

Prof. Dr. Christoph Kulgemeyer

AG Didaktik der Physik

Fakultät für Naturwissenschaften



Validität



Ziele der Lehrveranstaltung

- Überblick über grundlegende Ideen und Techniken zur Überprüfung von:
 - Inhaltsvalidität
 - Kriteriumsvalidität
 - Konstruktvalidität

Was ist Validität?

- Alle sind sich einig, dass Validität von höchster Wichtigkeit ist - aber was Validität ist, darüber wird lebhaft diskutiert (z.B. Journal of Educational Measurement)
- Gängigste Definition von Messick 1989 (übersetzt von Hartig 2008):

„Validität ist ein integriertes bewertendes Urteil über das Ausmaß, in dem die Angemessenheit und die Güte von Interpretationen und Maßnahmen auf Basis von Testwerten oder anderen diagnostischen Verfahren durch empirische Belege und theoretische Argumente gestützt sind.“

 - D.h. Validität betrifft die Frage, warum Schlüsse aus einer Erhebung überhaupt gerechtfertigt sind.
 - Validität ist kein Gütekriterium eines Tests, sondern eines Mess- und Auswertungsprozesses

Wie kann ich meine Schlüsse rechtfertigen?

- Beispiel:
 - Ich teste das physikalische Vorwissen von Studienanfängern.
 - Ich finde, dass die Neu-Studierenden sehr schlecht abschneiden.
 - Ich schließe, dass die Schulen physikalisches Fachwissen schlecht vermitteln

Ist dieser Schluss gerechtfertigt?

