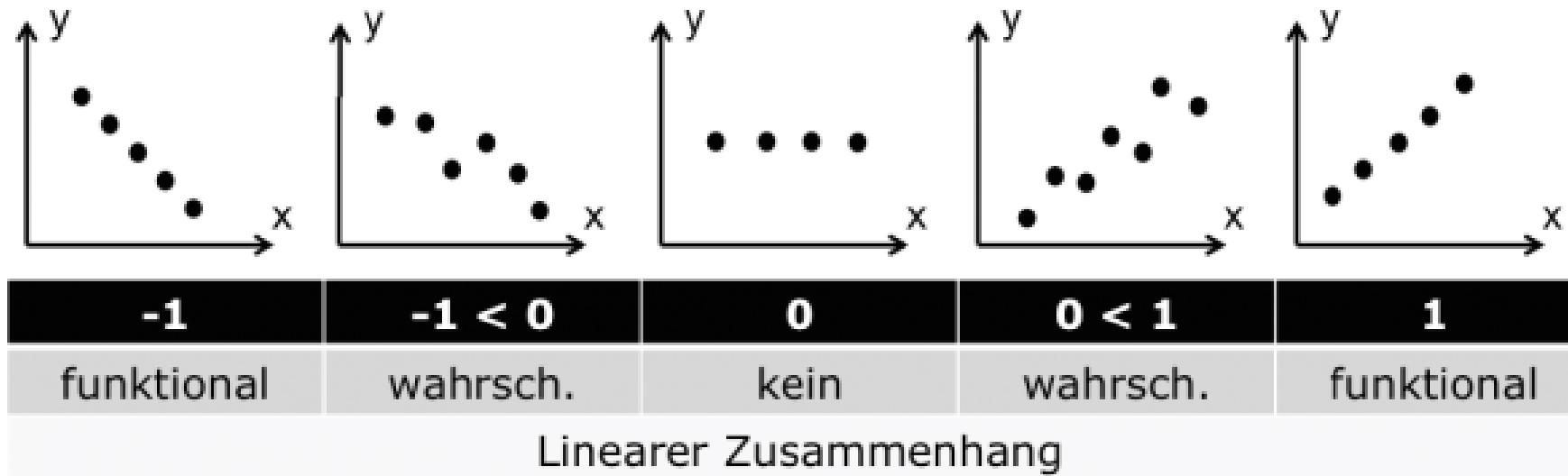
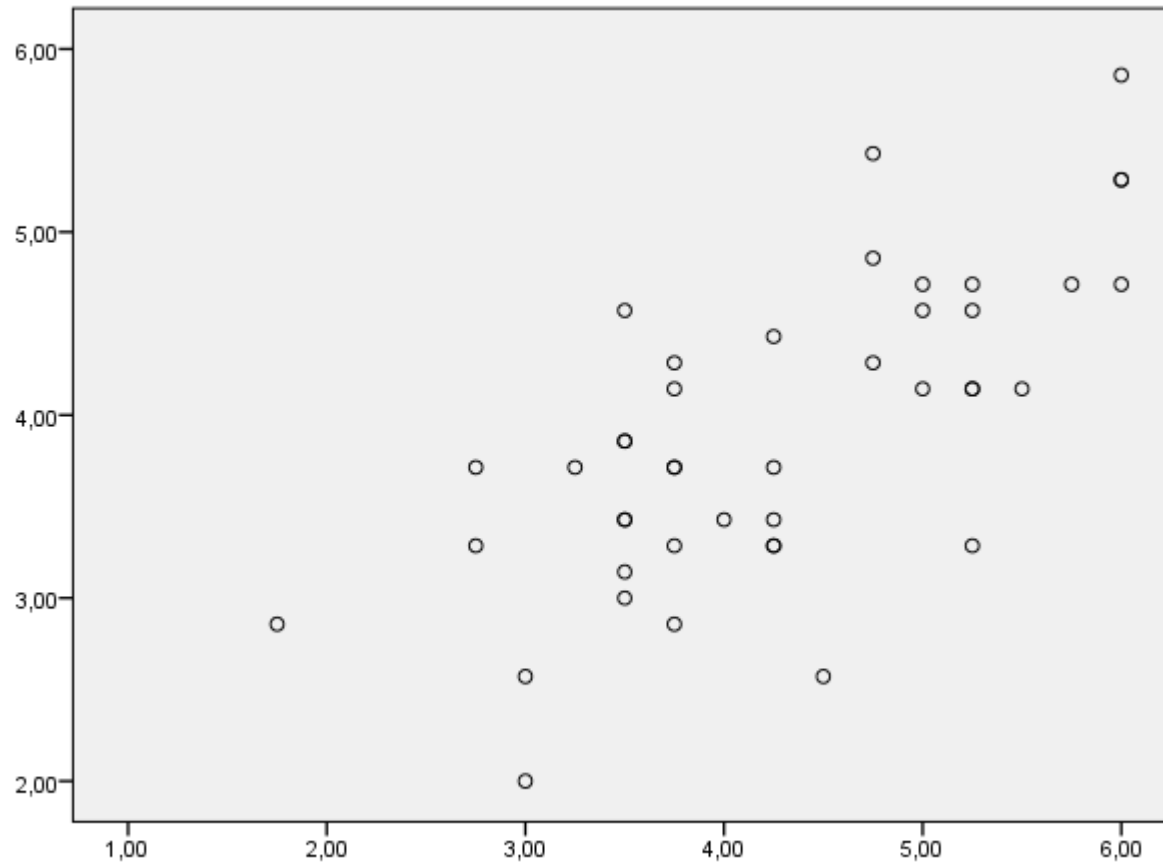


