

Kurs 03421

Bernd Marcus

Grundlagen der Testkonstruktion

Kultur- und Sozialwissenschaften

Inhaltsverzeichnis

GRUNDLAGEN DER TESTKONSTRUKTION

Bernd Marcus & Markus Bühner

Einführung in den Kurs.....	5
1 Grundlagen psychologischer Testverfahren.....	8
1.1 Anwendungsbereiche und Klassifikation von Testverfahren	10
1.2 Tests und Testtheorie	14
1.3 Grundlagen der Testtheorie.....	16
1.3.1 Einführung in die Klassische Testtheorie (KTT).....	17
1.3.2 Kritische Anmerkungen zur Klassischen Testtheorie.....	22
1.3.3 Grundideen der Probabilistischen Testtheorie (PTT)	24
2 Testkonstruktion: Von der Planung zum ersten Entwurf.....	27
2.1 Merkmalsbereich und Geltungsbereich	27
2.2 Festlegung des generellen Konstruktionsprinzips.....	30
2.3 Generierung eines Itempools	38
2.4 Sprachliche Formulierungsregeln und Antwortformate	42
3 Itemanalyse: Kürzung und Revision des Entwurfs	58
3.1 Itemanalyse nach Klassischer Testtheorie (KTT).....	58
3.1.1 Rekodierung, Deskriptive Itemanalyse und Itemschwierigkeiten	59
3.1.2 Trennschärfeanalyse.....	68
3.2 Überblick über die Probabilistische Testtheorie	79
3.2.1 Grundlagen des Rasch-Modells	80
3.2.2 Weitere probabilistische Testmodelle für dichotome Daten.....	85
3.2.3 Probabilistische Modelle für Ratingskalen	89
4 Skalenbildung: Faktorenanalyse, Aggregation von Items und Normierung.....	95
4.1 Explorative Faktorenanalyse (EFA).....	95
4.1.1 Grundlagen der Faktorenanalyse	95
4.1.2 Varianten und Vorgehensweise bei der Explorativen Faktorenanalyse	98
4.1.3 Beispiel für die Durchführung einer EFA mit SPSS.....	113
4.1.4 Bewertung der Explorativen Faktorenanalyse	127
4.2 Konfirmatorische Faktorenanalyse (CFA).....	129
4.2.1 Grundbegriffe Linearer Strukturgleichungsmodelle (SEM).....	130
4.2.2 Vorgehensweise bei einer Konfirmatorischen Faktorenanalyse..	135
4.3 Aggregation zu Skalenwerten, Normierung und Interpretation.....	153
4.4 Zusammenfassung der Schritte der Testkonstruktion.....	161
5 Evaluation psychologischer Testverfahren	165
5.1 Richtlinien und Beurteilungssysteme für Tests	165

5.2	Psychometrische Gütekriterien.....	168
5.2.1	Objektivität und Reliabilität	169
5.2.2	Validität	180
5.3	Weitere allgemeine Gütekriterien.....	197
5.3.1	Kriterien mit vorwiegend praktischem Anwendungsbezug	198
5.3.2	Kriterien mit Bezug zu Rechten und Reaktionen der Teilnehmer	199
5.3.3	Kriterien mit Bezug zum Verhalten der Teilnehmer	203
5.4	Gütekriterien für die Einzelfalldiagnostik	205
5.4.1	Anforderungen an die Normierung.....	206
5.4.2	Kennwerte mit Bezug zur Messgenauigkeit.....	208
	Literaturverzeichnis	217
	Stichwortverzeichnis	231