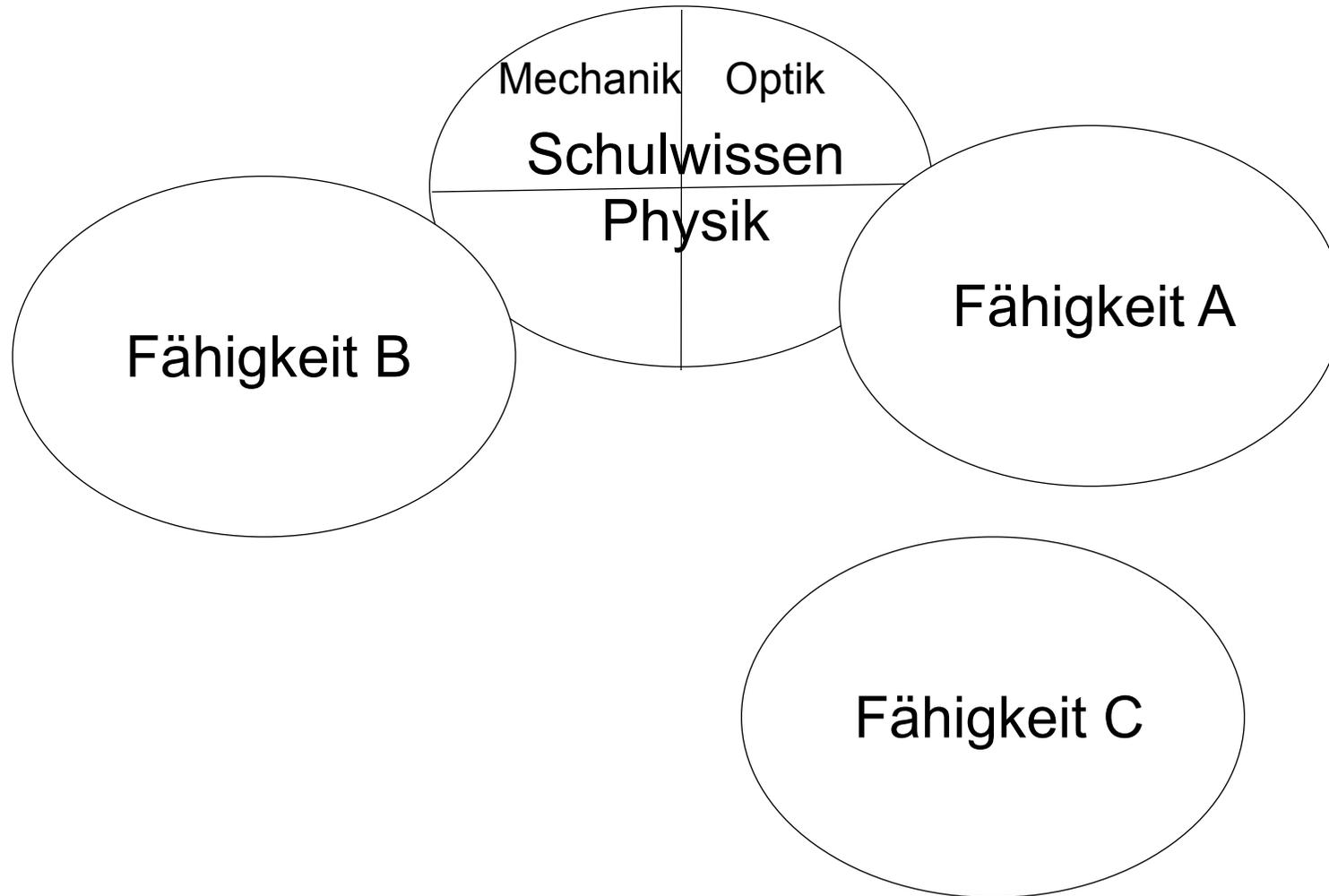
Drei (gängige) Arten von Validität

- **Inhaltsvalidität:** Deckt mein Test das zu testende Konstrukt ab?
- **Kriteriumsvalidität:** Wie hängt das Konstrukt mit externen Faktoren zusammen?
