

Karl-Heinz Renner

Testkonstruktion, Testen und Entscheiden

Fakultät für
Psychologie

Das Werk ist urheberrechtlich geschützt. Die dadurch begründeten Rechte, insbesondere das Recht der Vervielfältigung und Verbreitung sowie der Übersetzung und des Nachdrucks, bleiben, auch bei nur auszugsweiser Verwertung, vorbehalten. Kein Teil des Werkes darf in irgendeiner Form (Druck, Fotokopie, Mikrofilm oder ein anderes Verfahren) ohne schriftliche Genehmigung der FernUniversität reproduziert oder unter Verwendung elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden. Wir weisen darauf hin, dass die vorgenannten Verwertungsalternativen je nach Ausgestaltung der Nutzungsbedingungen bereits durch Einstellen in Cloud-Systeme verwirklicht sein können. Die FernUniversität bedient sich im Falle der Kenntnis von Urheberrechtsverletzungen sowohl zivil- als auch strafrechtlicher Instrumente, um ihre Rechte geltend zu machen.

Der Inhalt dieses Studienbriefs wird gedruckt auf Recyclingpapier (80 g/m², weiß), hergestellt aus 100 % Altpapier.

Inhaltsverzeichnis

Inhaltsverzeichnis	3
Einführung zum Kurs	5
Lehr- und Lernziele.....	8
Teil 1: Testkonstruktion: Vertiefung.....	9
1 Wie man einen Test vor dem Hintergrund der Klassischen Testtheorie konstruiert und anwendet	10
2 Item-Response-Theorie (IRT) und probabilistische Testtheorie (PTT).....	11
2.1 Unzulänglichkeiten der KTT	11
2.2 Probabilistische Testmodelle für dichotome und ordinale Itemantworten	13
2.3 Das 1 PL-Modell bzw. Rasch-Modell	14
2.4 Anwendungsbeispiele	17
3 Adaptives Testen	19
4 Latent-Class-Analyse	22
5 Latent-State-Trait-Theorie	25
6 Innovative Methoden der Testgestaltung: Diagnostik und Datenerhebung im Internet	27
Teil 2: Testen und Entscheiden	28
7 Personbeurteilung im Alltag.....	29
8 Klinische vs. statistische Urteilsbildung	31
9 Der diagnostische Prozess.....	34
10 Entscheidungstheoretische Modelle.....	37
11 Literatur.....	41

Abbildungsverzeichnis

Abbildung 1: Wahrscheinlichkeitsfunktion im Rasch-Modell.....	15
Abbildung 2: Rückkopplungsschleifen im diagnostischen Prozess (Quelle: Schmidt-Atzert & Amelang 2012, S. 387).....	35

Einführung zum Kurs

Bitte lesen Sie diese Einführung sorgfältig durch!

Der Kurs „Testkonstruktion, Testen und Entscheiden“ wiederholt im ersten Teil zunächst die wichtigsten Schritte der Klassischen Testtheorie (KTT) sowie der Anwendung von Tests bzw. Fragebögen, die vor dem Hintergrund der KTT konstruiert wurden. Im Mittelpunkt des ersten Teils steht die Item-Response- Theorie (IRT) bzw. stehen probabilistische Testmodelle, die eine Alternative zur KTT darstellen und Voraussetzungen für adaptives Testen schaffen. Zudem werden Sie innovative Methoden der Testgestaltung kennenlernen, die bei der Datenerhebung im Internet möglich sind.

Im Modul 7 des B.Sc. Psychologie wurden aus einschlägigen Definitionen folgende Bestimmungsstücke der Psychologischen Diagnostik abgeleitet:

1. Die Psychologische Diagnostik ist eine Methodenlehre im Dienste der Angewandten Psychologie.
2. Ihr Gegenstand ist die gezielte und regelgeleitete Sammlung und Verarbeitung von Daten, die für die Bearbeitung von Fragestellungen relevant sind. Spezielle diagnostische Verfahren (z. B. Tests, Fragebögen, Interviews, Verhaltensbeobachtung), die sowohl empirisch und praktisch bewährt als auch theoretisch fundiert sein müssen, werden zur Datensammlung eingesetzt. Da nicht nur in der Angewandten Psychologie, sondern auch in der Grundlagenforschung Daten erhoben und verarbeitet werden, ist die Psychologische Diagnostik auch für die Grundlagendisziplinen, insbesondere die Differentielle Psychologie, bedeutsam.
3. Zu den grundlegenden Fragestellungen der Diagnostik zählen die Beschreibung, Klassifikation, Vorhersage und Evaluation von Unterschieden zwischen und innerhalb von Personen im Hinblick auf psychische Zustände (z. B. Angst), Eigenschaften (z. B. Intelligenz) und deren Veränderungen sowie damit einhergehender relevanter Bedingungen.
4. Psychologische Diagnostik untersucht nicht nur Einzelpersonen, sondern auch Gruppen, Organisationen, Situationen und andere Merkmalsträger.
5. Psychologische Diagnostik bereitet Entscheidungen (z. B. zur Berufseignung, zur Wirksamkeit von Interventionsprogrammen) nach wissenschaftlichen Kriterien und ethischen Standards vor.
6. Psychologisches Diagnostizieren ist ein Prozess, der mehrere Phasen umfasst: Klärung der Fragestellung, Auswahl von psychologisch-diagnostischen Verfahren, Anwendung, Auswertung, Interpretation, Gutachtenerstellung, Interventions- bzw. Maßnahmenvorschlag.

