

## **Das Difficulty plus Guessing PL Modell – Wie sehr halten Modelltests das nominelle Risiko 1. Art ein?**

*Klaus D. Kubinger, Kathrin Gruber, Christine Hohensinn & Clemens Draxler:*

In psychologischen Leistungstests werden häufig Multiple-Choice-Items eingesetzt, ohne besonders auf das damit zusammenhängende Problem der Ratemöglichkeit zu achten, und damit darauf, dass Testpersonen relativ hohe Scores erzielen können, ohne über ein entsprechendes Fähigkeitsausmaß zu verfügen. Auf einem früheren Klagenfurter Statistiktag wurde bereits ein IRT-Lösungsversuch dafür vorgestellt, nämlich die Anwendung des Difficulty-plus-Guessing PL Modells (Kubinger & Draxler, 2006); Anwendungen von Kubinger und Draxler (2006) sowie von Hohensinn und Kubinger (in Vorb.) haben gezeigt, dass mit diesem Modellansatz gegenüber dem Rasch-Modell etliche Items für die Kalibrierung eines Tests „gerettet“ werden können. Im vorliegenden Referat werden nun umfangreiche Simulationsergebnisse vorgestellt, mit denen geprüft werden sollte, ob der einschlägige goodness-of-fit-Test überhaupt das Risiko 1. Art einhält. Nebenbei wurde auch die Macht bzw. das Risiko 2. Art analysiert.