- **Konstruktvalidität:** Wird die Struktur meines Konstrukts durch das Messinstrument reproduziert?

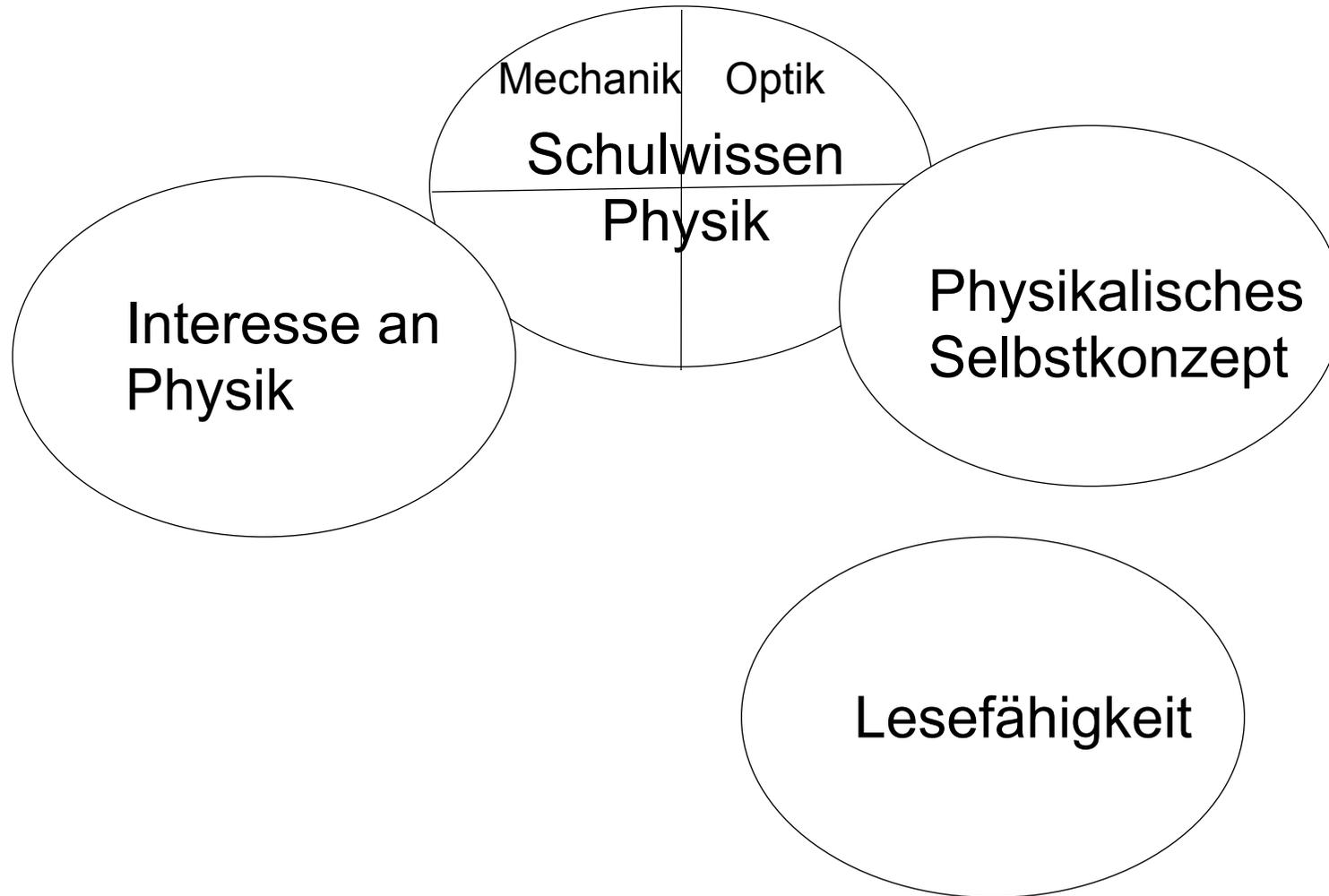
Erster Schritt! Ganz wichtig!!