Der zweite Teil des Kurses „Testen und Entscheiden“ fokussiert nun insbesondere die beiden zuletzt genannten Bestimmungsstücke, also psychologisch-diagnostisch fundierte Entscheidungen und den diagnostischen Prozess, der mit solchen Entscheidungen verbunden ist. Diagnostische Entscheidungen basieren auf diagnostischen Urteilen. Solche diagnostischen Urteile werden aus den Informationen gebildet, die mit Hilfe eines – zumeist aber mehrerer diagnostischer Verfahren gewonnen wurden.

Auch im Alltag bilden wir ständig diagnostische Urteile, um das Verhalten anderer Personen vorherzusagen und unser eigenes Verhalten darauf abzustimmen. Diese alltagspsychologische Urteilsbildung funktioniert manchmal erstaunlich gut. Andererseits zeigen viele Studien, dass der „gesunde Menschenverstand“ fehlerbehaftet ist. Da psychologische Diagnostiker/innen „auch nur Menschen“ sind, ist es sinnvoll, sich in der diagnostischen Ausbildung damit auseinanderzusetzen, wie die Personbeurteilung im Alltag funktioniert. In Ihrem bisherigen Bachelorstudium haben Sie bereits Wissen über die alltägliche Urteilsbildung in der Sozialpsychologie und in den Kursen zur Interview- und Beobachtungsmethode (z. B. Fehlerquellen bei der Beobachtung) erworben. Im fortlaufend nummerierten siebten Kapitel zum Teil 2 dieses Kurses wird das Thema Personbeurteilung im Alltag erneut aufgegriffen, systematisiert und vertieft.

Das achte Kapitel wendet sich dann einer langjährigen und immer noch aktuellen Kontroverse zu: der klinischen vs. statistischen Urteilsbildung. Klinische Urteilsbildung stützt sich auf intuitive, auf (klinischem) Erfahrungswissen basierende Informationsverarbeitung. Erhobene Daten werden auf der Basis von (klinischen) Erfahrungen intuitiv zusammengefasst. Bei der statistischen Urteilsbildung erfolgt dagegen eine Kombination und Gewichtung der Vorhersagevariablen anhand von statistischen Beziehungen, die vorher empirisch ermittelt wurden. Auch die Kontroverse um klinische vs. statistische Urteilsbildung haben Sie in Ihrem Bachelorstudium bereits kennengelernt. Wegen der hohen Relevanz für diagnostische Entscheidungen wird das Thema in diesem Kurs durch Pflichtliteratur und auch Originalartikel vertieft, die Studien zur klinischen und statistischen Urteilsbildung enthalten.

Das neunte Kapitel wendet sich dann verschiedenen Modellvorstellungen zum diagnostischen Prozess zu. Einigkeit besteht darüber, dass es sich beim psychologischen Diagnostizieren um einen mehrere Phasen umfassenden, rückgekoppelten Prozess handelt. In den meisten Lehrbüchern zur Psychologischen Diagnostik wird ausgeführt, dass dieser Prozess mit der Beantwortung einer Fragestellung, die eventuell in Form eines Gutachtens erfolgt, abgeschlossen ist. Eine Arbeitsgruppe der *European Association of Psychological Assessment* hat dagegen Richtlinien für den diagnostischen Prozess ausgearbeitet, die explizit auch diagnostische Aspekte bei der Planung, Durchführung und Evaluation von Interventionen berücksichtigen.

Kapitel 10 betrifft dann entscheidungstheoretische Modelle, die darauf abzielen, die Genauigkeit und insbesondere den Nutzen diagnostischer Entscheidungen zu optimieren. Diagnostische Urteile und daraus abgeleitete Entscheidungen haben Konsequenzen für den oder die Auftraggeber und die beurteilten Personen. Beispielsweise werden im Rahmen der Personalselektion Bewerber/innen akzeptiert oder abgelehnt; in der Klinischen Psychologie werden PatientInnen auf der Basis von Diagnosen der einen oder anderen Therapieform zugewiesen und auch in anderen Bereichen der Angewandten Psychologie geht es regelmäßig um die Frage, welches Treatment – in einem weit gefassten Sinne verstanden als jegliche diagnostische und interventionsbezogene Maßnahme, die eine Wirkung nach sich zieht – bei welchen Personen am besten geeignet ist.

