



PinguinSystem
T. Finke

Diagnostische Methoden und Verfahren
Datenerhebung

Borken, den
Name:

Grundlagen der Datenerhebung

Messen, Messskalen	2-4
Deskription (Beschreibung) von Messwerten	5-9
Gütekriterien von Messungen	10-14

 <p>PinguinSystem T. Finke</p>	<p>Diagnostische Methoden und Verfahren Datenerhebung</p>	<p>Borken, den Name:</p>
---	--	------------------------------

Messen

- Bestimmung der Ausprägung eines Merkmals (empirisches Relativ)
- Zuordnung von Zahlen (numerisches Relativ)
- Homomorphe Abbildung (empirisches → numerisches Relativ)

→ Messvorgang und Messskala müssen auf das zu messende Merkmal abgestimmt sein

Wie differenziere ich Merkmalen?

Welche Typen von Messskalen gibt es?

Differenzierung von Merkmalen

- **Qualitative** Merkmale = Ausprägung qualitativ beschreibbar
z. B. Geschlecht, Schultyp, Schulfach, soziales Herkunftsmilieu
- **Quantitative** Merkmale
= Ausprägung quantifizierbar über **direkte** Messung
z. B. Alter, Lernzeit, Dienstjahre, Größe, Gewicht
oder
= Ausprägung quantifizierbar nur über **indirekte** Messung
z. B. Schulleistung, Kompetenz, Fertigkeit, Intelligenz, Motivation
- Das Merkmale ist **direkt** erfassbar
→ Theorieebene ⇒ Beobachtungsebene
z. B. chronologisches Alter, Berufsjahre, Schuljahre, Geschlecht, Schultyp, Schulfach
- Das Merkmale ist nur **indirekt** erfassbar („Konstrukte“)
→ Theorieebene ⇒ Indikatoren ⇒ Beobachtungsebene
z. B. soziales Alter, soziales Herkunftsmilieu, Schulleistung, Kompetenz, Wissen, Fähigkeit, Fertigkeit, Intelligenz, Motivation, Interesse, Selbstkonzept etc.

 PinguinSystem T. Finke	Diagnostische Methoden und Verfahren Datenerhebung	Borken, den Name:
---	---	----------------------

Typen von Messskalen → Unterschiede im „Messniveau“

1. **Klassifikationsskala** bzw. Nominalskala
qualitative Kategorisierung nach „Gleichheit“ der Merkmalsausprägung
z. B. weiblich = 1; männlich = 2
2. **Rangskala** bzw. Ordinalskala
Abbildung einer qualitativen Rangordnung nach Merkmalsausprägung,
z. B. sehr gut = 1; gut = 2; befriedigend = 3; ausreichend = 4
3. **Intervallskala**
quantifizierend nach Ausprägungsgrad (gleiche Abstände), z. B. IQ = 80, 90, 100 ...
4. **Verhältnisskala** bzw. Proportionalskala (bzw. Ratioskala)
quantifizierend nach Ausprägungsgrad (gleiche Abstände), ausgehend von einem absoluten („natürlichen“) Nullpunkt, z.B. Alter = 1, 2, 3 Jahre

Skalentyp	Voraussetzungen (Merkmal)	stat. Kennwerte zur Deskription von Messungen bei Stichproben (Gruppen)	zuläss. Mathem. Operationen mit den Skalenwerten
Klassifikationskala	a) Äquivalenz (Gleichheit)	der Merkmalsausprägung → kategorisierbar	keine
Rangskala	a) Äquivalenz b) Aussagen z. Beziehungen zw. Merkmalsklassen	Modalwert Median (zentrale Tendenz) Quartile (Streuemaß)	> <
Intervallskala	a) Äquivalenz b) Beziehung c) gleiche Intervalle	Modalwert Median Arithmetisches Mittel (zentr. Tend.) Standardabweichung (Streuemaß)	Addition/ Subtraktion
Verhältnisskala	a) Äquivalenz b) Beziehung c) gleiche Intervalle d) absoluter Nullpunkt	Modalwert Median Arithmetisches Mittel (zentr. Tend.) Standardabweichung (Streuemaß)	Multiplikation/ Division



Skalenniveaus → Informationsgehalt der Messung



Ein hoher Informationsgehalt ist anzustreben!