- Bevor ich mir Gedanken über das Instrument machen kann, muss ich wissen, was ich messen will - Theorie vor Empirie
 - theoretisch umfassende Beschreibung des Konstrukts
 - Überlegungen über Subskalen und ihre gegenseitige Abhängigkeit (Struktur), Überlegungen über Zusammenhang zu anderen Merkmalen
 - Hier: Beschreibung des in der Schule vermittelten Fachwissens der Physik
 - Lehrpläne, Abiturklausuren, Expertenbefragungen (wer ist Experte?)

Resultat: Nomologisches Netz



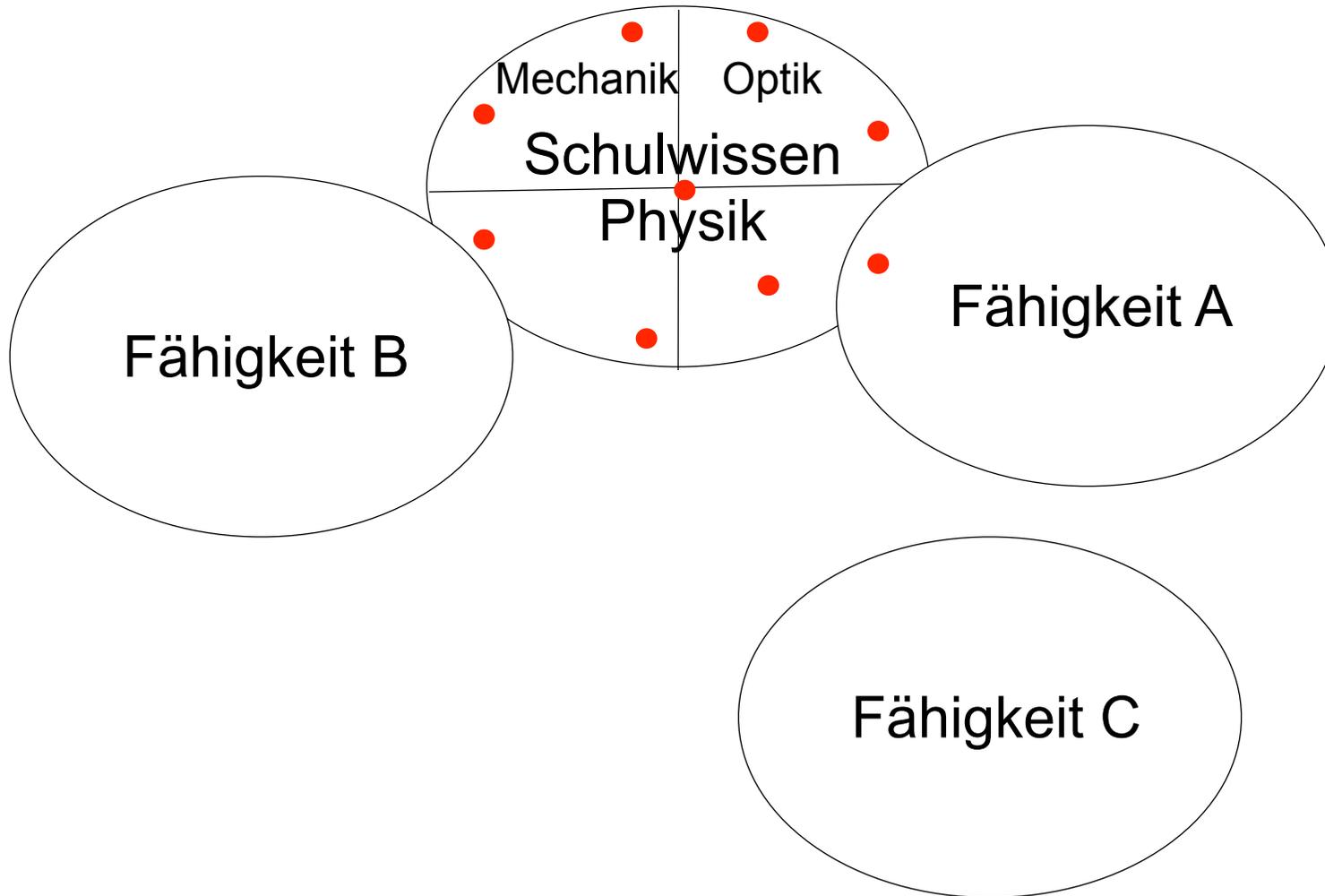
Resultat: Nomologisches Netz



Inhaltsvalidität

- Entwicklung von Items, die alle Bereiche des Konstrukts abdecken
- Items werden verschiedenen Bereichen des Konstrukts zugeordnet, dies kann überprüft werden
- Expertenrating für Itemzuordnungen
- Ganz wichtig: die Zielgruppe soll die Items wie intendiert verstehen - z.B. Lautes Denken, Stimulated Recall