Abb.7



Korrelationen: Beispiel 1



TEST YOUR KNOWLEDGE

(A) $r > 0$

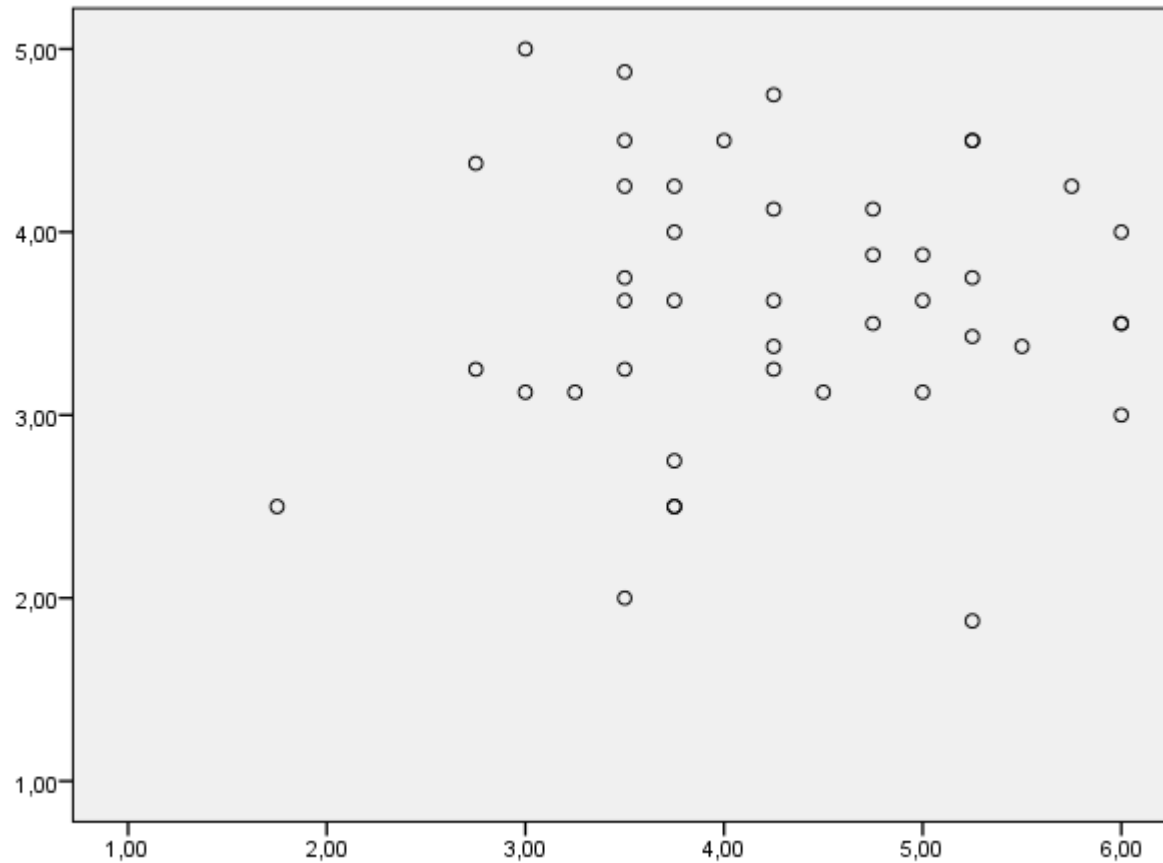
(B) $r \approx 0$

(C) $r < 0$

Entscheidet euch für eine der drei
Antworten und schreibt den
zugehörigen Buchstaben in den
Zoom Chat.