Beispiel	Skalenniveau
Geschwindigkeit in km/h	Ratioskala
Geburtsort	Nominalskala
Schulnoten	Ordinalskala
Alkoholpegel in Promille	Ratioskala
Religionszugehörigkeit	Nominalskala
Zustimmung auf einer Skala von 1-5	Ordinalskala
Kontostand in Euro	Intervallskala
An Punkten gewinnendes Börsenbarometer	Intervallskala
Top Ten der Charts	Ordinalskala
Familienstand	Nominalskala
Konfession	Nominalskala
Absicht, an einer Wahl teilzunehmen	Nominalskala
Temperatur in Grad Celsius	Intervallskala
Geschlecht	Nominalskala
Augenfarbe	Nominalskala
Temperatur in Grad Fahrenheit	Intervallskala
Einkommen	Ratioskala
Schuljahre	Ratioskala
Dauer einer Freundschaft	Ratioskala
Temperatur in Grad Kelvin	Ratioskala
Telefonnummer	Nominalskala
Haarfarbe	Nominalskala
Geschmacksrichtung	Nominalskala
Kontonummer	Nominalskala
Ränge der Polizei	Ordinalskala

Erklärung und Aufgaben der Uni Kassel.
Unbedingt bearbeiten!!

 PinguinSystem T. Finke	Diagnostische Methoden und Verfahren Datenerhebung	Borken, den Name:
---	---	----------------------

Deskription (Beschreibung) von Messwerten

sind abhängig vom Skalenniveau

Beispiel:

Leseverständnis (abstraktes Merkmal = Konstrukt)

indirekt erfasst mit einem Leseverständnistest n. GatesMc/Ginitie (2000)

- 11 Texte
- 48 Items (= Testaufgaben, Auswahl-Antwortform mit drei Distraktoren)
- Anzahl richtig beantworteter Aufgaben → Indikator für Leseverständnis

Text 11:

Jede Auflistung von symbiotischen Beziehungen würde zeigen, dass sie besonders häufig in der hochorganisierten, unpersönlichen Welt der Insekten vorkommen. Zum Beispiel ist die Geschichte von Ameisen, die ihre Milchkühen vergleichbaren Blattläuse schützen und „melken“, sehr bekannt. Sehr viel seltener lassen sich Symbiosen bei warmblütigen Wirbeltieren feststellen, und symbiotische Beziehungen, in denen sich taxonomische Klassen überschneiden, wie z.B. zwischen Vögeln und Säugetieren, sind besonders selten.

Frage 48. Der Abschnitt stellt Insektengesellschaften dar als

Item ←

- geordnet.
- hoch motiviert.
- klein im Umfang.
- hierarchisch



Formen der Deskription von Messwerten, z. B. Häufigkeitsverteilung einer Stichprobe

- **numerisch**, z. B. Häufigkeitstabelle zur Anzahl richtig gelöster Items
- **graphisch**, z. B. Balkendiagramm, Histogramm
- **verbal** → durch statistische Kennwerte (Maße zentraler Tendenz und Streuungsmaße)

Beispiel: Ergebnisse Leseverständnistest (N = 302)

Stichprobengröße
Anzahl der teilnehmenden
Schülerinnen und Schüler
(SuS)