Das didaktische Konzept des Kurses basiert auf vier Säulen:

- **Säule 1:** Dieser Studienbrief gibt Ihnen einen Überblick zu den genannten Themen und dient als „advanced organizer“ für die Pflichtliteratur und die Online-Vorlesungen.
- **Säule 2:** In der Pflichtliteratur, die zu jedem Thema angegeben ist, erhalten Sie vertiefte Einblicke, insbesondere zu den mathematischen Details der Item-Response-Theorie sowie der LCA und LST sowie zu den statistischen Details der entscheidungstheoretischen Modelle. Bei der Pflichtliteratur handelt es sich zumeist um Kapitel aus bewährten Lehrbüchern, die größtenteils als eBooks über die Bibliothek der FernUniversität verfügbar sind.
- **Säule 3:** In *Online*-Vorlesungen werden die Inhalte des Kurses in kompakter Form erläutert und durch Beispiele veranschaulicht.
- **Säule 4:** In einer *Moodle*-Umgebung zu diesem Kurs können Sie Fragen zu den einzelnen Themen des Kurses stellen und gemeinsam diskutieren. Dabei werden Sie von der Modulleitung unterstützt.

In der Präsenzveranstaltung zum Modul 2 wird die Erstellung von Gutachten (vgl. Kurs 3448) sowie damit verbundene Kompetenzen im Mittelpunkt stehen. Die Gutachtenerstellung wird auch in der Hausarbeit im Mittelpunkt stehen.

Die Betreuung für diesen Kurs wird schwerpunktmäßig über die virtuelle Lehr-Lernumgebung Moodle realisiert. Wir empfehlen Ihnen dringend sich in Moodle anzumelden und die bereitgestellten Materialien zu diesem Kurs zur Kenntnis zu nehmen. Sie finden den Zugang zu Moodle unter:

<http://moodle-psy.fernuni-hagen.de>

Die Lernumgebung zu diesem und zu den anderen Kursen des zweiten M.Sc-Moduls werden im Sommersemester in der Regel jeweils Mitte April und im Wintersemester jeweils Mitte Oktober freigeschaltet.

In diesem Kurs wird aus Gründen der besseren Lesbarkeit manchmal auf die explizite Nennung beider Geschlechter verzichtet. In solchen Fällen sind aber immer beide Geschlechter gemeint.

Lehr- und Lernziele

Der erste Teil des Kurses vertieft Wissen und Kompetenzen zur Testkonstruktion und -anwendung. Im Einzelnen sollen Sie...

- das bereits im B.Sc. Psychologie erworbene Wissen zur Konstruktion und Anwendung von Tests vor dem Hintergrund der KTT festigen.
- die konzeptuellen und mathematischen Prinzipien einiger probabilistischer Testmodelle verstehen, insbesondere das dichotom-logistische Rasch-Modell, aber auch die Latent-Class-Analyse.
- die Vor- und Nachteile probabilistischer Testmodelle gegenüber der KTT diskutieren und kritisch beurteilen können.
- wissen, dass probabilistische Testmodelle Voraussetzungen für adaptives Testen schaffen, Prinzipien des adaptiven Testens sowie einige adaptive diagnostische Verfahren kennen.
- die Bedeutsamkeit der Latent-State-Trait-Theorie (LST) als Erweiterung der KTT sowie deren Relevanz bei der Konstruktion von Verfahren zur Erfassung von *States* und *Traits* einschätzen können.
- wissen, wie man diagnostische Verfahren mit UniPark im Internet administrieren kann und welche ethischen und methodischen Aspekte damit verbunden sind.

Der zweite Teil des Kurses vermittelt Wissen und Kompetenzen zur psychologisch-diagnostischen Urteilsbildung, zum diagnostischen Prozess und zu entscheidungstheoretischen Modellen. Im Einzelnen sollen Sie...

- die Grenzen der Personbeurteilung im Alltag und Methoden zur Analyse vorwissenschaftlicher Urteilsbildung kennen.
- Moderatoren der Urteilsbildung im Alltag sowie damit verbundene empirische Befunde kennen und deren Relevanz für die eigene psychologisch-diagnostische Tätigkeit einschätzen können.
- vor dem Hintergrund aktueller Befunde und Methoden kritisch reflektieren können, unter welchen Bedingungen die klinische Urteilsbildung und unter welchen Bedingungen die statistische Urteilsbildung sinnvoll eingesetzt werden kann.
- die wesentlichen Phasen und Merkmale verschiedener Modelle des diagnostischen Prozesses kennen.
- Nutzenerwägungen im diagnostischen Prozess einschätzen können.
- Wissen zu den Parametern diagnostischer Entscheidungen sowie zu verschiedenen Entscheidungsstrategien erworben haben.
- das Zusammenwirken verschiedener Determinanten der Entscheidungsgüte kennen.
- wissen, wie Entscheidungen mit Hilfe von ROC-Kurven optimiert werden können.

Ich wünsche Ihnen viel Erfolg bei der Erarbeitung des Kurses und viel Spaß beim Studieren!

Karl-Heinz Renner

[Überarbeitet durch das Lehrgebiet Persönlichkeits-, Rechtspsychologie und Diagnostik]