



Methodology and Research Practice

Most Psychological Researchers Assume Their Samples Are Ergodic: Evidence From a Year of Articles in Three Major Journals

Craig P. Speelman¹^a, Laura Parker², Benjamin J. Rapley², Marek McGann³

¹ Experimental Psychology Unit, Edith Cowan University, Joondalup, WA, Australia, ² School of Arts & Humanities, Edith Cowan University, Joondalup, WA, Australia, ³ Mary Immaculate College, University of Limerick, Limerick, Ireland

Keywords: Research methodology, ergodicity, replication

<https://doi.org/10.1525/collabra.92888>

Collabra: Psychology

Vol. 10, Issue 1, 2024

Conventional statistics methods in most psychological research, such as null-hypothesis significance tests (NHSTs), use aggregated values (i.e., the sample means) of group behaviours to make inferences about individuals. Such inferences are possibly erroneous because groups of humans rarely, if ever, constitute an ergodic system. To assume ergodicity without checking is to commit the ‘ergodic fallacy’. The aim of the current study was to examine the prevalence of this error in contemporary psychological research. We analysed three highly cited ‘Q1’ journals in the fields of clinical, educational and cognitive psychology for statements that indicated this error. As hypothesized, the ergodic fallacy was found in the vast majority of the papers investigated here. We also hypothesised that the prevalence of this error would be highest in cognitive psychology papers because this field typically assesses theoretical claims about universal cognitive mechanisms, whereas clinical and educational psychology are more concerned with empirically supported interventions. This hypothesis was also supported by our results. Nonetheless, the prevalence of the ergodic fallacy was still high in all fields. Implications are discussed with respect to the reporting of research findings and the validity of theories in psychology.

Introduction

The most frequently used approach of researchers conducting quantitative psychology studies is to collect data from groups of people, and to aggregate their results to estimate some population parameter(s), with the aim of comparing behaviour under different conditions, or exploring associations between different measurements of the same people. One problem with this approach is examined in this paper – that a group-based outcome is used to characterize some feature of the individuals in the group and/or to extrapolate to people like these individuals in the population.

Speelman and McGann (2020) observed that it is common practice for psychology researchers to conclude far more from their results than is justified by the statistical techniques used to analyse their data. As illustration of a common form of phrasing, take the following: “...people integrate ensemble information about the group average expression when they make judgments of individual faces’ expressions” (Griffiths et al., 2018, p. 311, Abstract). The problem with statements such as this is that the aggregation of group data only allows conclusions about group results, such as the average group performance, and there-

fore generalisations to average population performance. It is not possible to unscramble the metaphorical omelette and draw conclusions about the individual components (the people) of that population. We know this because of work on ergodicity by the work of mathematician George Birkhoff, whose ergodic theorem posits that for a pooled statistic (e.g., the mean) derived from a group to be legitimately used to describe an individual of that group, two conditions must be met – first, the individuals must be so similar that they are virtually interchangeable, and second, the individuals’ characteristics are temporally stable (i.e., do not change over time) (Molenaar, 2008). These conditions are met in situations such as a vessel containing only hydrogen atoms. However, the psychological phenomena and processes of most interest to psychology researchers are by their very nature non-uniform between individuals, and variable over time both within individuals and between them (Molenaar & Campbell, 2009). Therefore, any results derived from averaging measures of multiple individuals’ behaviours, cognitions, or emotional states (i.e., the sample mean), for example, does not accurately describe any one of those individuals at one time point, and cannot account

a Corresponding author: c.speelman@ecu.edu.au

for changes in those variables for any individual over time (Molenaar & Campbell, 2009).

So, in the example cited above, when the word “people” is used to characterize the effect observed in the experiment, the reader does not know what this means. Does it mean everyone in the experiment exhibited the effect, or some sizeable majority of the group (most?), or perhaps just more than half? The observation of a significant effect, as determined by some form of statistical analysis (e.g., a null hypothesis significance test (NHST) or a Bayesian equivalent, effect sizes, confidence intervals), only provides information about some difference at the group statistic level (Danziger, 1994). Information about the relative performance of individuals in the groups is lost once group-based statistics are calculated. Thus, unless some form of pervasiveness analysis (Speelman & McGann, 2020) is performed, it is difficult to know how to interpret the meaning of “people” in this context. As a result, it is overselling the experiment’s outcome to say that a feature of ‘people’ beyond the experiment has been observed.