numerisch: Häufigkeitstabelle

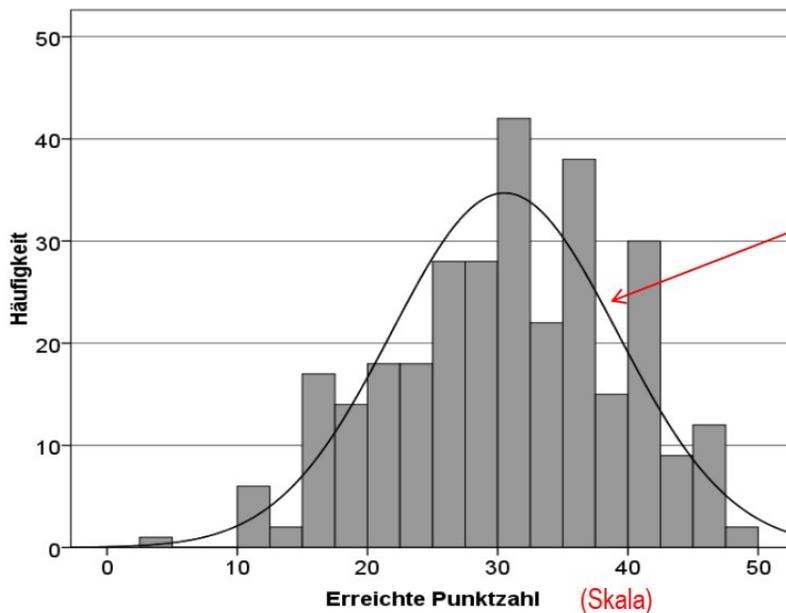
absolute Anzahl
Schüler/innen je Punktzahl

relativer Anteil SuS in der
Stichprobe je Punktzahl

Punkte	Häufigkeit „f“	Prozent (N=302)	Kumulierte Prozent
10	1	0,33	0,33
11	6	1,99	2,32
12	2	0,66	2,98
13	18	5,96	8,94
14	15	4,97	13,91
15	18	5,96	19,87
..



graphisch: Histogramm mit Normalverteilung



Mittelwert = 30,51
Standardabweichung = 8,678
N = 302

Normalverteilung

Balkendiagramm und Histogramm

Ein **Balkendiagramm** entsteht, wenn einzelne Ergebnisse als Balken abgebildet werden. Dabei liegt der Balken an der Stelle des Ergebnisses z.B. erreichte Punkte und die Länge des Balkens beschreibt die Häufigkeit dieses Ereignisses (so viel Prozent der Teilnehmer haben diese Anzahl von Punkten erreicht).

Ein **Histogramm** liegt vor, wenn alle Balkenlängen zusammen 100% ergeben (es wurden alle Teilnehmer berücksichtigt) und alle Balkenbreiten zusammen 100% ergeben (es wurden alle Ergebnismöglichkeiten berücksichtigt).

 <p>PinguinSystem T. Finke</p>	<p>Diagnostische Methoden und Verfahren Datenerhebung</p>	<p>Borken, den Name:</p>
---	---	------------------------------

Normalverteilung

Entstehung: Wir tragen die Häufigkeiten eines Zufallsereignisses zu einem Histogramm auf.

Nun legen wir (**nach einer bestimmten Formel**) eine Kurve, sodass die Balken bestmöglich „in“ die Kurve „passen“. Diese Kurve nennen wir Normalverteilung, Gauß-Verteilung oder Gaußsche Glockenkurve, nach dem Mathematiker Carl Friedrich Gauß und der Form der Kurve, die an eine Glocke erinnert.

Zur **Veranschaulichung** wie sich Diagramm und Glockenkurve verhalten.

In diesem Zusammenhang wird auch der Begriff **Binomialverteilung** oft genannt. „Sie beschreibt die Anzahl der Erfolge in einer Serie von gleichartigen und unabhängigen Versuchen, die jeweils genau zwei mögliche Ergebnisse haben („Erfolg“ oder „Misserfolg“). Solche Versuchsserien werden auch Bernoulli-Prozesse genannt“ (Wikipedia), nach dem Mathematiker Daniel Bernoulli.

Beispiel: Wir haben einen Würfel und würfeln. Jeder Wurf ist von der Art her gleich und unabhängig vom Ergebnis des vorherigen Wurfs (gleichartig und unabhängig).

Nun legen wir fest, dass eine 6 ein Erfolg ist und alle anderen Zahlen ein Misserfolg. Wir haben nun zwei Zustände. „Binom“ kommt aus dem Lateinisch und bedeutet „bi“ „zwei“ und „nomen“ = „Namen“. Dieses System kennen wir auch aus der Informatik (0,1).

