

All for one or some for all? Bayesian evaluation of multiple N=1 hypotheses

Klaassen, F.^{1*}, Zedelius, C. M.², Veling, H.³, Aarts, H.⁴ and Hoijtink, H.^{1,5}

¹ Department of Methodology and Statistics, Utrecht University, The Netherlands

² Department of Psychology, University of Santa Barbara, USA

³ Behavioural Science Institute, Radboud Universiteit Nijmegen, The Netherlands

⁴ Department of Psychology, Utrecht University, The Netherlands

⁵ Cito Institute for Educational Testing, Arnhem, The Netherlands

* Presenting author, PhD student, supervisor: H. Hoijtink

Suggested talk duration (15-60 minutes)

20 (25-30 including discussion?)

Summary (max. 500 words)

Analyses are mostly executed at the population level, whereas in many applications the interest is on the individual level instead of the population level. This research project considers multiple N=1 experiments, where participants perform multiple trials with a dichotomous outcome in various conditions. Expectations with respect to the performance of participants can be translated into so-called informative hypotheses. These hypotheses can be evaluated for each participant separately using Bayes factors. A Bayes factor expresses the relative evidence for two hypotheses based on the data of one individual. We propose to "average" these individual Bayes factors in the gP-BF, the average relative evidence. The gP-BF can be used to determine whether one hypothesis is preferred over another for all individuals under investigation. This measure provides insight into whether the relative preference of a hypothesis from a pre-defined set is homogeneous over individuals. Two additional measures are proposed to support the interpretation of the gP-BF: the Evidence Rate (ER), the proportion of individual Bayes factors that support the same hypothesis as the gP-BF, and the Stability Rate (SR), the proportion of individual Bayes factors that express a stronger support than the gP-BF. These three statistics can be used to determine the relative support in the data for the informative hypotheses entertained. Software is available that can be used to execute the approach proposed and to determine the sensitivity of the outcomes with respect to the number of participants and within condition replications.

Relevance to conference theme

Synthesis of knowledge from a small number of within-person experiments.

Keywords (max. 3)

Within-subject experiment, Bayes factor, Synthesis.