Resultat

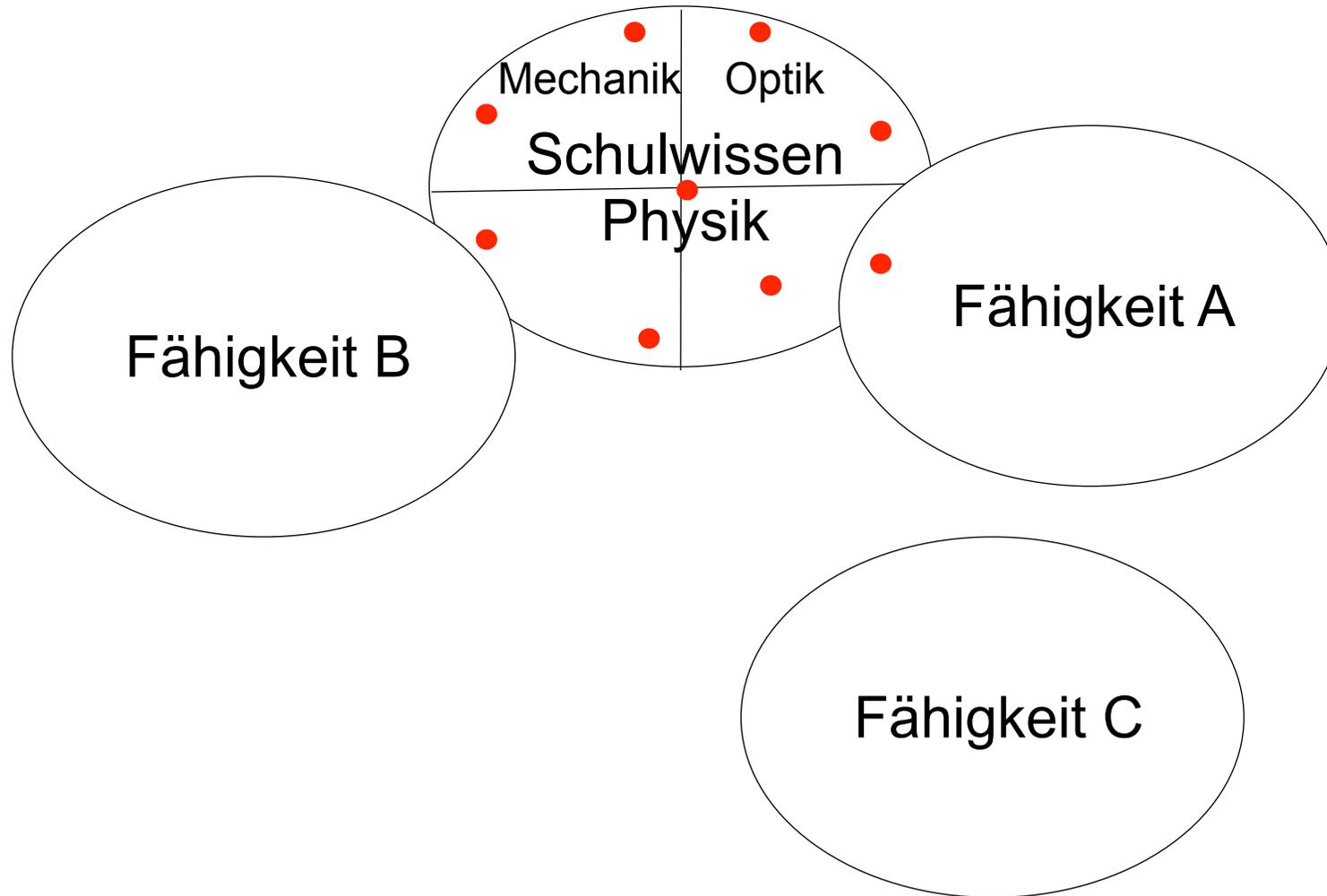


● Items

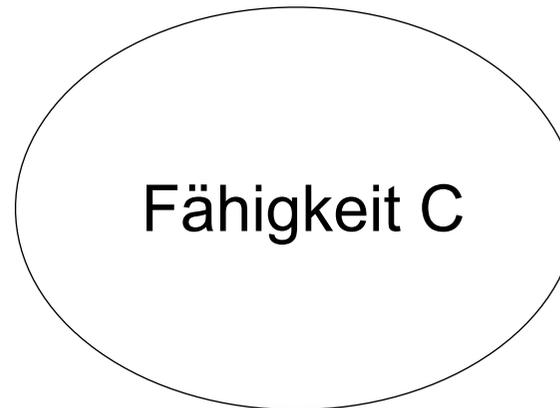
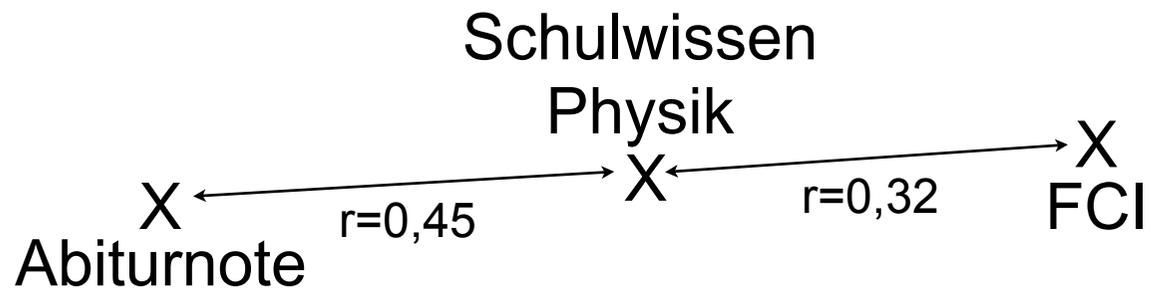
Kriteriumsvalidität