Korrelationen: Beispiel 2



TEST YOUR KNOWLEDGE

(A) $r > 0$

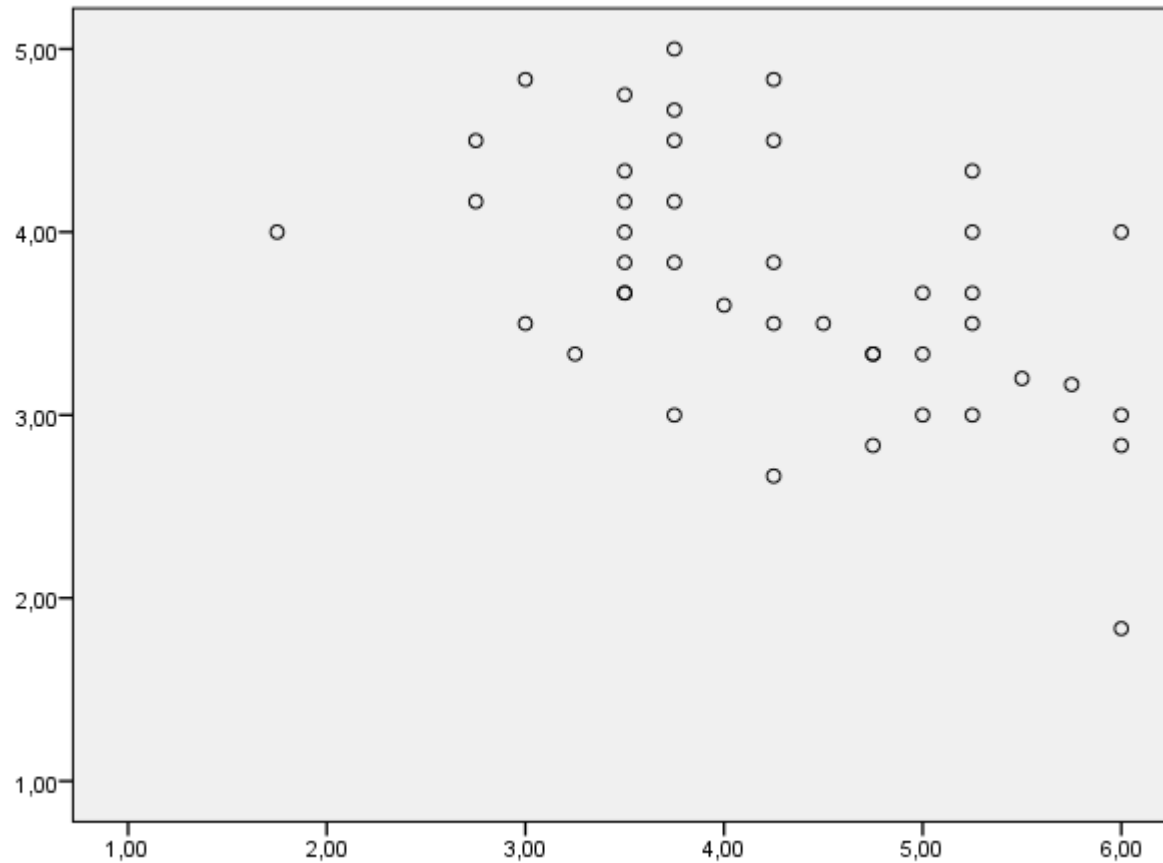
(B) $r \approx 0$

(C) $r < 0$

Entscheidet euch für eine der drei
Antworten und schreibt den
zugehörigen Buchstaben in den
Zoom Chat.



Korrelationen: Beispiel 3



TEST YOUR KNOWLEDGE

(A) $r > 0$

(B) $r \approx 0$

(C) $r < 0$

Entscheidet euch für eine der drei
Antworten und schreibt den
zugehörigen Buchstaben in den
Zoom Chat.



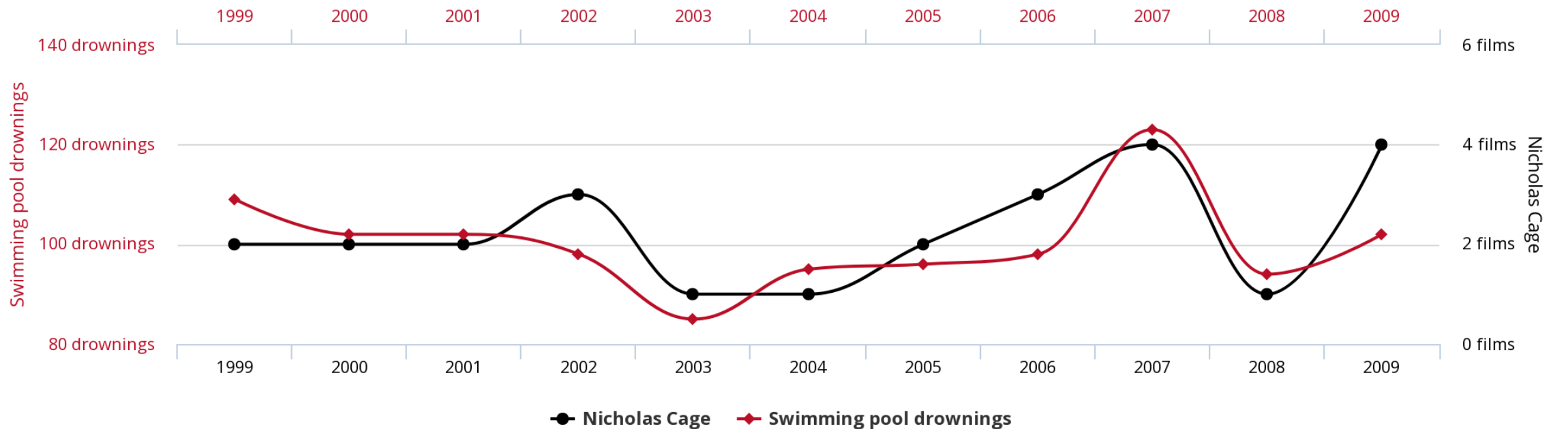
Korrelation: Hinweis

- **Vorsicht** bei Korrelation und Kausalität!
- Zwei Probleme:
 - **Richtung des Zusammenhangs**
 - Bsp.: Leistung \leftrightarrow Motivation
 - keine „Wenn dann Beziehung“, eher „je/ desto“
 - **Mögliche Drittvariablen (Scheinkorrelation)**
 - Bsp.: Schuhgröße \leftrightarrow Intelligenz
 - Bsp.: Eiskonsum \leftrightarrow Tote durch Ertrinken



Korrelation: Scheinkorrelationen

Number of people who drowned by falling into a pool
correlates with
Films Nicolas Cage appeared in



tylervigen.com

Abb.8: Chart by Tyler Vigen



Übersicht

- Prozessablauf (inkl. Konstrukt, Operationalisierung)
- Deskriptive Statistik
- Inferenzstatistik
- **Anforderungen an Testverfahren**



Abb.1



Messgüte: Wann ist ein Test ein guter Test?