Binomialverteilungen werden verwendet, wenn wir ganzzahlige Ereignisse haben. Beim Würfeln wäre das 1, 2, 3, 4, 5 und 6. Würden wir die Körpergröße einer Versuchsgruppe oder den Durchmesser von Schrauben darstellen wollen, können wir auch Dezimalzahlen (Kommazahlen) erhalten. Diese können nur mit einer Normalverteilung abgebildet werden.

Noch ein Video zur **Veranschaulichung** .

Wir können auch eine Kurve über die Balken legen, die noch besser passen würde, aber diese standardisierte Glockenform der Kurve hat einige Vorteile. Wir können recht einfach erkennen und berechnen wie viel Prozent der Teilnehmer sich in welchem Ergebnisbereich aufhalten.

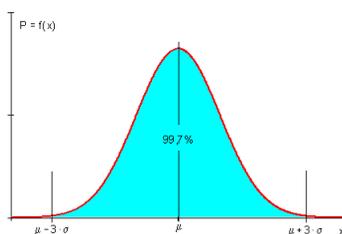
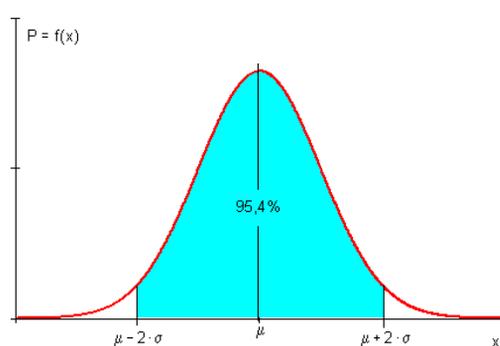
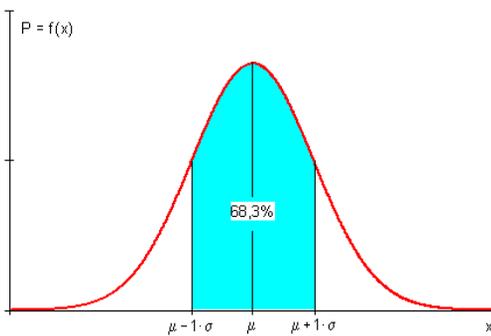
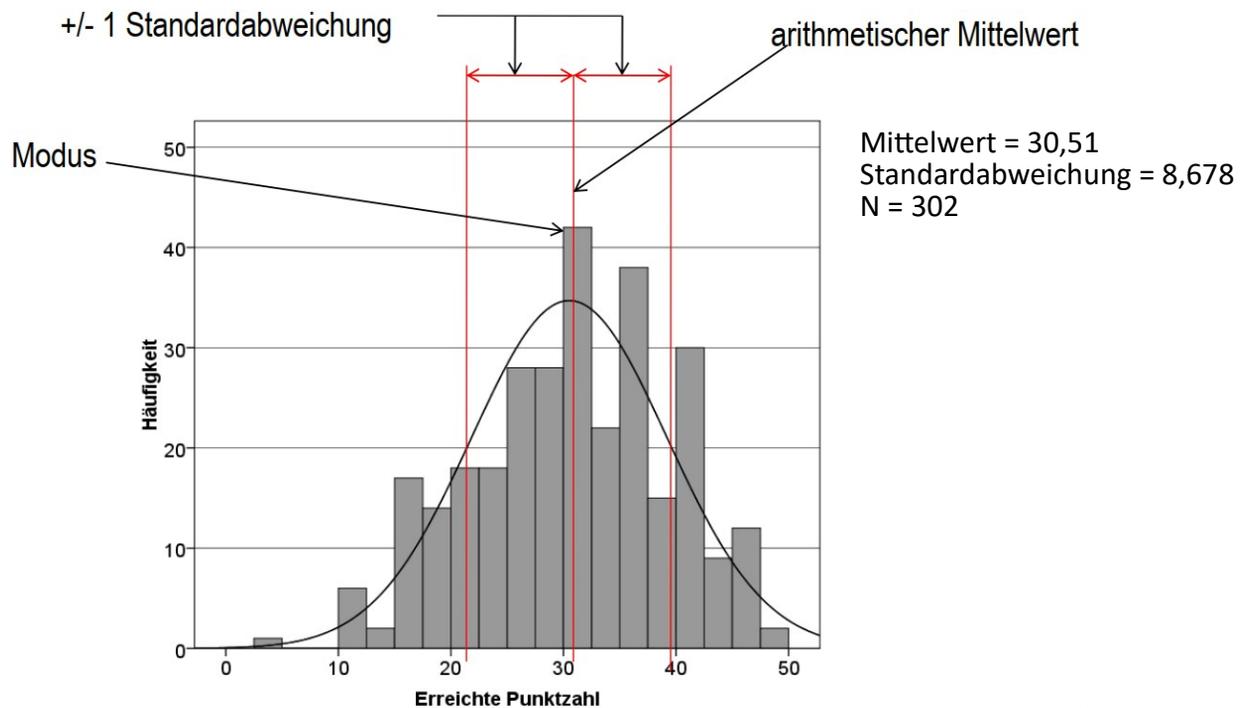