- Zusammenhang der Testergebnisse (!) mit anderen Merkmalen
 - Übereinstimmungsvalidität: Zusammenhang mit anderen Testinstrumenten für das gleiche / ein verwandtes Konstrukt (z.B. FCI...)
 - Vorhersagevalidität: Prognose von späteren Ergebnissen, die auf den gemessenen Fähigkeiten beruhen (z.B. Studienerfolg)
 - Retrospektive Validität: Zusammenhang zu vorherigen Ergebnissen (z.B. Abiturnote)
- Kriteriumsvalidität kann durch Korrelationen ausgedrückt werden (nicht zu hohe Zusammenhänge...)

Resultat



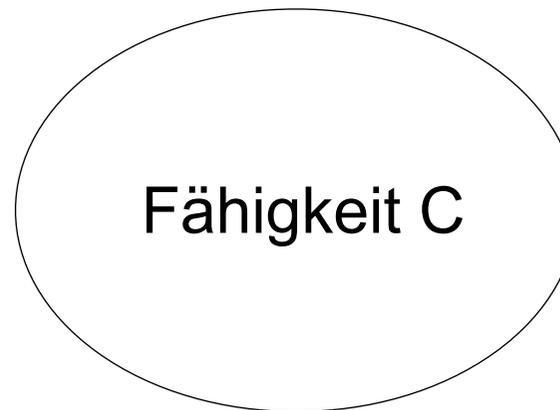
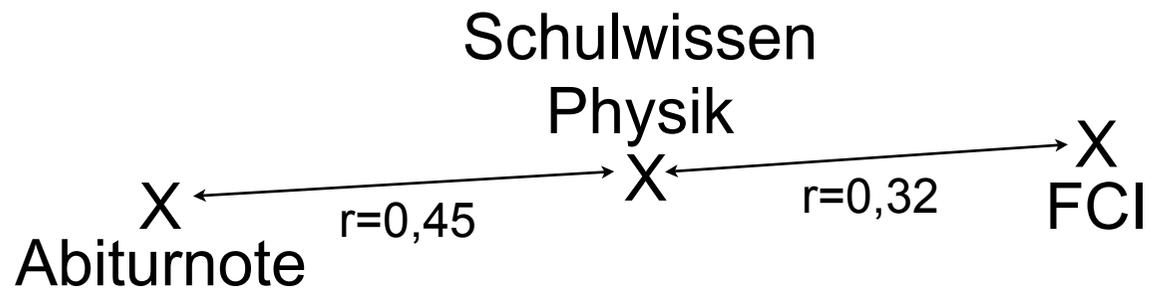
Resultat