Die drei wichtigsten Gütekriterien:

1. **Objektivität**
2. **Reliabilität**
3. **Validität**



Messgüte: Objektivität

Ein Test ist dann objektiv, wenn er dasjenige Merkmal, das er misst, unabhängig von Testleiter und Testauswerter misst. Außerdem müssen klare und anwenderunabhängige Regeln für die Ergebnisinterpretation vorliegen.

- **Durchführungsobjektivität**
- **Auswertungsobjektivität**
- **Interpretationsobjektivität**



Abb.9



Objektivität: Arten

- **Durchführungsobjektivität:** Durchführung eines Tests muss immer gleich sein – Instruktion, Zeitvorgabe, etc.
 - Beispiel: Testleiter; Schule: Lehrperson
- **Auswertungsobjektivität:** Alle Auswerter ermitteln dieselben Leistungswerte
 - Schablonen
 - Offene Antworten: Kodierschemata – Übereinstimmungskoeffizienten?!
 - Schule: Auswertungsbogen
- **Interpretationsobjektivität:** Unterschiedliche Auswerter kommen zur selben Interpretation
 - Normen, PISA: „Risikoschüler“?
 - Schule: Was bedeuten 23 Punkte in einem Deutschaufsatz -> Skala



Messgüte: Reliabilität

Ein Test ist dann reliabel (zuverlässig), wenn er das Merkmal, das er misst, exakt, d.h. ohne Messfehler, misst.

- **Retest-Reliabilität**
- **Paralleltest-Reliabilität**
- **Innere Konsistenz**



Reliabilität: Arten

- **Retest-Reliabilität (= Stabilität):** zu unterschiedlichen Messzeitpunkten wird ein ähnlicher Wert gemessen
 - Problem: Zwischenzeitliche Veränderung?
- **Paralleltest-Reliabilität:** mit „vergleichbaren“ Testhälften wird ein ähnliches Ergebnis erzielt
 - **Interne (innere) Konsistenz:** Homogenität des Tests – Korrelation der Items untereinander, häufig durch Cronbachs' Alpha bestimmt, Werte zwischen 0 und 1, (.90 exzellent, > .80 gut, > .70 akzeptabel)



Messgüte: Validität

Ein Test gilt dann als valide („gültig“), wenn er das Merkmal, das er messen soll, auch wirklich misst und nicht irgendein anderes.

- **Inhaltsvalidität**
- Augenscheinvalidität
- **Konstruktvalidität (konvergente, diskriminante)**
- **Kriteriumsvalidität**

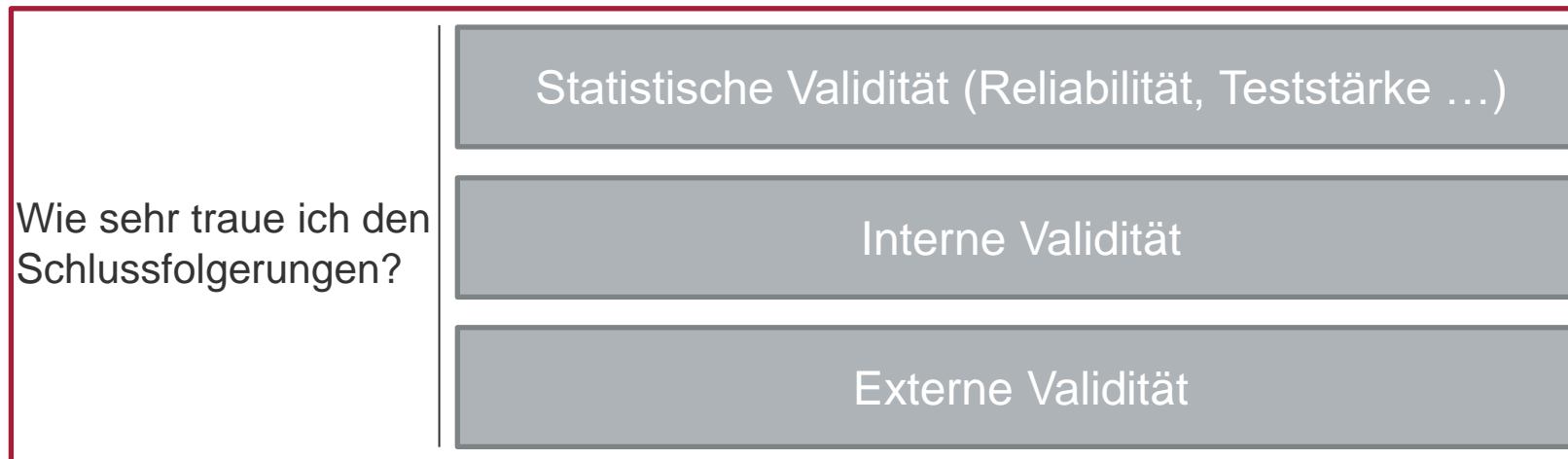
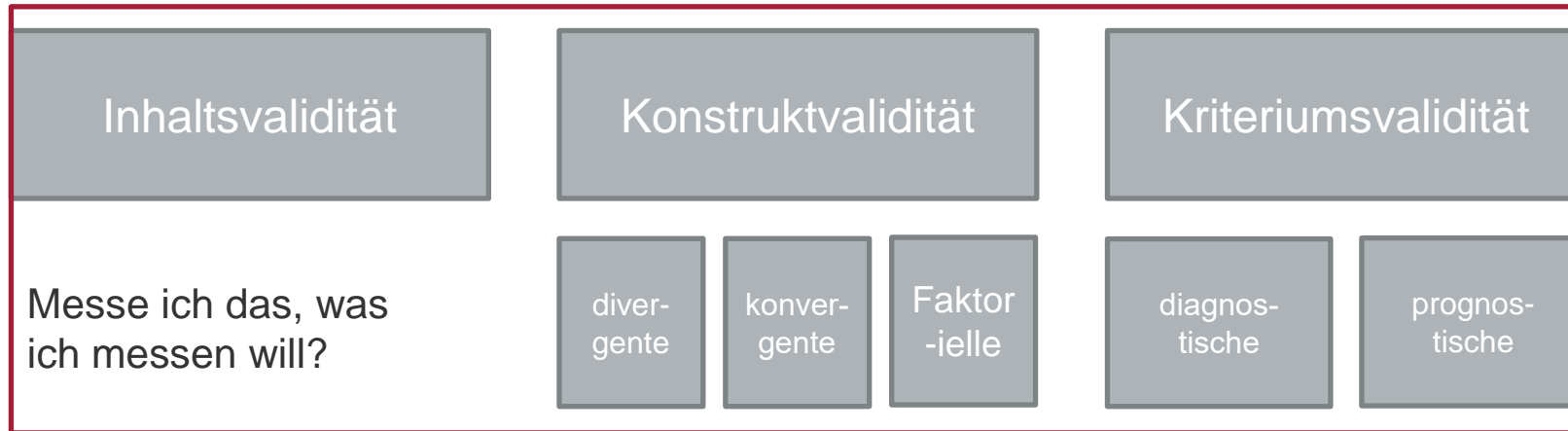


