The Growth of Psychology and its Corrective Mechanisms:  
A Bibliometric Analysis (1950-2014)

Jelte M. Wicherts, Tilburg University

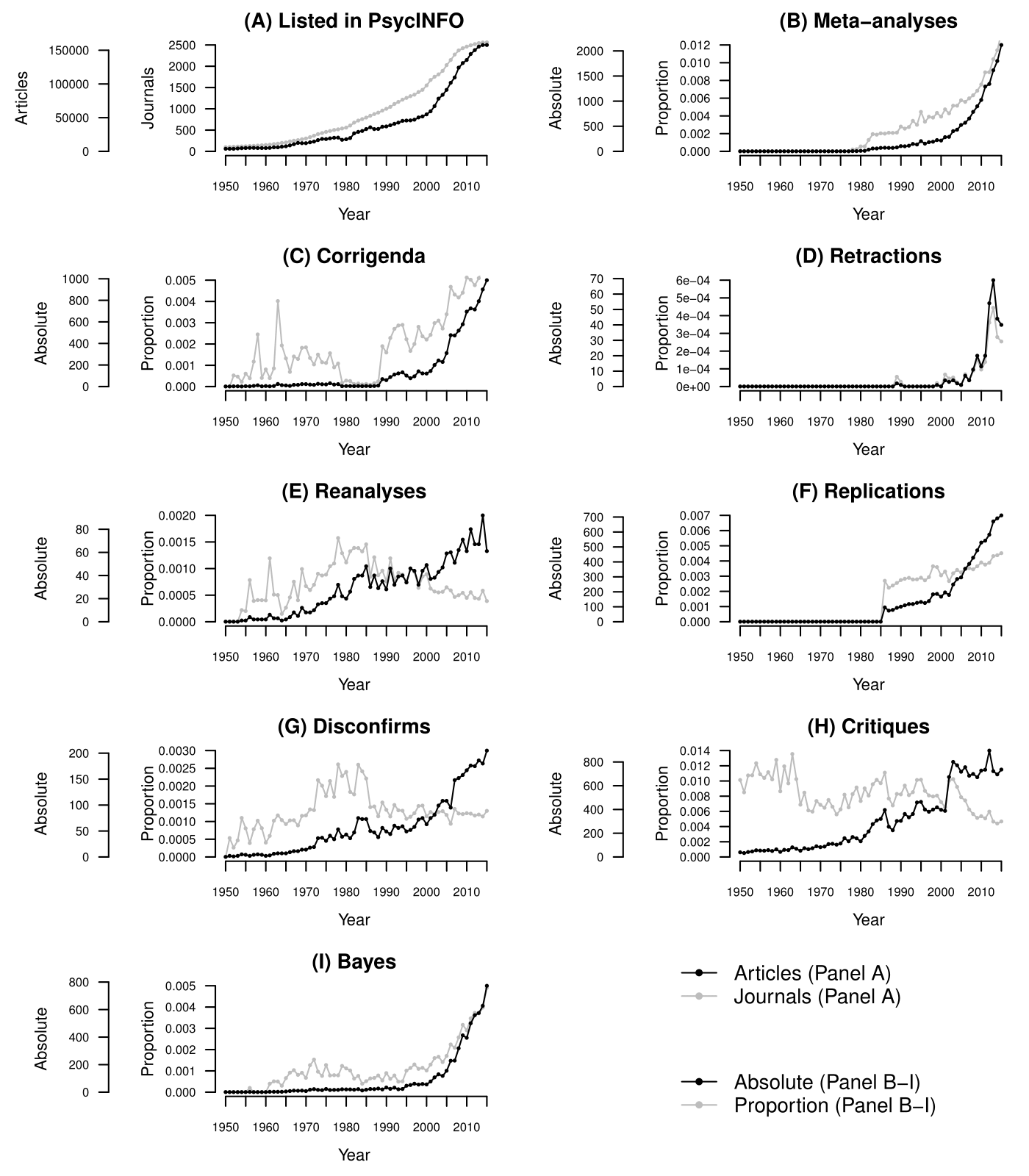
Chris H.J. Hartgerink, Tilburg University

Raoul P. P. P. Grasman, University of Amsterdam

A search in PsycINFO learns that the psychological literature anno 2014 covered 136,191 articles published in 2,557 peer-reviewed journals. Specific searches show that in the last year, 1,581 psychological articles used the term meta-analysis in the abstract, 750 were corrigenda/errata, 31 were retraction notices, and 580 involved comments or critiques. Moreover, 629 articles contained the term replication in their abstract, while the abstracts of 75 articles featured the term reanalysis. Another 159 articles appear to concern disconfirmations as they included “falsify/falsifies’”, “contradict(s)”, or “disconfirm(s)” in the abstract. “Bayes” was mentioned only 559 times, indicating that bayesian methods are still scarcely represented in the literature.