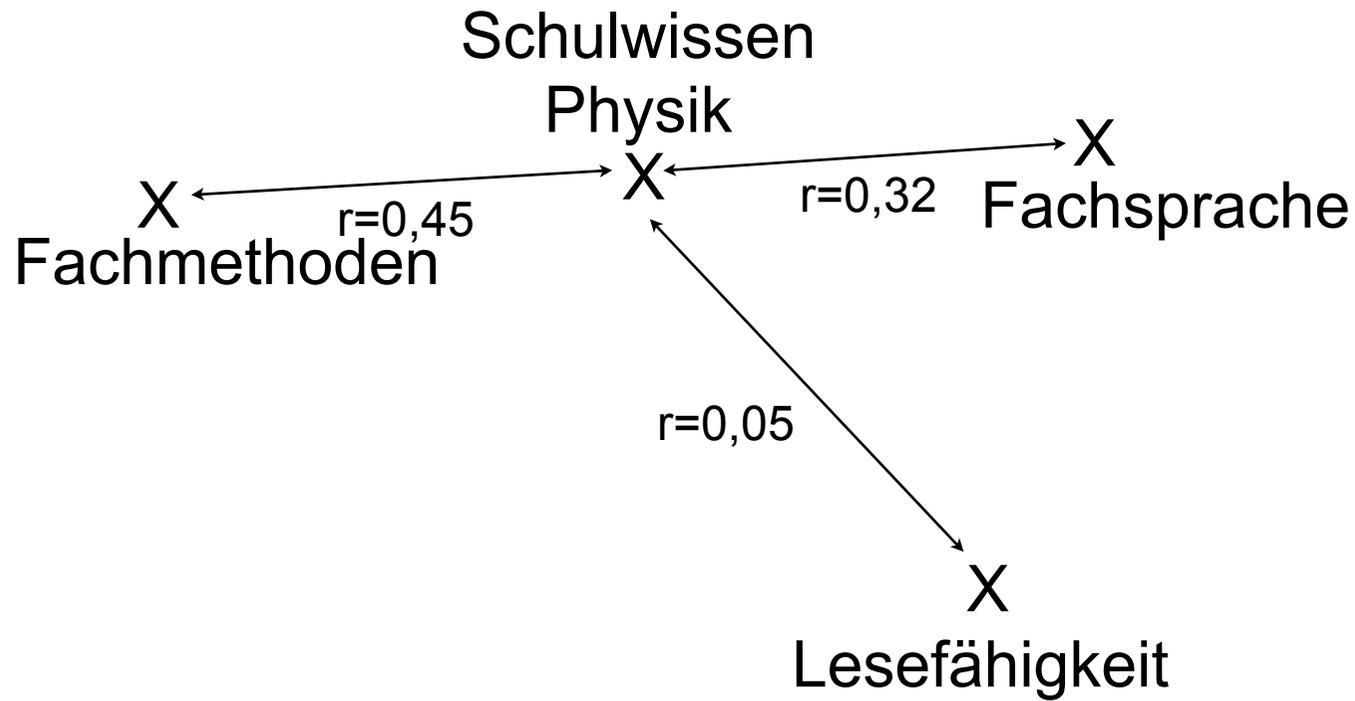
Konstruktvalidität

- Konstruktvalidität als umfassendste Form von Validität
- Nomologisches Netz soll durch Testergebnisse reproduziert werden, innere Struktur des Konstrukts soll reproduziert werden
- Konvergente/ Diskriminante Validität

Resultat



Resultat



Konstruktvalidität

- Konstruktvalidität als umfassendste Form von Validität
- Nomologisches Netz soll durch Testergebnisse reproduziert werden, innere Struktur des Konstrukts soll reproduziert werden
- Konvergente/ Diskriminante Validität
- Multitrait-Multimethod-Matrix
- Konstruktinterne Hypothesen (z.B. über Itemschwierigkeiten)
- Problem: was ist Hypothese, was Voraussetzung?
Worauf kann man unter keinen Umständen verzichten

Zurück zum Beispiel

- Beispiel:
 - Ich teste das physikalische Vorwissen von Studienanfängern.
 - Ich finde, dass die Neu-Studierenden sehr schlecht abschneiden.
 - Ich schließe, dass die Schulen physikalisches Fachwissen schlecht vermitteln

Ist das alles überprüft, habe ich gute Argumente für meinen Schluss. Validitätsnachweise sind evidenzgestützte Argumente!

Es bleibt die Frage: wieviel Validität muss ich überprüfen, damit die Fachgemeinschaft mir meine Schlüsse abnimmt?