Einführung in den Kurs

„Was Ihre Frisur über Sie verrät“. Eine Online-Frauenzeitschrift beantwortet diese Frage, indem einfach aus sechs verschiedenen Frisuren ausgewählt werden kann. Der Erstautor dieses Studienbriefs hat diesen „psychologischen Test“ nach bestem Wissen unter Auswahl der subjektiv ähnlichsten Frisur (Stirn frei) bearbeitet, mit folgendem Ergebnis: „Sie begegnen dem Leben auf direkte Weise und Sie haben keine Angst, direkt auf Ihr Ziel zuzugehen. Sie sind offen und ehrlich und haben dynamische Qualitäten. Durch Ihre aktive Energie kann es Ihnen manchmal an Feinheit fehlen, denn Sie kümmern sich nicht unbedingt immer um die Details. Sie gehen schnell zum Wesentlichen über. Sie lieben den Kontakt mit Ihren Mitmenschen und Sie sind optimistisch. Ihre Frisur zeigt, dass Sie der Welt und dem Leben gegenüber offen sind.“ Unter Zuhilfenahme eines Kammes wäre es ihm nach diesem Test auf recht einfache Weise gelungen, seine Persönlichkeit grundlegend zu verändern (Mittelscheitel): „Sie möchten in einer harmonischen und ausgeglichenen Atmosphäre leben. Mit einer gewissen Strenge versuchen Sie, in allen Dingen gerecht zu sein. Sie sind geradeheraus und Sie streben nach einer einfachen, direkten und soliden Existenz und ebensolchen Beziehungen. Sie wünschen sich eine moralische Ordnung, die Ihnen entspricht. Sie brauchen geistreiche Kontakte und ein Leben, das reibungslos und ohne zu viel Aufregung verläuft. Ihre Frisur offenbart Ihren leidenschaftlichen und energischen Charakter.“ (alle Zitate nach Gofeminin.de, n.d.)

Wenn Sie es sich nicht ohnehin schon dachten, sollte Ihnen nach der Lektüre dieses Studienbriefs klar werden, dass es sich bei „Was Ihre Frisur über Sie verrät“ *nicht* um einen psychologischen Test handelt. Zwischen dem, was sich eine Redakteurin vermutlich bei einer Tasse Kaffee ausgedacht hat und dem, was Psychologinnen in jahrelanger intensiver Arbeit entwickeln, bestehen bestenfalls oberflächliche Ähnlichkeiten. Die Konstruktion psychologischer Tests nach psychometrischen Prinzipien gilt als eine der bedeutendsten Leistungen der wissenschaftlichen Psychologie und als ein wesentlicher Beitrag der Verhaltenswissenschaften zur Gesellschaft (z.B. American Educational Research Association, American Psychological Association & National Council on Measurement in Education, 1999). Die vertiefte Kenntnis dieser Prinzipien gehört zu den Qualifikationen, welche die Psychologie normalerweise von allen anderen Disziplinen unterscheidet und ist damit auch ein Alleinstellungsmerkmal für Psychologen im Wettbewerb auf dem Arbeitsmarkt. In diese Kenntnisse soll der vorliegende Studienbrief einführen. Allerdings werden Tests keineswegs nur von Psychologen angewandt (vgl. Kap. 1.1).

Die Konstruktion psychologischer Tests gilt als eine der bedeutendsten Leistungen der wissenschaftlichen Psychologie.

Psychologische Testkonstruktion ist ein im Wesentlichen methodisches Teilgebiet der Psychologie, das eine bestimmte statistisch fundierte Sozialtechnologie bereitstellt. Viele Aspekte werden in der Fachliteratur auf einem mathematisch recht anspruchsvollen Niveau diskutiert. Natürlich kann eine Einführung hier nur die Grundprinzipien vermitteln. Vorausgesetzt werden Kenntnisse zu statistischen

Psychologische Testkonstruktion als statistisch fundierte Sozialtechnologie

Grundlagen wie z.B. Skalenniveaus oder Korrelationskoeffizienten, die in den Modulen 1 und 2 des Studiengangs vermittelt wurden. Wir haben uns bemüht, hier vor allem das grundsätzliche Verständnis des Vorgehens bei der Testkonstruktion in den Vordergrund zu rücken und mathematisch-technische Details wie z.B. Formeln nur insoweit darzustellen, wie es die Nachvollziehbarkeit des Gesagten erfordert. Auf eine vertiefende Zusatzlektüre kann jedoch in vielen Fällen nicht verzichtet werden. Am Ende der einzelnen Abschnitte des vorliegenden Textes werden dazu jeweils Literaturhinweise gegeben.

Der Studienbrief beruht in geringen Teilen auf dem Lehrbuch „Einführung in die Test- und Fragebogenkonstruktion“ von Markus Bühner in der Auflage von 2006 (aktuell Bühner, 2011). Einige kurze Abschnitte wurden dabei in gekürzter und überarbeiteter Fassung für den vorliegenden Text übernommen, ohne dies, um der besseren Lesbarkeit willen, im Detail durch Zitate kenntlich zu machen. Das Lehrbuch geht jedoch vielfach inhaltlich erheblich über diesen Studienbrief hinaus und wird dementsprechend auch als vertiefende Lektüre empfohlen. Weitere vertiefende und ergänzende Elemente können, wie Sie es aus anderen Modulen gewohnt sind, in Form von Online-Vorlesungen, Texten und Aufgaben in der Lernumgebung Moodle abgerufen und bearbeitet werden.