An important implication of not being aware of the ergodicity assumption was highlighted by Speelman and McGann (2020). They contrasted four different data sets, with identical means, slightly different standard deviations, and similar inferential statistical effects. Despite these similarities, the underlying distributions and pervasiveness of effects amongst the individuals within each set were wildly different. A blanket statement such as “Individuals in one condition improved more than individuals in the other condition” would dramatically misrepresent the situation in each case, and yet a reliance on just the result of the inferential statistical test would have typically produced an interpretation such as this statement. Thus, not only would the underlying results be misrepresented by relying on the group results only, there are serious implications for applying such research outcomes to inform applied practice (e.g., clinical psychology) where the focus is on outcomes for individuals. Data from group experiments with aggregated statistics do not provide information that is useful for predicting how individuals will respond in such situations.

Speelman and McGann (2020) used the term “ergodic fallacy” to label the behaviour of researchers who assume the individuals in their research groups are ergodic when there is little evidence that individuals are ever ergodic (see also Fisher et al., 2018; Molenaar, 2013). Speelman and McGann suggested that this behaviour was “common” but provided no information as to how prevalent it is. The aim of the present study was to determine the prevalence of the ergodic fallacy amongst psychology researchers across a demonstrative sample of published empirical work when they report the findings of their research.

A secondary aim of the study was to assess the relative prevalence of this behaviour in three broad fields of psychology as represented by major journals in the area: cognitive psychology, educational psychology, and clinical psychology. The rationale for this comparison was that the aims of the three fields are arguably different in terms of their focus on theoretical versus applied outcomes. For in-

stance, the aim of research in cognitive psychology is often to assess theoretical arguments about universal cognitive mechanisms. As a result, it might be expected that researchers in cognitive psychology, with an interest in general or universal mechanisms underlying behaviours, would exhibit a tendency to overestimate the pervasiveness of observed average effects in the individuals of their sample. Researchers would thus be likely to draw erroneous conclusions about the results and implications of their studies. In contrast, in the field of educational psychology, where the aim is to provide information about, for instance, effective educational techniques, there may be more of a focus on the practical outcomes of such techniques for individuals. For this reason, it might be expected that a greater proportion of research in this area would better reflect individual variance and the ergodic fallacy would not be as prevalent. Finally, in clinical psychology, where the findings are often added to the evidence base for empirically-supported treatments (ESTs) and, by extension, applied to individuals, it might be expected that more researchers would be sensitive to the ergodic issue and less likely to draw erroneous conclusions. In summary, the tendency to assume ergodicity in experimental samples is likely to be lower in educational and clinical psychology compared to cognitive psychology.

Method

Approval for this study was granted by the Edith Cowan University Human Research Ethics Committee, application reference number: 2021-02358-PARKER.

Corpus

All papers in the 2020 editions of the *Journal of Experimental Psychology: Learning, Memory and Cognition (JEP:LMC)*, the *Journal of Educational Psychology (JEdP)*, and the *Journal of Consulting and Clinical Psychology (JCCP)* comprised the sample of psychology research papers. These journals were chosen because they are part of the same stable of journals (i.e., produced by the American Psychological Association), and so could be expected to reflect similar editorial policies and standards. All of these journals are considered to be in the top ranks of quality in their fields, as indicated by high *impact factors* (JEP:LMC: 3.140; JEdP: 6.856; JCCP: 7.156) and being ranked as Q1 by SciMago (<https://www.scimagojr.com>). At the least, papers published in these journals were considered to be representative of the types of papers published in the three fields under consideration here, and so behaviours in those papers should be reflective of research behaviour in those fields. It is worth noting, though, that there has been a greater emphasis in recent years on the perils of the ergodic fallacy in the research methodology literature (e.g., Fisher et al., 2018; Molenaar, 2013; Speelman & McGann, 2020). As a result, the snapshot of researcher behaviour provided by this sampling of the 2020 issues of these journals could well be supplanted by a different picture as more researchers become aware of the problems associated with

Table 1. Tallies of papers exhibiting particular behaviours with respect to the Ergodic Fallacy.