Validität: Arten

- **Inhaltsvalidität:** wird angenommen, wenn ein Verfahren zur Messung eines bestimmten Konstrukts alle Aspekte dieses Konstrukts erschöpft
 - Test zur Rechenleistung: alle Grundrechenarten
- **Kriteriumsvalidität:** Test sollte mit einem Kriterium einen hohen Zusammenhang (z.B. Korrelation) aufweisen
 - Hochschulzulassungstest und Studienerfolg
- **Konstruktvalidität:**
 - Konvergente Validität: Zusammenhang mit „ähnlichem“ Test (Intelligenztest A - Intelligenztest B)
 - Diskriminante/divergente Validität: niedrigerer Zusammenhang mit „anderem“ Konstrukt (Aufmerksamkeitstest - Intelligenztest)



Validität: Arten





Mögliche Probleme beim Messen

- Verletzung von Gütekriterien (teilweise vermeidbar)
- Schwellenproblem (unvermeidbar)
- Rundungsproblem (unvermeidbar)
- Stichprobenprobleme (teilweise vermeidbar)



Abb.10

Probleme beim Messen: Stichprobenproblem

- **Systematische Fehler** (keine Repräsentativität)
- **Unsystematische Fehler** (sampling error; Abb. rechts)

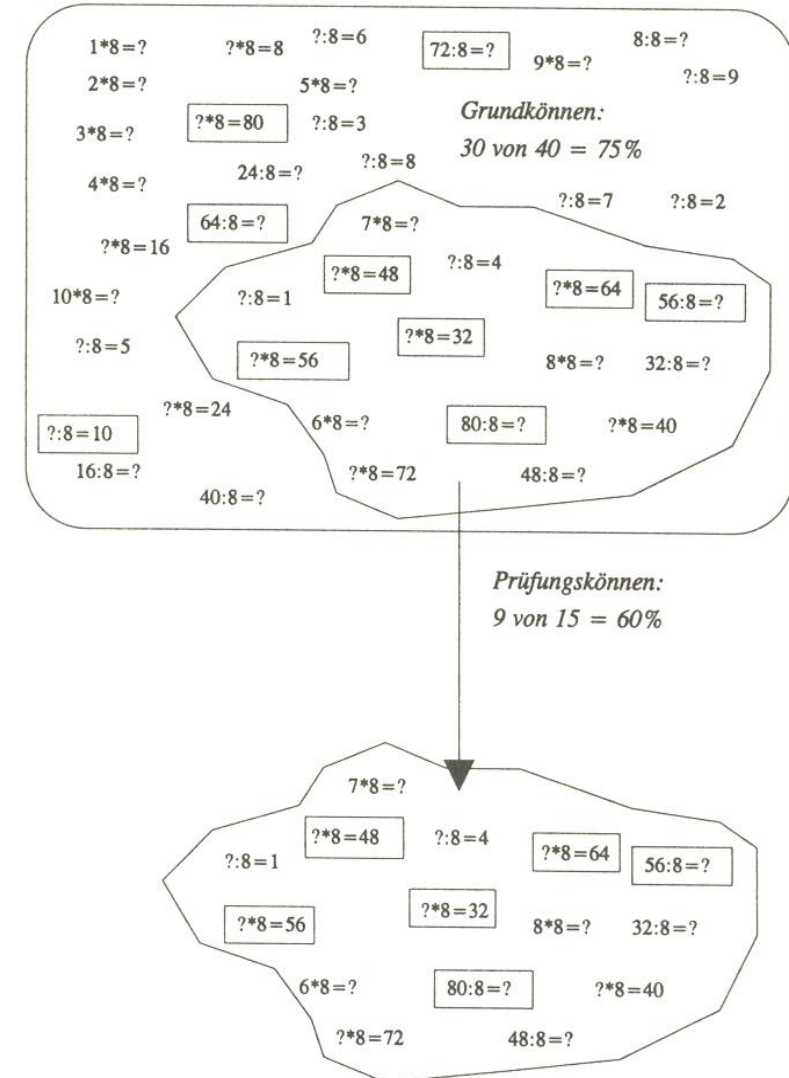


Abb.11: Quelle

Sacher, W.: Leistungen entwickeln, überprüfen und beurteilen. Bewährte und neue Wege für die Primar- und Sekundarstufe, Klinkhardt, Bad Heilbrunn 2014, S. 111



