

Ein neues Modell für PISA 2012? Ein Kommentar zu 'Der schiefe Turm von PISA...'

Rost, Jürgen

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Rost, J. (2006). Ein neues Modell für PISA 2012? Ein Kommentar zu 'Der schiefe Turm von PISA...'. *ZA-Information / Zentralarchiv für Empirische Sozialforschung*, 59, 50-52. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-198399>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Ein neues Modell für PISA 2012?

Ein Kommentar zu ‚Der schiefe Turm von PISA...‘

von Jürgen Rost¹

Die Autoren stellen ein neues Item-Response Modell vor, das ein Grundbedürfnis vieler Testpraktiker erfüllt, nämlich Schülern, die in einem Test eine bestimmte Anzahl schwerer Testaufgaben richtig bearbeiten, eine größere Fähigkeitsausprägung zuzuschreiben als Schülern, die dieselbe Anzahl leichter Aufgaben korrekt bearbeiten. Das ist beim *Rasch*-Modell nicht der Fall, jedenfalls dann nicht, wenn alle Schüler denselben Test bearbeiten. Hier, d.h. in einem Test, für den das *Rasch*-Modell gilt, erhalten alle Schüler mit derselben Anzahl gelöster Aufgaben denselben Fähigkeitsparameter, unabhängig davon, *welche* Aufgaben sie gelöst haben.

Das irritiert viele Testpraktiker, z.B. Lehrer, die gewohnt sind, für die Lösung schwerer Testaufgaben mehr Punkte zu vergeben als für die Lösung leichter Aufgaben und den Test mit der Gesamtpunktzahl zu benoten. In der Tat, das widerspricht dem *Rasch*-Modell. Die Notenverteilung wäre eine andere, wenn man die Testaufgaben mit 0 und 1 kodieren und mit dem *Rasch*-Modell auswerten würde.

Ein kleiner, aber bedeutsamer Unterschied zum hier vorgestellten trigonometrischen Modell besteht darin, dass der Lehrer die pro Aufgabe zu vergebende Punktzahl *vorher* festlegt, während die Personenparameter im vorgeschlagenen Modell allein anhand der Daten geschätzt werden. Aber dieser kleine Unterschied soll nicht darüber hinwegtäuschen, dass es beim *Rasch*-Modell tatsächlich nicht darauf ankommt, *welche* Items eine Person gelöst hat, sondern nur *wie viele*. In der theoretischen Ableitung des Modells spielt diese Annahme sogar eine zentrale Rolle, es ist die Annahme, dass die Anzahl der gelösten Items eine suffiziente Statistik für die Schätzung der entsprechenden Personen- oder Itemparameter darstellt. Wer diese Annahme für seinen Test als nicht gegeben oder als unplausibel ansieht, der sollte nicht das *Rasch*-Modell verwenden.

¹ Dr. *Jürgen Rost* ist Professor am IPN - Leibniz-Institut für die Pädagogik der Naturwissenschaften an der Universität Kiel, Olshausenstraße 62, 24098 Kiel, jrost@ipn.uni-kiel.de

Warum ist das *Rasch*-Modell denn so berühmt geworden, wenn es ‚noch nicht einmal‘ berücksichtigt, *welche* Items eine Person gelöst hat? Nun, mir liegt es näher, diese Annahme der suffizienten Statistik nicht als *Voraussetzung* anzusehen, aus der man das *Rasch*-Modell ableitet, sondern als *Folgerung*, die sich daraus ergibt, dass alle übrigen Annahmen des *Rasch*-Modells gültig sind. Diese sind:

- Die Annahme der Eindimensionalität, die besagt, dass die Antwortwahrscheinlichkeiten nur von einer Personenvariable, und zwar derselben für alle Items abhängen,
- Die lokale stochastische Unabhängigkeit, die besagt, dass es keine Abhängigkeiten zwischen den Antwortwahrscheinlichkeiten und den Items mehr gibt, wenn man die latente Variable konstant hält, und
- Die Annahme, dass sich die Items in ihrer Schwierigkeit unterscheiden, aber die itemcharakteristischen Funktionen sich nicht überschneiden.