Journal	N	n	Ergodic Fallacy	Ergodic Awareness	Ambiguous	Aware
JCCP	89	77	60 (77.9%)	6	11	17 (22.1%)
JEdP	100	84	75 (89.3%)	2	7	9 (10.7%)
JEP:LMC	137	135	126 (93.3%)	4	5	9 (6.7%)
Total	326	296	261 (88.2%)	12	23	35 (11.8%)

Notes: JCCP = Journal of Consulting and Clinical Psychology; JEdP = Journal of Educational Psychology; JEP:LMC = Journal of Experimental Psychology: Learning, Memory and Cognition

the ergodicity assumption. Thus researchers themselves may not constitute an ergodic system as they could change in this respect over time.¹

Papers were only included for analysis if they described studies that reported new data or new analyses on previously reported data. Literature reviews and meta-analyses were not included in this analysis because, by their nature, details about individual behaviour are glossed over. The total number of papers published in 2020 in each journal, and the number of papers that were included in the analysis, are reported in [Table 1](#).

Procedure

Three of the four authors (CS, LP & BR) independently assessed each paper. Discrepancies were resolved by discussion amongst the assessors. The final assessments reflect a consensus position.

The primary focus of assessment was whether a research paper exhibited the ergodic fallacy. This was assessed by considering conclusions reported in the Abstract and Discussion sections of each paper and determining whether they exhibited the ergodic fallacy. That is, did the conclusions imply that the finding that was derived from group data applied also to the individuals in the group, and was not just a result that applied only at the group level. Such behaviour could include specific reference to the individuals observed in the study (e.g., “The current study found a significant difference in the improvement in symptoms for individuals in the ‘new treatment’ condition compared to participants in the ‘treatment as usual’ condition.”), or an implication that the effect observed in the study applies to individuals beyond the study (e.g., “The current study demonstrates that people exposed to the treatment investigated here could exhibit some improvement in their symptoms.”). Each paper was assessed as exhibiting one of three outcomes: 1. Ergodic Fallacy behaviour; 2. Ergodic Awareness (Non-ergodic Fallacy) behaviour (i.e., some awareness that a group result may not accurately reflect the results for individuals); 3. Ambiguous (i.e., typically some recognition of the ergodic issue, but conclusions were ultimately consistent with the ergodic fallacy).

A range of other features of each paper was recorded, including the statistical analyses used and how results were

presented (i.e., were frequency tables or measures of pervasiveness reported?). This data was used to validate decisions as to whether researchers were reasoning according to the ergodic fallacy. For example, if the main statistical analysis compared two means, and frequency tables revealed information about individual behaviours, which the researchers commented upon, then this could result in an assessment of “Ambiguous” because there was some sensitivity to the fact that aggregating results obscures the underlying individual behaviour patterns. In such a study, if each individual contributed a data point in two conditions, and the researchers considered the pervasiveness of an effect by counting the number of people where the difference between their two responses were in a particular direction (i.e., a pervasiveness analysis as described by Speelman & McGann, 2020), then this would be assessed as “Ergodic Awareness”.

Results

Tables summarising details of each paper with respect to the variables of interest in this study are provided as Supplementary Material (Tables S1 (JEP:LMC), S2 (JEdP) & S3 (JCCP)) and also available at osf.io/v8p6b. Summary data from all papers is presented in [Table 1](#). The first column (N) displays the total number of papers published in each journal in 2020. The second column (n) displays the number of papers that reported new data or new analyses of previously reported data (i.e., after the exclusion of papers reporting literature reviews, meta-analyses, qualitative studies or historical accounts).

The first result of interest in [Table 1](#) is that the majority of papers (88.2%) exhibited behaviour consistent with the ergodic fallacy. The remainder of papers were assessed as showing some awareness of the ergodicity issue (4.1%) or ambiguous (7.8%). This is one of those rare times in psychology where the outcome is so clear no form of statistical analysis is required to establish its veracity.