Leistungsbeurteilung: Skalen

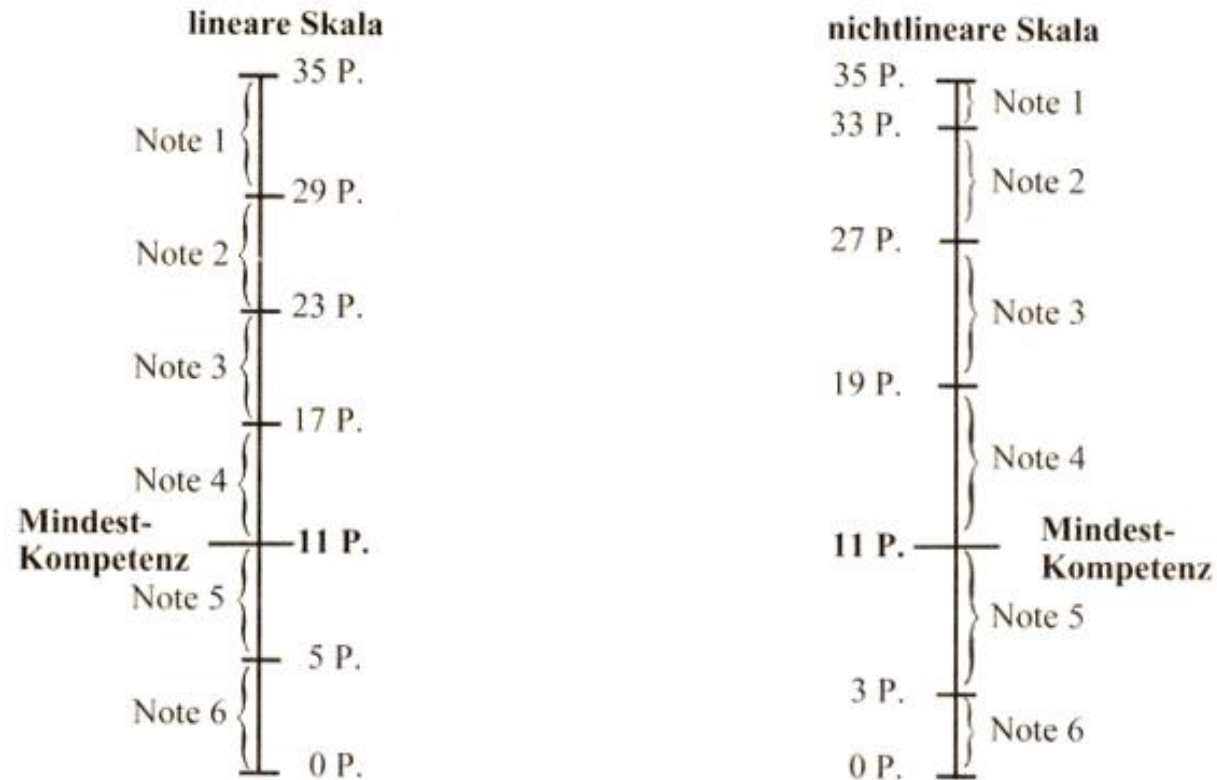
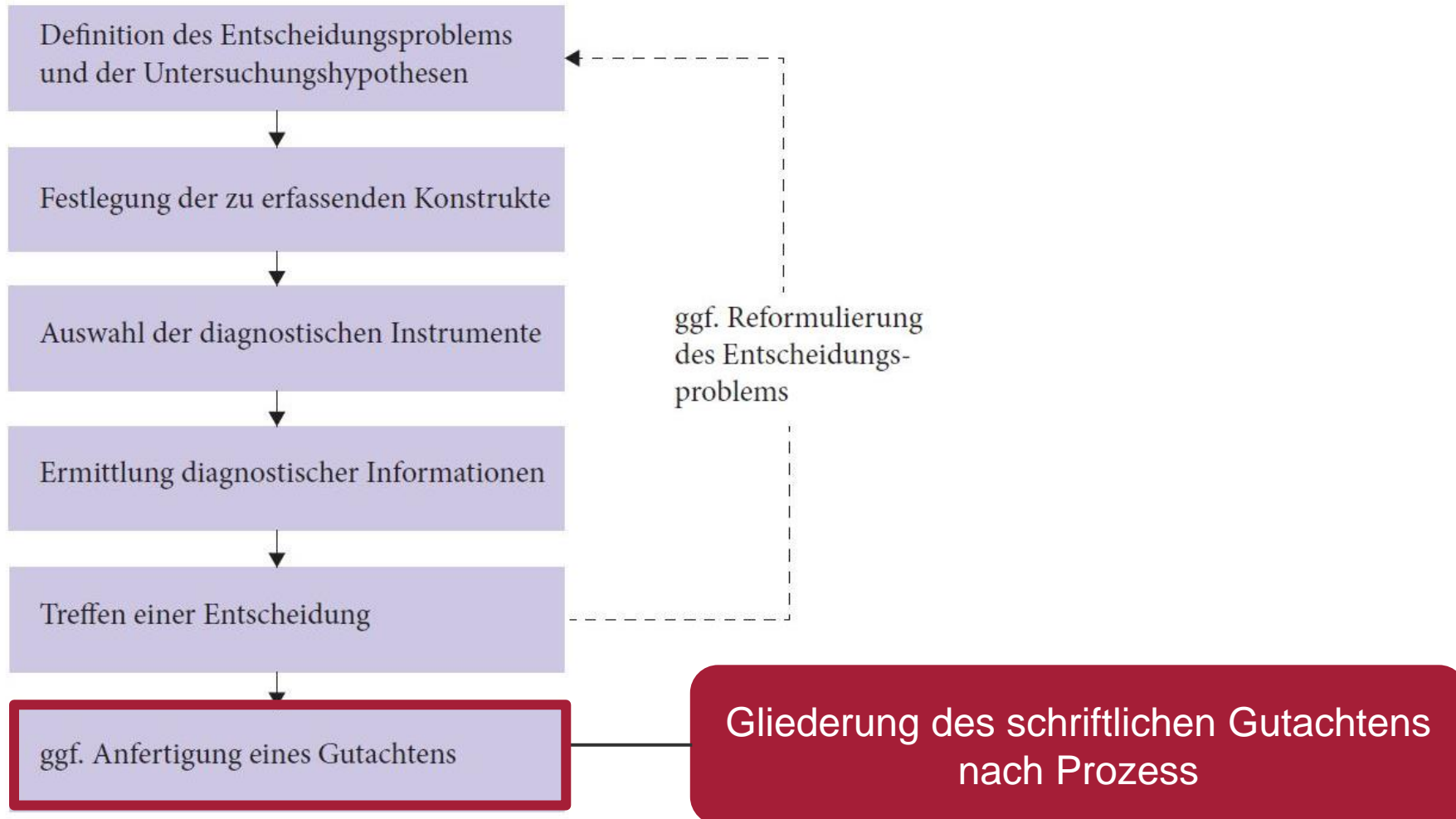


