

Site: **Wiki of Science** at <http://wikiofscience.wikidot.com>
 Source page: **20120426 - Misinterpretation of 'p' (2000) (2e)** at
<http://wikiofscience.wikidot.com/print:20120426-no-significance-haller2000>

20120426 - Misinterpretation of 'p' (2000) (2e)

[Data] [[<Normal page](#)] [**PEREZGONZALEZ Jose D [ed] (2012)**. [Misinterpretation of 'p' \(2000\) \(2e\)](#)⁵. Journal of Knowledge Advancement & Integration ([ISSN 1177-4576](#)), 2012, pages 146-148.]

Misinterpretations of 'p' and 'sig'

Haller and Krauss (2000²) carried out a study on common misinterpretations of tests of significance among German psychology students and academics, which partly replicates one done by Oakes (1986³). Typically, most of these misinterpretations confuse [p-values](#) (ie, the probability of the data when assuming that the null hypothesis is true) and, especially, [statistical significance](#)⁷, with the probability of proving or disproving hypotheses (be this the null hypothesis or an alternative hypothesis). Another misinterpretation is the so-called "replication fallacy", which occurs when the probability of the data is assumed to represent the probability of finding similar results if the research were to be repeated.

Haller and Krauss found that most participants held at least one misinterpretation out of the six presented (see table 1). They also found that, overall, 100% of psychology students held one or more misinterpretations (mean=2.5), almost 90% of psychology researchers also held one or more misinterpretations (mean=2), and 80% of instructors of statistics in psychology also held one or more misinterpretations (mean=1.9). The authors thought worrisome the high percentage of instructors with misinterpretations, as these may pass those misinterpretations down to students. Another interesting result, one not highlighted by the authors, though, is the high percentage of researchers (including instructors when carrying out and publishing research) with misinterpretations, as these would perpetuate those when publishing, peer-reviewing others' publications, and making research-informed decisions (such as chairing committees, granting funding, etc).

[Fold](#)

Table of Contents

[Misinterpretations of 'p' and 'sig' Methods](#)

[Research approach](#)

[Sample](#)

[Materials](#)

[Analysis](#)

[Generalization potential](#)

Table 1. Percentages of misinterpretations regarding tests of significance

Common misinterpretations ⁶	Stat. instructors	Researchers	Students
Significance disproves the null hypothesis	10%	15%	34%
The p-value informs of the probability of the null hypothesis	17%	26%	32%
Significance proves the alternative hypothesis	10%	13%	20%
The p-value informs of the probability of the alternative hypothesis	33%	33%	59%
'P' informs of the probability of a wrong decision when rejecting the null	73%	67%	68%
The p-value informs of the probability of the results if replicated	37%	49%	41%
<i>(Participants who answered that all of above were false)</i>	20%	10%	0%

Methods

Research approach

Replication study using a German sample. The original study had been carried out by Oakes (1986³) with a British sample of psychology academics 15 years earlier.

Sample

A convenient sample of 113 participants from departments of psychology in 6 German universities. 44 participants were psychology students, 39 participants were research psychologists not involved with teaching statistics, and 30 participants were instructors of statistics in psychology (including lecturers and tutors).

Materials

Oakes's (1986³) questionnaire translated into German:

- The questionnaire consisted of a small scenario and six statements. The scenario described a small research with two-independent samples, and provided the relevant results: a t-test with 'p=0.01'.
- The six statements asked for a true / false decision regarding whether each particular statement reflected a logical interpretation of the results. Unbeknownst to the participants, all statements were false, representing six common misinterpretations of tests of significance.
- The study also provided the 'hint' that "several or none of the statements may be correct".

Analysis

Descriptive statistics.

Generalization potential

This particular research was done with a sample of psychology academics and students from different universities in Germany, and its design appears to be more valid than that of previous studies. It also found similar trends than did Oakes, 1986³, in the U.K., and Falk and Greenbaum, 1995¹, in Israel. Thus, these results may be generalizable to the following populations (in order of decreasing generalization power):

- German, British and Israeli psychology academics and researchers (including students).
- Psychology professionals trained in German, British and Israeli universities.
- Psychology professionals and academics elsewhere.
- Other scientists (especially from the social sciences, medicine and business) which rely on the [null hypothesis significance testing](#) (NHST) procedure.

References

1. **FALK Ruma & Charles W GREENBAUM (1995)**. *Significance tests die hard: the amazing persistence of a probabilistic misconception*. *Theory & Psychology*, 1995, volume 5, number 1, pages 75-98. DOI 10.1177/0959354395051004.
2. **HALLER Heiko & Stefan KRAUSS (2000)**. *Misinterpretations of significance. A problem students share with their teachers*. *Methods of Psychological Research Online*, 2002, volume 7, number 1, pages 1-20.
3. **OAKES Michael (1986)**. *Statistical inference: a commentary for the social and behavioral sciences*. John Wiley & Sons (Chichester, UK), 1986.
4. **PEREZGONZALEZ Jose D [ed] (2011)**. *Misinterpretation of 'p' (2000)*. *Journal of Knowledge Advancement & Integration (ISSN 1177-4576)*, 2011, pages 110-112.

+++ Notes +++

5. This second edition updates the original edition⁴ by reducing confusion between p-values and statistical significance (see [tests of significance](#)).
6. The original research statements have been rephrased here.
7. The example provided to participants used $p=0.01$, thus it can be interpreted as having the dual role of a '[p-value](#)' and a '[conventional level of significance](#)'.

Want to know more?

[Wiki of Science - Hypotheses testing \(disambiguation\)](#)

This Wiki of Science page lists alternative methods for testing the probability of data or hypotheses.

[Wiki of Science - Null hypothesis significance testing](#)

This Wiki of Science page reflects on the pseudoscientific bases of the null hypothesis significance testing (NHST) procedure typically used in the social sciences and medicine.

Wiki of Science - Related studies

You can find more information on two related studies in Wiki of Science. One was the original study done by [Oakes](#) in 1986; the other study was a replication done by [Falk and Greenbaum](#) in 1995.

Editor

Jose D PEREZGONZALEZ (2012). Massey University, Turitea Campus, Private Bag 11-222, Palmerston North 4442, New Zealand. ([JDPerezgonzalez](#) [JDPerezgonzalez](#)).



Other interesting sites



[Journal KAI](#)



[Wiki of Science](#)



[AviationKnowledge](#)



[A4art](#)



[The Balanced Nutrition Index](#)

page revision: 3, last edited: 26 Apr 2012, 10:43 GMT+12 (17 seconds ago)

Unless stated otherwise Content of this page is licensed under [Creative Commons Attribution-ShareAlike 3.0 License](#)