Journal of Statistical Software

MMMMMM YYYY, Volume VV, Issue II. doi: 10.18637/jss.v000.i00

Vazul: An R Package for Analysis Blinding

Tamás Nagy
ELTE Eotvos Lorend University

Alexandra Sarafoglou
University of Amsterdam

Abstract

The abstract of the article.

Keywords: keywords, not capitalized, Java.

1. Introduction

In data analysis, researchers face the challenge of maintaining objectivity and minimizing bias. A key issue in this context is the considerable analytic flexibility available at various stages of the process, including data preprocessing, variable selection, model specification, and statistical testing. This flexibility has been described as the "garden of forking paths'' (Gelman and Loken 2013), as a multitude of plausible analytic strategies can be applied to the same dataset. Depending on the complexity of the analysis, this can result in hundreds or thousands of plausible, non-redundant analytic pipelines (Steegen, Tuerlinckx, Gelman, and Vanpaemel 2016; Simonsohn, Nelson, and Simmons 2020). Importantly, different analytic paths can yield different results, and analysts may, intentionally or unintentionally, gravitate toward those that align with their expectations. Such flexibility can therefore lead to biased outcomes if researchers make analytic choices that are influenced by their hypotheses or prior beliefs about the data.

To ensure the reliability of results and the validity of conclusions drawn from empirical data, it is crucial to prevent the idiosyncrasies of the data from feeding back into the formulation of the hypotheses being tested (HARKing, Kerr 1998) or from prompting researchers to exploit their analytic freedom to accentuate desired outcomes (Simmons, Nelson, and Simonsohn 2011).

Analysis blinding (MacCoun and Perlmutter 2015) is a methodological approach designed to protect the confirmatory status of an analysis by concealing crucial test-relevant aspects of the data from the analysts—for example, by scrambling dependent variables or masking key labels. The procedure typically involves two independent parties: a data manager, who has full access to the raw data and is responsible for implementing the blinding steps, and an analyst, who is tasked with developing the analysis plan. After the data manager applies the blinding procedure, the analyst receives the resulting modified dataset: the blinded data.

Working with the blinded data, the analyst can explore and preprocess the dataset and develop an analysis plan without being influenced by whether a particular analytic choice produces the desired result, as any patterns in the blinded data are, by design, only a product of chance. Importantly, although the test-relevant elements are concealed, other meaningful characteristics remain intact, including demographic information, secondary variables, and the distributional properties of both dependent and independent variables. As a result, the analyst retains all information necessary to tailor the analytic approach to the specific, and potentially unexpected, features of the data.

Once the analysis plan has been finalized using the blinded data, the analyst is granted access to the raw, unblinded dataset, and the confirmatory analysis can then be carried out strictly according to the pre-established plan.

Analysis blinding ties in with other methodologies to minimize bias in the research cycle, such as single-blind or double-blind experimental designs, where participants and/or experimenters are unaware of certain aspects of the study to prevent bias in data collection (REF). Here, the analysts is intentionally kept unaware of certain aspects of the data to prevent bias in data analysis, which is why the methodology is also referred to as triple-blind (REF).

The **vazul** package in R `(https://CRAN.R-project.org/package=vazul` provides the tools needed for data managers to implement blinding procedures. The package includes functionality for data scrambling, data masking, and subset-based analyses. In the remainder of this article, we introduce the methodology in more detail, describe the functionalities of the **vazul** package, and illustrate its application through practical examples. We conclude with a brief discussion and outline future directions.

1.1. Analysis Blinding: Safeguarding Against Bias Without Preregistration

Originally, analysis blinding gained traction in astrophysics in the early 2000s (REF) as a means of safeguarding analysts against bias. Since then, the methodology has also been advocated in the social and behavioral sciences (Dutilh, Sarafoglou, and Wagenmakers 2019; Nagy, Hergert, Elsherif, Wallrich, Schmidt, Waltzer, Payne, Gjoneska, Seetahul, Wang, Scharfenberg, Tyson, Yang, Skvortsova, Alarie, Graves, Sotola, Moreau, and Rubínová 2025; MacCoun 2021; MacCoun and Perlmutter 2015; MacCoun and Perlmutter 2018), where preregistration of analysis plans, that is, requiring researchers to specify their planned analyses before data collection begins or before they view the data, has been the predominant approach for limiting analytic flexibility (e.g., existing estimates of preregistration rates in empirical psychology published in prominent journals range from about 14% Hardwicke, Thibault, Clarke, Moodie, Crüwell, Schiavone, Handcock, Nghiem, Mody, Eerola, and Vazire (2024) to about 40% Pfadt, Bartoš, Godmann, Waaijers, Groot, Heo, Mensink, Nak, de Ruiter, Sarafoglou, Siepe, Arena, Akrong, Aust, van den Bergh, Brenner, Doekemeijer, Donzallaz, van Doorn, Ormeño Echevarria, Finneman, Geller, Hato, Koskinen, Krijgsman, Kulbe, Lüken, Marsman, Ott, Pawel, Piestrak, de Ron, Sekulovski, Serry, Stefanów, Stevenson, Sadowski, Sopuch, Vasileiou, Visser, Völler, Wiechert, de Wit, Wuth, and Wagenmakers (2025)). In the literature, analysis blinding has been proposed as an alternative or complementary method to preregistration (e.g., Dutilh *et al.* 2019) as it addresses several of preregistration's major shortcomings. Specifically, preregistration can restrict analysts so rigidly that they are unable to adapt their analyses to unexpected peculiarities in the data. Although deviations from preregistered plans are both possible and accepted when transparently disclosed in the final manuscript, such deviations run counter to the spirit of preregistration, as they reintroduce data-dependent decisions that may bias the results. At the same time, text-based preregistrations, that is, the description of analysis plans in a preregistration template, often fail to be 'specific, precise, and exhaustive' (Wicherts, Veldkamp, Augusteijn, Bakker, van Aert, and van Assen 2016, 2), which leaves substantial degrees of freedom unaccounted for.

