



STEVE AYAN ist Diplompsychologe und G&G-Redaktionsleiter.

Murks per Definition

In einem neuen Projekt wollen Forscher die Ergebnisse psychologischer Studien reproduzieren. Ein längst überfälliges Vorhaben!

Es passieren ja manchmal die unwahrscheinlichsten Dinge. Das gilt auch für Psychologen. Bei ihrer Arbeit funkt ihnen ständig der Zufall dazwischen, denn selbst wenn eine Theorie eigentlich Quatsch ist, kann ihr empirischer Test positiv ausfallen; und stimmt sie doch, zieht man eventuell trotzdem eine Niete. In der Psychologie dreht sich eben alles um Wahrscheinlichkeiten.

Als Kriterium für die Richtigkeit eines Studienresultats gilt die statistische Signifikanz, Forscherdeutsch für »Bedeutsamkeit«. Hierfür legen Psychologen (nach ziemlich komplizierten Formeln) meist eine Fehlertoleranz von fünf Prozent zu Grunde. Das heißt, in einem von 20 Fällen sitzt man einem Irrtum auf. Anders gesagt: Jedes 20. Resultat ist Murks – per Definition! Tatsächlich dürfte die Quote der falsch positiven Befunde in den Fachjournalen noch weit höher ausfallen. Denn Arbeiten, die keine Effekte ergeben, haben kaum Chancen, publiziert zu werden. Folglich stochern Forscher – getrieben vom Veröffentlichungsdruck – in ihren Daten herum, bis etwas Vorzeigbares herauskommt.

Wie ein Team um Joseph Simmons von der University of Pennsylvania Ende 2011 zeigte, lassen sich die absurdesten Annahmen überraschend leicht statistisch untermauern. Zum Beispiel: Wer dem Song »When I'm sixty-four« lauscht, wird jünger verglichen mit Leuten, die ein Kinderlied hören! Zum Beweis braucht man nur eine ausreichend kleine Stichprobe (bei der einzelne Ausreißer hart durchschlagen) sowie eine Fülle verschiedener Variablen, die man auf Signifikanz testet. Wie gesagt, wer lange genug herumrechnet, wird früher oder später fündig.

Nun hat eine Gruppe von Forschern um Brian Nosek von der University of Virginia ein ehrgeiziges Projekt gestartet. Sie wollen alle im Jahr 2008 in den drei renommierten Fachjournalen »Psychological Science«, »Journal of Personality and Social Psychology« sowie »Journal of Experimental Psychology: Learning, Memory, and Cognition« erschienene Studien replizieren. Lässt sich ein Ergebnis nicht wiederholen, beweist dies zwar noch längst nicht, dass es falsch ist – nur Vorsicht scheint geboten. Das Problem: Noch nie wurde annähernd so systematisch untersucht, wie viele

Studien einen »Re-Test« bestehen. Das so genannte Reproducibility Project macht die (längst überfällige!) Probe aufs Exempel.

Das schlägt in Fachkreisen überraschend hohe Wogen. Manche Kollegen sehen in Nosek und seinen Mitstreitern Nestbeschmutzer, die dem Ansehen des Fachs schaden. Andere erwidern, es lasse tief blicken, dass Psychologen so empfindlich darauf reagieren, wenn man ihre Studien unter die Lupe nimmt – schließlich sei Reproduzierbarkeit eine Kernforderung jeder Wissenschaft. Oder komme es den Seelenkundlern gar nicht so sehr darauf an, dass ihre Ergebnisse wahr sind, solange sie nur gut klingen?

Fest steht: Auch wenn eine Studie nach allen Regeln der Experimentierkunst durchgeführt wird, kommt trotzdem womöglich Unsinn dabei heraus (siehe G&G 7-8/2011, S. 43). Jedes Ergebnis ist interpretationsbedürftig – und eines allein sagt wenig aus. Wissenschaft ist immer Teamwork. Wer sich also wie Nosek anschickt, den Kollegen auf die Finger zu schauen, tut genau das, wofür Forscher da sind: kritisch nachzuhaken, statt in Ehrfurcht zu erstarren.

WEBLINK

Das Reproducibility Project im Internet:
<http://openscienceframework.org>

QUELLE

Simmons, J.P. et al.: False-Positive Psychology. In: Psychological Science 22, S. 1359–1366, 2011

MEHR ZUM THEMA

Die Fallstricke der Seelenkunde – von der Alltagspsychologie bis zur Laborforschung – erkundet G&G-Redaktionsleiter Steve Ayan in seinem gerade erschienenen Buch »Hilfe, wir machen uns verrückt!« (Pendo, München 2012, 256 S., € 19,99).