verbal → durch „statistische Kennwerte“

Beispiel:

Beim Leseverständnistest konnten zwischen 0 und 48 Punkten erreicht werden. Die **Verteilungsform** der Stichprobe ist unipolar und annähernd **normalverteilt (nv)**. Der **Mittelwert** (zentrale Tendenz) liegt bei 30,51 Punkten, bei einer **Standardabweichung** von 8,68 (Streumaß).

„normalverteilt“ → zwei Drittel der Stichprobe liegen zwischen +/- einer Standardabweichung (haben zwischen 22 und 39 Punkten erreicht).



Die Standardabweichung σ ist die steilste Stelle in der Glockenkurve. Mit einer mathematischen Funktion ist sie leicht zu bestimmen.

 PinguinSystem T. Finke	Diagnostische Methoden und Verfahren Datenerhebung	Borken, den Name:
---	--	----------------------

Gütekriterien sozialwissenschaftlicher Messungen

Hauptgütekriterien

- **Objektivität**
↓
- **Reliabilität (Zuverlässigkeit)**
↓
- **Validität (Gültigkeit)**

Nebengütekriterien

- Normierung
- Vergleichbarkeit
- Nützlichkeit
- Zumutbarkeit
- Unverfälschbarkeit

Qualitätsstandards (DIN 33430)

 PinguinSystem T. Finke	Diagnostische Methoden und Verfahren Datenerhebung	Borken, den Name:
--	---	------------------------------------

Objektivität

Wird die Messung durch den Messvorgang beeinflusst?

Voraussetzung für Reliabilität und Validität!

Objektivität ist gegeben, wenn die Messergebnisse unabhängig vom Messenden und der Messsituation sind.

Durchführungsobjektivität

Das Ergebnis ist unabhängig davon, wer die Messung durchführt.

→ Gleiche Anforderungen unter gleichen Durchführungsbedingungen

- zeitlich, räumlich
- Erläuterungen zur Bearbeitung
- Bearbeitungszeiten
- Hilfsmittel

Auswertungsobjektivität

- Das Ergebnis ist unabhängig davon, wer die Messung auswertet.
- Ein gleiches Verhalten wird von unterschiedlichen Beobachtern gleich gewertet (mit dem gleichen „Messwert“ beziffert)

Interpretationsobjektivität

- Ergebnis ist unabhängig davon, wer den Messwert interpretiert.
- Die Interpretation bezieht sich auf eine „allgemeine“ Bezugsnorm.

 <p>PinguinSystem T. Finke</p>	<p>Diagnostische Methoden und Verfahren Datenerhebung</p>	<p>Borken, den Name:</p>
---	---	------------------------------

Exkurs zur Auswertungsobjektivität

Forschungsbefunde zur Auswertungsobjektivität mündlicher Prüfungen (Birkel 1978; Birkel/Pritz 1980; Pritz 1981)

- unterschiedliche Schwierigkeitsgrade von Fragen im Klassenunterricht
- abweichende Urteile von Prüfenden:
- Einflussfaktoren auf die Urteilsbildung
- Kontrasteffekte (8%)
- Informationen zu Vorzensuren (7%, 2%)
- Sprachgeschwindigkeit (17%)
- Lehr- und Prüferfahrungen der Bewertenden