In contrast to preregistration, analysis blinding allows researchers to maintain flexibility in their analysis plans, as they can explore the data without being constrained by a rigid preregistered protocol. This flexibility is especially important in analyses involving advanced statistical modeling or preprocessing, where not all decisions can be anticipated in advance and researchers must make data-dependent choices. For instance, in longitudinal network analysis, statistical models typically assume stationarity–an assumption which is frequently violated in psychological data and one that must be detected and addressed through detrending techniques (Epskamp, Waldorp, Mõttus, and Borsboom 2018; Burger, Hoekstra, Mansueto, and Epskamp 2022; Hoekstra, Huth, Sekulovski, Delhalle, and Sarafoglou 2025). The flexibility gained by analysis blinding enables researchers to develop analysis strategies that are appropriately tailored to the specific characteristics of the dataset without the need to anticipate and specify all eventualities in advance. Consequently, researchers who analyze their data with analysis blinding are less likely to deviate from their developed analysis plan compared to researchers who preregister their analytic strategy in a text-based format (Sarafoglou, Hoogeveen, and Wagenmakers 2023).

Moreover, similar in spirit to approaches proposed in code-based preregistration (Peikert, Van Lissa, and Brandmaier 2021; Van Lissa, Brandmaier, Brinkman, Lamprecht, Peikert, Struiksma, and Vreede 2021), analysis blinding naturally produces analysis plans that are specific, precise, and exhaustive, because they are written directly in code rather than described verbally in preregistration templates. As with preregistration, analysis scripts based on the blinded data can be uploaded or attached to preregistration platforms (e.g., the Open Science Framework) to provide a transparent record of the planned analysis before the data are unblinded. In short, analysis blinding can serve as a valuable alternative or additional safeguard to preregistered analysis plans: it effectively protects against bias while providing analysts with the flexibility needed to develop an analysis strategy that is optimally suited to their data.

References

Burger J, Hoekstra RHA, Mansueto AC, Epskamp S (2022). "Network estimation from time series and panel data." In *Network psychometrics with R*, pp. 169–192. Routledge.

Dutilh G, Sarafoglou A, Wagenmakers EJ (2019). "Flexible Yet Fair: Blinding Analyses in Experimental Psychology." *Synthese*, **198**, S5745–S5772. doi:10.1007/s11229-019-02456-7.

Epskamp S, Waldorp LJ, Mõttus R, Borsboom D (2018). "The Gaussian graphical model in cross-sectional and time-series data." *Multivariate behavioral research*, **53**(4), 453–480. doi:10.1080/00273171.2018.1454823.

Gelman A, Loken E (2013). "The garden of forking paths: Why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time." *Department of Statistics, Columbia University*, **348**. URL http://www.stat.columbia.edu/~gelman/research/unpublished/p_hacking.pdf.

Hardwicke TE, Thibault RT, Clarke B, Moodie N, Crüwell S, Schiavone SR, Handcock SA, Nghiem KA, Mody F, Eerola T, Vazire S (2024). "Prevalence of Transparent Research Practices in Psychology: A Cross-Sectional Study of Empirical Articles Published in 2022." *Advances in Methods and Practices in Psychological Science*, **7**(4), 25152459241283477. ISSN 2515-2459, 2515-2467. doi:10.1177/25152459241283477.

Hoekstra RHA, Huth KBS, Sekulovski N, Delhalle M, Sarafoglou A (2025). "Safeguarding Against Bias Without Preregistration: A Tutorial on Analysis Blinding for Network Analysis." *PsyArXiv*. doi:10.31234/osf.io/gruw5_v1.