To put these numbers in historical perspective, we repeated these searches for the last 65 years. The resulting data shed light on the growth of the psychological literature and trends in corrective mechanisms therein. We used unambiguous search-strings (i.e., avoiding terms compounded by substance) related to common corrective mechanisms, in order to shed light on the appearance of relevant articles. Absolute and relative results per year are depicted in Figure 1 for total number of articles and journals, publications related to meta-analysis, critiques/comments, disconfirmations, errata/corrigenda, retractions, replications, and re-analyses. The data, search strings, and database coverage are given in the appendix.

*Figure 1.* Absolute yearly frequencies (black lines) and yearly proportions (grey lines) of articles in peer-reviewed journals in PsycINFO from 1950-2014 for meta-analyses (Panel B), errata/corrigenda (Panel C), comments/critiques (Panel D), reanalyses (Panel E), replications (Panel F), disconfirmations (Panel G), retractions (Panel H), false negatives (Panel I), false positives (Panel J), and Bayes (Panel K). Panel A shows the absolute number of journals (black) and articles (grey) listed in PsycINFO. Data available at osf.io/



These results indicate that despite an absolute rise of bibliometric indices that indicate critical assessment of research results, relative increases are not omnipresent. Whereas meta-analyses, corrigenda, retractions, and the mention of Bayes have clearly increased in the last 65 years, reanalyses, replications, disconfirmations, and critiques show no clear increasing trend. More specifically, reanalyses and critiques show a downward trend; disconfirmations have been relatively stable in recent decades; replications are slowly becoming more frequent again after drooping off prior to 1985.

Inspecting whether the search terms actually pertained to those mechanisms we aimed to inspect, 25 abstracts per search item were coded manually. Retractions were already manually checked (see Appendix), leaving 7 search terms and 175 abstracts. Of these 175 abstracts, the first two authors agreed in 148 cases, and of these 82% pertained to the mechanism of interest. This indicates that the searches are plausibly effective in determining the prevalence of certain corrective mechanisms in the literature.

The results highlight that despite the increasing amount of research, prevalence of corrective mechanisms is not improving across the board. Research is accumulated, corrected, and retracted more frequently, which may be due to growing scrutiny by readers in the online era and rising awareness of misconduct (Van Noorden, 2011)[. Comments of frequentist statistics (e.g.,](#_ENREF_3) Rouder, Speckman, Sun, & Morey, 2009; Wagenmakers, Wetzels, Borsboom, & van der Maas, 2011) seem to have increased the uptake of Bayesian methods, but this nevertheless remains a niche in the literature. The relative number of articles that concern comments, reanalyses of data, or disconfirmations did not increase in the last decades as compared to earlier periods, reflecting perhaps a growth in journals’ emphasis on publishing novel results. Moreover, the relative frequency for published reanalyses actually decreased in recent decades, which may also reflect the common failure to share data for reanalysis due to substandard documentation and archiving of data (Vanpaemel, Vermorgen, Deriemaecker, & Storms, 2015; Wicherts, Bakker, & Molenaar, 2011; Wicherts, Borsboom, Kats, & Molenaar, 2006)[. However, replications have been on the rise since approximately 2000; whether this pertains to conceptual or direct replications remains unclear. Even though initially it might have pertained more to conceptual replications, changes in recent years anecdotally seemed to have focused on direct replications, as indicated by projects such as Many Labs (Klein et al., 2014) and the Reproducibility Project (Open Science Collaboration, 2015). It seems that the recent replication crisis (Pashler & Wagenmakers, 2012) has affected the psychology literature only very little and other important corrective mechanisms (i.e., reanalyses, critiques, disconfirmations), have hardly changed. There remains room for improvement in the psychology literature, which stresses that further cultural change is required to improve the quality of psychology as a whole.](#_ENREF_5)

\*\*OLD\*\*

The number of articles (Panel A) has grown sharply from 3850 in 1950 to 136,191 in 2014. This aligns with the increasing number of journals listed in PsycINFO (line in Panel A), which went from 97 to 2,557. Increases in specific types of articles should be viewed relative to the total number of articles by year, so Panels B-H include both the raw counts (black lines) and the proportions (grey lines).

Panel B shows an increasing number of articles with “meta-analysis” in the abstract, starting with the development of meta-analysis in the late 1970s. The increase in meta-analytic articles is disproportionate in the sense that their relative occurrence went up from approximately 0.2% in the 1980s to 1.2% in recent years.

Panel C depicts the trend line and the number of articles with “erratum” or “corrigendum” in the title (we did not search with “correction” because of its substantive meaning). Results show a near-absence of errata/corrigenda in most of the 1980s and a disproportionate increase in (self-)corrections since 1987.

Papers may be retracted because of (honest) errors, failures to replicate, or scientific misconduct. Panel D highlights a sharp increase in the number and relative frequency of retracted papers in the last decade. The increase is disproportionate even if we exclude the year 2012, which features many retractions by Diederik Stapel. In psychology, the yearly retraction rate of published articles is currently around 0.02%, which aligns with the recent (increasing) rate documented across the whole of science (Van Noorden, 2011)[.](#_ENREF_3)

A core corrective mechanism of science concerns (independent) reanalyses of the data. Although the number of articles with the term reanalysis in the abstract (Panel E) has increased over the years, the relative frequency dropped steadily since 1985. This is striking given that sharing electronic data for reanalysis should have become easier in the last decades.