Maßnahmen zur Verbesserung der Messgüte mündlicher Prüfungen:

- nur sprachgebundene Leistungen mündlich prüfen
- Vorbereitung (auch informeller Prüfungen)
 - Lernzielbezogene Indikatoren festlegen
 - Schwierigkeitsgrade der Fragen beachten
- Reihenfolgeeffekte beachten
- für entspannte Atmosphäre sorgen
- „Eisbrecherfragen“ bereithalten

Auswertungsobjektivität mindernde Faktoren bei schriftlichen Prüfungen, z. B.

- Vorinformationen (z. B. Weiss 1995)
- Länge der Textproduktion (z. B. Birkel/Birkel 2002)
- grammatikalische und orthografische Fehler (z. B. Birkel/Birkel 2002)
- Handschrift (z. B. Briggs 1970)
- Reihenfolgeeffekte (z. B. Baurmann 1995)
- Beliebtheit der Schüler/innen (Hadley 1954, 1995)
- Differenzen i. d. Bewertung schriftlicher und mündlicher Leistungen (Grabowski 1999)

Reliabilität (Zuverlässigkeit)

Wie zuverlässig ist die Messung?

heißt: Mit welcher Sicherheit und Genauigkeit misst das Messinstrument (z.B. der Fragebogen)?

Reliabilität bei Messwiederholung (gleiche Bewerter)

		2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
18	Erste Einstufung											1						
17																		
16																		
15				1					1	1	1				2			
14				1						1								1
13			1		1	1			1									
12				1		1	2			1	1	1	2		2			
11	Mittelwert = 10,4									1	2	1						
10				1	2				2	3		3	2					
9								1	1	1	1							
8				1		1	1			1	1							
7							2			1								
6						1	1											
5											3							
4																		
3	Zweite Einstufung	1					1											
2																		

Studie: Bewertung der gleichen Arbeiten (drei Kurzaufsätze) in Geografie zu unterschiedlichen Zeiten (11 Wochen), n = 61 Lehrer (Eells 1930, 1995)

→ r = 0,25

Reliabilitäts- bzw. Zuverlässigkeitskoeffizienten (Lienert/Raatz 1998)
 individuelle Leistungen $r_{\min} = 0,9$
 Gruppenvergleiche r = 0,5 bis 0,7 (Mittelwerte)

Wichtiger Reliabilitätskennwert bei individuellen Messungen:

Standardmessfehler (SMF) = Messgenauigkeit des Messinstruments

z. B. SMF +/- 4

→ bei 24 Punkten im Test liegt der „wahre“ Leistungswert der/des SuS zwischen 20 und 28 Punkten

→ Angenommener Messfehler des Zensurensystems: +/- 1 Notenstufe!!

Interne Konsistenz

„Skala“ zur Erfassung „introjizierter“ Motivation (vgl. Prenzel et al.1996)

In der letzten Unterrichtsphase

... versuchte ich, alles so zu erledigen, wie es von mir erwartet wird ① ① ~~②~~ ③ ④ ⑤

... habe ich mich angestrengt, wie sich das für ordentliche Auszubildende gehört ① ① ② ③ ④ ~~⑤~~

... habe ich mich selbst unter Druck gesetzt, um alles möglichst richtig zu machen ~~①~~ ① ② ③ ④ ⑤

keine interne Konsistenz!

Probleme bei Messungen an Personen

- instabile Merkmale bzw. Merkmalsausprägungen
- Verzerrungen bei Testwiederholung (z. B. Lerneffekte)

→ Die **Reliabilität** ist abhängig von er **Stabilität** des zu messenden **Merkmals!**

 PinguinSystem T. Finke	Diagnostische Methoden und Verfahren Datenerhebung	Borken, den Name:
--	---	------------------------------------

Validität (Gültigkeit)

Wird tatsächlich das Merkmal gemessen, was gemessen werden soll?

bzw.

Ist das diagnostische Verfahren für den Zweck, für den es verwendet werden soll, tatsächlich aussagefähig?