Struktur des Studienbriefs

Dieser Studienbrief unterscheidet sich außerdem von vielen Lehrtexten zur Testkonstruktion dadurch, dass der Aufbau sich weniger an den statistischen Grundlagen orientiert als an dem idealtypischen Ablauf einer Testkonstruktion. Nach einer generellen Einführung in Kapitel 1 werden in Kapitel 2 zunächst Möglichkeiten besprochen, nach welchen grundsätzlichen Konzepten Tests und ihre einzelnen Elemente (Skalen, Aufgaben bzw. Items) generiert werden können. Das dritte Kapitel beschäftigt sich mit empirischen Informationen, die genutzt werden können, um die in einem so entstandenen Rohentwurf enthaltenen einzelnen Aufgaben oder Fragen zu beurteilen. Dazu zählen insbesondere Verfahren der Itemanalyse nach der klassischen und probabilistischen Testtheorie. Im folgenden vierten Kapitel geht es um die Zusammenfassung einzelner Aufgaben zu Skalen mittels verschiedener Varianten der Faktorenanalyse. Im fünften und letzten Hauptteil geht es um Anforderungen an eine anwendungsbezogene Dokumentation und Evaluation (psychometrische und anwendungsbezogene Gütekriterien) von Testverfahren. Dabei ist es gar nicht so erheblich, ob diese Kenntnisse genutzt werden, um ein eigenes Testverfahren zu entwickeln (was eher die Ausnahme sein dürfte) oder ob vorhandene Testverfahren für eine bestimmte Fragestellung bewertet und daraus die bestmögliche Auswahl getroffen werden soll. Auch für eine qualifizierte Bewertung ist es erforderlich, mit den Prinzipien der Testkonstruktion vertraut zu sein.

Im Einzelnen sollen in diesem Kurs die folgenden Ziele und Lernergebnisse erreicht werden. Sie sollen nach erfolgreicher Absolvierung des Kurses...

Ziele und Lernergebnisse

- das Spektrum psychologischer Testverfahren kennen und deren Relation zur zugrunde liegenden Testtheorie einschätzen können. Um übrigens ein mögliches Missverständnis zu vermeiden: Die Begriffe „Test(verfahren)“, „Testkonstruktion“, „Testtheorie“ etc. beziehen sich hier stets auf Tests im Sinne bestimmter psychologischer Messverfahren, nicht auf statistische Signifikanztests (vgl. Modul 2) und deren Theorie.
- mit den grundsätzlichen Möglichkeiten, psychologische Konstrukte mittels Tests zu operationalisieren und die Testitems in einer dem Anwendungszweck und den Teilnehmern angemessenen Weise zu formulieren vertraut sein.
- die wichtigsten Itemkennwerte berechnen und angemessen interpretieren können und in der Lage sein, diese Informationen für die Konstruktion der Endform von Tests und die Beurteilung vorhandener Testverfahren zu nutzen sowie dabei...
- mit den Voraussetzungen und Grenzen der nach der klassischen Testtheorie berechneten Kennwerte vertraut sein und die Möglichkeiten der probabilistischen Testtheorie einschätzen können.
- die wichtigsten Verfahren der exploratorischen Faktorenanalyse kennen und hinsichtlich ihres Stellenwerts für die Skalenbildung beurteilen sowie deren Nutzen vor dem Hintergrund eines Grundverständnisses der konfirmatorischen Faktorenanalyse einschätzen können.
- die Einflussfaktoren und die Möglichkeiten zur Bestimmung der Reliabilität und zur Validierung von Tests kennen und diese Gütekriterien in ihrer Bedeutung einschätzen können. Darüber hinaus besitzen Sie ein Verständnis für die Bedeutung anwendungsbezogener Kriterien und können diese in die Beurteilung der Güte von Testverfahren einbeziehen.
- in der Lage sein, anhand der in Testmanualen und aus anderen Quellen zugänglichen Informationen eine dem Anwendungszweck angemessene Auswahl aus vorhandenen Tests zu treffen.

Wir wünschen Ihnen bei der Erarbeitung des Kursmaterials viel Erfolg und viel Freude.

Bernd Marcus und Markus Bühner