

Peer Community in Registered Reports: Stage 1 Snapshot

1. **Provisional title.** Psychology: The science of behavior?
2. **Authors and affiliations.** Sabrina F. Norwood¹, Malte Elson¹, A. Solomon Kurz^{2, 3}, Ian Hussey¹
¹University of Bern, ²VISN 17 Center of Excellence for Research and Returning War Veterans, ³Central Texas Veterans Health Care System
3. **Field and keywords.** meta-science, measurement
4. **Research question(s) and/or theory.** Psychology's self-image is that it is "the study of the mind and behavior" (APA Dictionary of Psychology). Provocatively, Baumeister et al. (2007) called this self-image into question by pointing out that we measure actual behavior with increasing rarity. However, the invention of smartphones, wearable devices, smart home devices, and other digital technologies have greatly expanded our potential to collect real-world behavioral data (Harari et al., 2016). This raises the question of whether the field of psychological science has capitalized on this potential in the intervening 15 years. This study seeks to estimate the prevalence and trends in the use of multiple different predefined classes of measurement in psychology. Equally, this study goes beyond Baumeister et al. (2007) by a) quantifying the use of other types of psychological measurement beyond behavior, and b) classifying whether each article's research question was relevant to overt behavior, and therefore the eventual measurement of actual behavior may be relevant for that line of work.
5. **Hypotheses (where applicable).** This study takes an estimation rather than hypothesis testing approach. It will address the following questions: For each of the different defined types of measurement, what is the prevalence in the psychology literature in the period from 2009 to 2024? How does this vary between subfields and journals? What is the trend in their use over time (i.e., increasing or decreasing) during this period of technological change? More specifically, has the downward trend in the use of behavioral measures up until 2007 described by Baumeister et al. (2007) continued?
6. **Study design and methods.** Articles reporting the results of quantitative original research will be randomly sampled from 5 journals in each of 6 subfields of psychology (clinical, cognitive, developmental, industrial and organizational, social and personality, and general) in the last 15 years (2009-2024). Journals will be selected based on a combination of their citation metrics and community feedback on representativeness. 15 articles will be sampled from each journal based on the availability of resources, for a total of 450 articles. While this proportion of sampling is small at the journal level, partial pooling in the Bayesian models should provide meaningful estimates of prevalence and trend at the field and subfield level. Two raters will code types of psychological measurement each study employed, using our already-validated definitions and coding scheme: 1) direct behavioral measures, 2) behavioral proxy measures, 3) neuro/bio measures, 4) self-reports of subjective states, 5) self-reports of behavior, and 6) generative tasks.
7. **Key analyses that will test the hypotheses and/or answer the research question(s).** The counts of each type of measure will be employed as the DV in Bayesian negative-binomial multilevel models, using year, subfield, and journal as (nested) random effects, and using weakly informative priors. Posterior distributions (95% HDIs) will be used to estimate (a) absolute prevalence (across years) and (b) trends (between years). Depending on convergence, nonlinear trends over time may also be estimated. This will be done 1) at the overall field level, 2) the subfield level, and 3) the journal level. The (non)overlap of 95% HDIs will be used to determine differences between journals, fields, and types of measurement.
8. **Conclusions that will be drawn given different results.** Results will be used to quantify the degree to which psychological science's research questions relate to behavior, and how this varies between subfields and journals. Having established the degree to which the field states an aspiration to make claims about behavior, results will then speak to the degree to which our actual measurement practices align with these aspirations. These conclusions will be necessarily descriptive as no point-hypotheses regarding prevalences are being tested. Specific hypotheses will be tested regarding the trends in the use of different types of measurement over time at the field/subfield/journals level (i.e., evidence of upward or downward trends).
9. **Key references.** 1. APA Dictionary of Psychology (Retrieved April 2, 2024) <https://dictionary.apa.org/psychology> 2. Baumeister et al. (2007) <https://doi.org/10.1111/j.1745-6916.2007.00051.x> 3. Harari et al. (2016). <https://doi.org/10.1177/1745691616650285>