Abb.12: Quelle

Sacher, W.: Leistungen entwickeln, überprüfen und beurteilen. Bewährte und neue Wege für die Primar- und Sekundarstufe, Klinkhardt, Bad Heilbrunn 2014, S. 106



Diagnostik: Prozessablauf





Anfertigung von Gutachten



Abb. 13

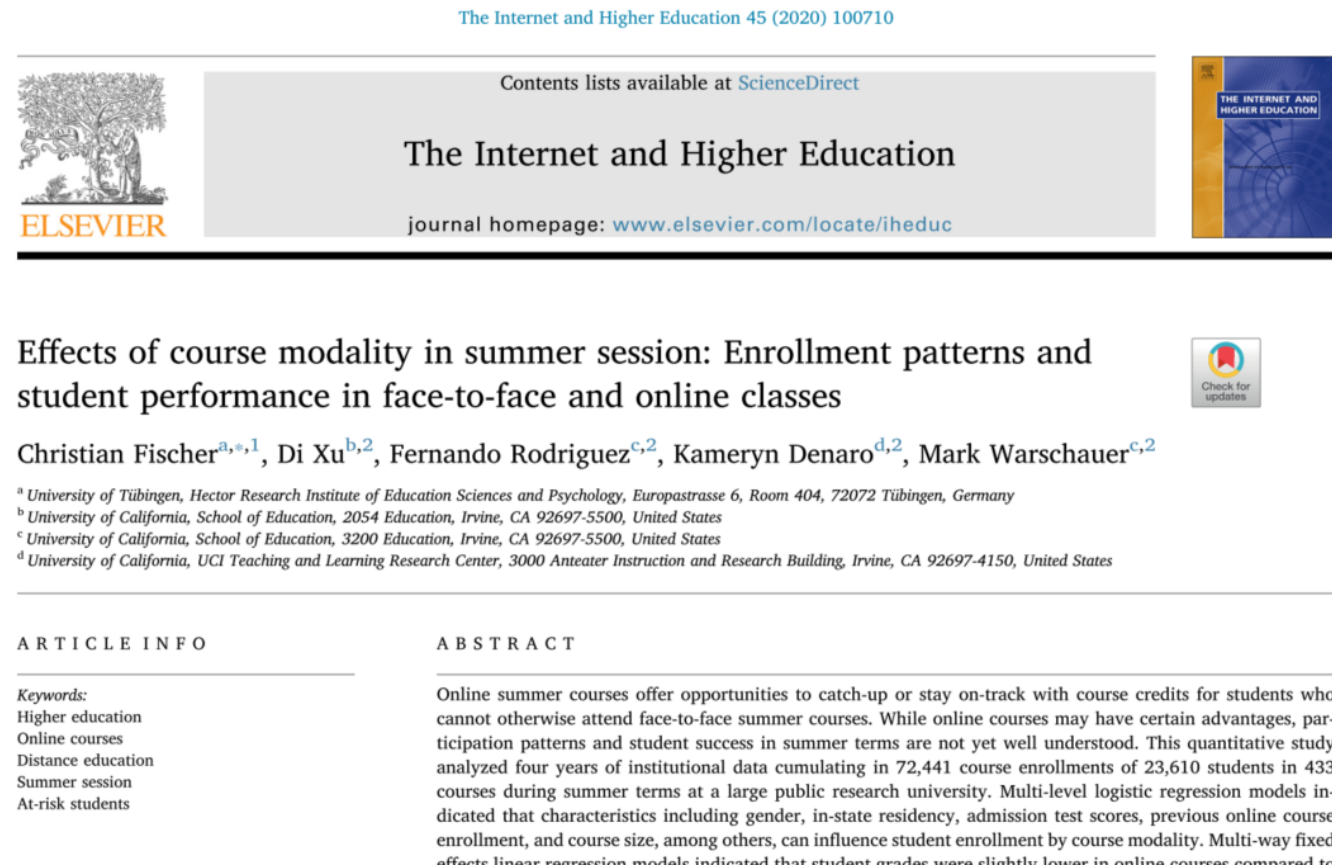


Abb.14



Fassen wir zusammen ...

In der Praxis ermöglichen diagnostische Informationen:

- hinlänglich präzise **Vorhersagen (Prognosen)** über das zukünftige Verhalten und Erleben bzw. deren Veränderung
- **Entscheidungen** bezüglich des Merkmalsträgers zu vorzubereiten und zu unterstützen

Diagnostik ist eine Methodenlehre, die das Ziel hat, **Personen „richtig“ zu beurteilen** und zu erfahren, **wie sich Menschen voneinander unterscheiden.**



Semesterplan

Woche	Datum	Thema
01	10.11.2020	Einführung
02	17.11.2020	Entwicklung, Sozialisation und Lernen
03	24.11.2020	Gedächtnismodelle und kognitive Basisfunktionen
04	01.12.2020	Intelligenz
05	08.12.2020	Selbstkonzept
06	15.12.2020	Motivation
07	22.12.2020	Diagnostik und Evaluation
08	12.01.2021	Diagnostik und Testverfahren
09	19.01.2021	Selbstregulation
10	26.01.2021	Lernstrategien
11	02.02.2021	Unterrichtsmodelle
12	09.02.2021	Unterrichtsqualität
13	16.02.2021	Digitale Technologien
14	23.02.2021	Wiederholung
15	02.03.2021	Klausur



Abbildungsverzeichnis

- Abb. 1: https://www.km.bayern.de/bilder/km_absatz/foto/932_mnnchen_auf_leiter.jpg
- Abb. 2: Brunner, M., Stanat, P., & Pant, H.A. (2014). Diagnostik und Evaluation. In T. Seidel & A. Krapp (Eds.), *Pädagogische Psychologie*, 489.
- Abb. 3: Lehrevaluation Pädagogische Psychologie, Wintersemester 2019/20
- Abb. 4: Lehrevaluation Pädagogische Psychologie, Wintersemester 2019/20
- Abb. 5: Dr. Gero Tacke, „Lese- Rechtschreibförderung: Informationen und Fördermaterialien“, <http://www.leserechtschreibfoerderung.de/Informationen/Beruf.html>