Another core corrective mechanism of science concerns replication (Asendorpf et al., 2013)[. Panel F depicts the number and rate of articles that contained the word “replication” in their abstract. Only few such (apparent) failures to replicate get published](#_ENREF_1) (cf. Makel, Plucker, & Hegarty, 2012)[. The rate of these articles was around 0.37% in the period 1950-1982 and slightly lower at around 0.32% since 1983.](#_ENREF_2)

Panel G includes the number of articles with “disconfirm(s)”, “falsify/falsifies’”, or “contradict(s)” in the abstract. Despite the growing count per year, the relative appearance of articles that appear to fit Popper’s falsificationist philosophy was on the rise throughout 1950-1980 to approximately .25%, to subsequently decrease and stabilize around .1% in recent years.

Panel H reports the number of articles with “critique” or “comment” in the title. Despite the rising numbers of such critical articles in PsycINFO, their relative appearance shows a slight decrease over the years.

Additional searches for statistical decision errors, be they false negative (Panel I) or false positive (Panel J) decision errors, increase in absolute numbers but show no substantial changes throughout the years. However, in recent years the false positive debate has been more present in the psychological literature, which is accompanied by a minor rise in relative mentions of false negatives, rising to just below .1% of the literature.

Panel K shows the number of abstracts mentioning anything related to Bayes (i.e., “bayes\*”), with an absolute and relative increase. This could indicate that Bayesian methods are becoming more widely spread and possibly also its application. It could also be that there are more mentions of Bayes because of the false positive debate. Note that the mentions only take up approximately .4% and this increase should not be exaggerated.

**References**

Asendorpf, J. B., Conner, M., Fruyt, F. D., Houwer, J. D., Denissen, J. J. A., Fiedler, K., et al. (2013). Recommendations for increasing replicability in psychology. *European Journal of Personality, 27*, 108-119. doi: 10.1002/per.1919

Klein, R. A., Ratliff, K. A., Vianello, M., Jr., R. B. A., Bahník, Š., Bernstein, M. J., … Nosek, B. A. (2014). Investigating variation in replicability. *Social Psychology*, *45*(3), 142–152.

Makel, M. C., Plucker, J. A., & Hegarty, B. (2012). Replications in Psychology Research How Often Do They Really Occur? *Perspectives on Psychological Science, 7*, 537-542. doi: 10.1177/1745691612460688

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251). http://doi.org/10.1126/science.aac4716

Pashler, H., & Wagenmakers, E. (2012). Editors’ introduction to the special section on replicability in psychological science: A crisis of confidence? *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *7*(6), 528–530.

Vanpaemel, W., Vermorgen, M., Deriemaecker, L., & Storms, G. (2015). Are we wasting a good crisis? The availability of psychological research data after the storm. *Collabra Open Access Journal*. http://doi.org/10.1525/collabra.13

Van Noorden, R. (2011). Science publishing: The trouble with retractions. *Nature, 478*, 26-28. doi: 10.1038/478026a

Wicherts, J. M., Bakker, M., & Molenaar, D. (2011). Willingness to share research data is related to the strength of the evidence and the quality of reporting of statistical results. *PLoS ONE, 6*, e26828. doi: 10.1371/journal.pone.0026828

Wicherts, J. M., Borsboom, D., Kats, J., & Molenaar, D. (2006). The poor availability of psychological research data for reanalysis. *American Psychologist, 61*, 726-728. doi: 10.1037/0003-066X.61.7.726

**Appendix**

Table A1. Search terms used in PsycINFO

|  |  |
| --- | --- |
| Category | Boolean/Phrase |
| articles | PY 1950 AND PT Peer Reviewed Journal |
| meta-analyses | PY 1950 AND PT Peer Reviewed Journal AND AB meta-analysis |
| corrigenda | PY 1950 AND PT Peer Reviewed Journal AND (TI corrigendum or TI erratum) |
| critiques | PY 1950 AND PT Peer Reviewed Journal AND (TI critique OR TI comment) |
| reanalyses | PY 1950 AND PT Peer Reviewed Journal AND AB reanalysis |
| replications | PY 1950 AND PT Peer Reviewed Journal AND AB replication |
| disconfirms | PY 1950 AND PT Peer Reviewed Journal AND (AB falsifies OR AB contradicts OR AB disconfirms) |
| retractions\* | PY 1950 AND PT Peer Reviewed Journal AND TI retraction |
| bayes | PY 1950 AND PT Peer Reviewed Journal AND AB bayes\* |

Note: Searches conducted on May 6, 2015; \*For retractions, we manually inspected search results to avoid inclusion of irrelevant papers