Kerr N (1998). "HARKing: Hypothesizing After the Results are Known." *Personality and Social Psychology Review*, **2**, 196–217. doi:10.1207/s15327957pspr0203_4.

MacCoun R (2021). "Blinding to remove biases in science and society." In R Hertwig, C Engel (eds.), *Deliberate ignorance: Choosing not to know*, pp. 51–64. MIT Press, Cambridge.

MacCoun R, Perlmutter S (2015). "Blind analysis: Hide results to seek the truth." *Nature*, **526**, 187–190. doi:10.1038/526187a.

MacCoun R, Perlmutter S (2018). "Blind Analysis as a Correction for Confirmatory Bias in Physics and in Psychology." In SO Lilienfeld, I Waldman (eds.), *Psychological Science Under Scrutiny: Recent Challenges and Proposed Solutions*, pp. 297–322. John Wiley and Sons.

Nagy T, Hergert J, Elsherif MM, Wallrich L, Schmidt K, Waltzer T, Payne JW, Gjoneska B, Seetahul Y, Wang YA, Scharfenberg D, Tyson G, Yang YF, Skvortsova A, Alarie S, Graves K, Sotola LK, Moreau D, Rubínová E (2025). "Bestiary of Questionable Research Practices in Psychology." *Advances in Methods and Practices in Psychological Science*, **8**(3), 25152459251348431. ISSN 2515-2459, 2515-2467. doi:10.1177/25152459251348431.

Peikert A, Van Lissa CJ, Brandmaier AM (2021). "Reproducible research in R: A tutorial on how to do the same thing more than once." *Psych*, **3**(4), 836–867. doi:10.3390/psych3040053.

Pfadt MJ, Bartoš F, Godmann RH, Waaijers M, Groot L, Heo I, Mensink L, Nak J, de Ruiter PJ, Sarafoglou A, Siepe B, Arena G, Akrong E, Aust F, van den Bergh D, Brenner W, Doekemeijer R, Donzallaz CM, van Doorn J, Ormeño Echevarria N, Finneman A, Geller G, Hato T, Koskinen E, Krijgsman B, Kulbe L, Lüken M, Marsman M, Ott LV, Pawel S, Piestrak O, de Ron J, Sekulovski N, Serry M, Stefanów A, Stevenson N, Sadowski B, Sopuch M, Vasileiou A, Visser I, Völler M, Wiechert S, de Wit K, Wuth J, Wagenmakers EJ (2025). "Methodological Metamorphosis: The Rapid Rise of Bayesian Inference and Open Science Practices in Psychology." *PsyArXiv*. doi:10.31234/osf.io/ck3js_v1.

Sarafoglou A, Hoogeveen S, Wagenmakers EJ (2023). "Comparing analysis blinding with preregistration in the Many-Analysts Religion Project." *Advances in Methods and Practices in Psychological Science*, **6**(1), 25152459221128319. doi:10.1177/25152459221128319.

Simmons J, Nelson L, Simonsohn U (2011). "False–positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant." *Psychological Science*, **22**, 1359–1366. doi:10.1177/0956797611417632.

Simonsohn U, Nelson LD, Simmons JP (2020). "Specification curve analysis." *Nature Human Behaviour*, **4**, 1208–1214. doi:10.1038/s41562-020-0912-z.

Steegen S, Tuerlinckx F, Gelman A, Vanpaemel W (2016). "Increasing transparency through a multiverse analysis." *Perspectives on Psychological Science*, **11**, 702–712. doi:10.1177/1745691616658637.

Van Lissa CJ, Brandmaier AM, Brinkman L, Lamprecht AL, Peikert A, Struiksma ME, Vreede BM (2021). "WORCS: A workflow for open reproducible code in science." *Data Science*, **4**(1), 29–49. doi:10.3233/DS-210031.

Wicherts JM, Veldkamp CLS, Augusteijn HEM, Bakker M, van Aert RCM, van Assen MALM (2016). "Degrees of Freedom in Planning, Running, Analyzing, and Reporting Psychological Studies: A Checklist to Avoid p-Hacking." *Frontiers in Psychology*, **7**, 1832. doi:10.3389/fpsyg.2016.01832.

Affiliation:

Tamás Nagy
ELTE Eotvos Lorend University
Institute of Psychology,
ELTE Eotvos Lorend University,
Budapest, Hungary
E-mail: nagy.tamas@ppk.elte.hu

Alexandra Sarafoglou
University of Amsterdam
Department of Psychology,
University of Amsterdam,
Amsterdam, The Netherlands
E-mail: a.s.g.sarafoglou@uva.nl

Journal of Statistical Software http://www.jstatsoft.org/
published by the Foundation for Open Access Statistics http://www.foastat.org/
MMMMMM YYYY, Volume VV, Issue II *Submitted:* yyyy-mm-dd
doi:10.18637/jss.v000.i00 *Accepted:* yyyy-mm-dd