- Abb. 6: Diedrich, Jennifer & Schiepe-Tiska, Anja & Ziernwald, Lisa & Tupac-Yupanqui, Ana & Weis, Mirjam & McElvany, Nele & Reiss, Kristina. (2019). Lesebezogene Schülermerkmale in PISA 2018: Motivation, Leseverhalten, Selbstkonzept und Lesestrategiewissen. https://www.researchgate.net/publication/338713822_Lesebezogene_Schulermerkmale_in_PISA_2018_Motivation_Leseverhalten_Selbstkonzept_und_Lesestrategiewissen
- Abb. 7: Technische Universität Darmstadt, „STAT5 Regression“, <http://wiki.ifs-tud.de/fm/stat/stat05>, Direktlink zum Bild: http://wiki.ifs-tud.de/_detail/fm/stat/r_linearerzusammenhang.png?id=fm%3Astat%3Astat05
- Abb. 8: Chart by Tyler Vigen, „Spurious correlations“, <http://www.tylervigen.com/spurious-correlations>
- Abb. 9: Bild von [PublicDomainPictures](#) auf [Pixabay](#). <https://pixabay.com/de/photos/lesen-buch-junge-kinder-kind-316507/>



Abbildungsverzeichnis

- Abb. 10: Image by Gerd Altmann from Pixabay. Link zum Bild: <https://pixabay.com/illustrations/problem-arrow-direction-5464444/>
- Abb. 11: Quelle: Sacher, W.: Leistungen entwickeln, überprüfen und beurteilen. Bewährte und neue Wege für die Primar- und Sekundarstufe, Klinkhardt, Bad Heilbrunn 2014, S. 111
- Abb. 12: Quelle: Sacher, W.: Leistungen entwickeln, überprüfen und beurteilen. Bewährte und neue Wege für die Primar- und Sekundarstufe, Klinkhardt, Bad Heilbrunn 2014, S. 106
- Abb. 13: nationales MINT Forum, PISA Studie 2018: Grundbildung im internationalen Bereich: https://www.nationalesmintforum.de/fileadmin/medienablage/content/themen/aktuelles/2018/aktuelle-studien/2019/PISA_2018_Berichtsband_online_29.11.pdf
- Abb.14: Fischer, C., Xu, D., Rodriguez, F., Denaro, K., & Warschauer, M. (2020). Effects of course modality in summer session: Enrollment patterns and student performance in face-to-face and online classes. The Internet and Higher Education, 45, 1-9. <https://doi.org/10.1016/j.iheduc.2019.100710>



Danke.