Aus diesen Annahmen folgt, dass alle Antwortmuster mit demselben Summenscore unterschiedliche Wahrscheinlichkeiten haben beobachtet zu werden. In diese Patternwahrscheinlichkeiten gehen natürlich die Itemschwierigkeiten ein. Ein Pattern, in dem nur die beiden leichtesten Items gelöst wurden ist sehr viel wahrscheinlicher als das Pattern, in dem nur die beiden schwersten Items gelöst wurden. Das ist plausibel. Tritt nun der unwahrscheinliche Fall ein, dass eine Person nur die beiden schwersten Items gelöst hat, so muss dies als ein von dem Modell abweichendes Ereignis angesehen werden, das aufgrund der Probabilistik eintreten kann und auch in entsprechend geringem Umfang eintreten wird.

Aber es widerspricht doch jeder Intuition, dieser Person einen höheren Parameter einzuräumen, wo sie doch alle Items, die leichter sind, als die beiden gelösten, *nicht* gelöst hat. Mit den beiden gelösten (schweren) Aufgaben mag die Person ihre Kompetenz unter Beweis gestellt haben, mit den vielen ungelösten (leichten) Aufgaben hat sie aber gleichzeitig ihre Inkompetenz unter Beweis gestellt.

Wenn man einen Test will, bei dem man die Schwierigkeiten der von einer Person gelösten Items bei der Berechnung ihrer Fähigkeit berücksichtigen will, so sollte man nicht das *Rasch*-Modell verwenden, sollte seinen Test anders konstruieren als herkömmliche Leistungstests und sollte – vielleicht – das von den Autoren propagierte trigonometrische Modell anwenden. Schade ist es, dass die Autoren ihr neues Modell nur so zaghaft beschreiben. Mich hätte eine Grafik der Itemcharakteristiken interessiert, die Frage ob, die ICC's sich überschneiden, die allgemeinen Schätzgleichungen für die Modellparameter, Bedingungen der Identifizierbarkeit, die Erwartungstreue der Schätzer, Modellgeltungstests und – ja und eine Anwendung

auf einen echten Leistungstest inklusive eines empirischen Vergleiches mit dem *Rasch*-Modell.

So erfährt der Leser nur, dass so ziemlich alles am *Rasch*-Modell falsch ist, dass das *Rasch*-Modell nicht mehr spezifisch objektiv ist, wenn man den Begriff durch den neuen Begriff der Forminvarianz ersetzt, das das trigonometrische Modell das bessere *Rasch*-Modell ist und dass die PISA-Studien völlig inadäquat ausgewertet wurden. Alles Feststellungen, die man kaum glauben mag, die zu entkräften aber mehr Platz erfordert als der Beitrag selbst. Gewinnt das trigonometrische Modell nur dann an wissenschaftlicher Bedeutung, wenn zugleich das *Rasch*-Modell als unbrauchbar abqualifiziert wird? Ich glaube nicht. Konkurrenz belebt das Geschäft und wer weiß, vielleicht wenden wir bei der PISA-Studie 2012 schon das trigonometrische Modell an.

Antwort der Autoren auf den Kommentar von Jürgen Rost

von Klaus Harney, Christoph Fuhrmann und Hanns L. Harney

Unsere Kritik geht von der Beobachtung aus, dass das logistische *Rasch*-Modell uniforme Antwortmuster erlaubt, sie jedoch nicht interpretieren kann. *Jürgen Rost* hebt hervor, dass die Lösung – das trigonometrische *Rasch*-Modell – eine weitere interessante Eigenschaft hat: Der Parameter einer Person hängt von den Schwierigkeiten der von dieser Person gelösten Aufgaben ab. Im Rahmen des logistischen Modells werden die gelösten Aufgaben im Wesentlichen nur gezählt. (Das mittlere Schwierigkeitsniveau geht ein). In beiden Gestalten des *Rasch*-Modells gelten Kompetenzen als stark oder schwach ausgeprägt je nach einer Bewertung der gelösten Aufgaben. Die Ausprägung der Kompetenzen bildet der Logik des Modells zufolge eine jeweilige Welt der Erreichbarkeit von Aufgabenlösungen aus. Hinter der Schwierigkeit von Aufgaben stecken grundsätzlich gesehen Wirklichkeiten, die sich durch eine erfolgreiche Aufgabenbearbeitung erschließen bzw. im negativen Fall für Personen verschlossen bleiben. Umgekehrt stecken hinter der Kompetenz von Personen unterschiedliche Handlungs- und Erkenntnisniveaus für Wirklichkeiten, die ohne Wissen und Kompetenz nicht entstehen können.