The second question we were interested in was whether the field of research had any impact on the likelihood that researchers would exhibit behaviour consistent with the ergodic fallacy. To address this question, we conducted a chi squared analysis of the contingency table in [Table 1](#). The small number of papers classified as Ergodic Awareness cre-

¹ We thank Caspar van Lissa for this observation.

ated problems for the analysis due to small expected frequencies in some cells. Given that papers in the Ergodic Awareness and Ambiguous columns all exhibited some awareness of the ergodicity issue, we combined the frequencies in these two columns into one to create a new column, labelled Aware. A chi-squared analysis of the 3 x 2 contingency data in the Ergodic Fallacy and Aware columns of [Table 1](#) revealed a significant association between journal and behaviour ($\chi^2(2) = 11.308, p = .004$, Cramer's $V = .195, p = .004$). Inspection of [Table 1](#) reveals that the likelihood of Ergodic Fallacy behaviour being the most commonly exhibited by a paper was greatest when the paper was published in JEP:LMC, less likely in JEdP, and least likely in JCCP.

Discussion

The main purpose of this study was to evaluate whether psychology researchers use the results of group-based statistical analyses to make claims about the individuals in their samples or the populations to which they belong, hence committing the ergodic fallacy. On this question the results of our analysis are clear. The vast majority of the 296 papers assessed from the 2020 volumes of three APA journals exhibited this type of behaviour. For example, one article discussed how pupil changes can predict levels of cognitive engagement and this was moderated by participant attention (Hutchinson et al., 2020). The researchers concluded that "More important, trial type effects in pupil diameter emerged only when participants reported being "on-task," but disappeared during periods of mind wandering. These results demonstrate that changes in pupil diameter reflect the degree of preparatory control exerted for an upcoming trial but only when attention is actively focused on the upcoming task" (p. 280). There is no reference in the discussion as to how many individuals in the sample exhibited something like the average result in their performance. Thus, it is implied that pupil dilation is the mechanism through which cognitive engagement can be measured for 'participants'. Certainly no qualification to 'participants' was provided. This illustrates the typical behaviour of most researchers in the 2020 edition of the journals assessed here. That is, they use between-person designs, aggregated data and statistical inference to reach conclusions about the population, and then on the basis of these conclusions, and without any indication of individual performance, they also imply conclusions about the individuals in their sample. If the aim is to define and explain mechanisms at an individual level, then researchers need to measure individuals throughout different times and under many contexts (Molenaar, 2013; Molenaar & Campbell, 2009). As demonstrated by the current study, the reporting behaviour of psychology researchers in the articles examined here does not respect this necessity, and the methods do not allow the type of conclusions researchers are apparently wanting to make (as evidenced by the phrasing used in Results and Discussion sections of research reports).

The second aim of this study was to explore differences in the likelihood of exhibiting the ergodic fallacy that might be associated with research area. To this end we assessed

three different journals that publish papers in three distinct research areas of psychology: cognitive psychology, educational psychology and clinical psychology. Our analyses found that the majority of papers in the 2020 volumes of each of the three journals exhibited the ergodic fallacy, but the prevalence of this behaviour did vary by research area. That is, the ergodic fallacy was more prevalent in JEP:LMC, less prevalent in JEdP, and the least prevalent in JCCP. This difference possibly reflects the different aims typical of research in the three broad areas of psychological research. The study highlighted above (Hutchinson et al., 2020) is a typical example of the way the research in JEP:LMC was reported. In the JEP:LMC papers, most experiments were designed as though the individuals in the sample each had a common cognitive mechanism that underlies the behaviour in focus, and the group-based analyses of aggregated data reinforced this impression. Relying on such forms of analysis makes it difficult to discover that individuals may behave differently, let alone consider the possibility that their behaviours result from different cognitive mechanisms.

It might be expected that in the field of educational psychology researchers would assume that people are not the same and change overtime, and so research in this field would be designed to reflect this view. Certainly some studies reported in the 2020 volume of JEdP were consistent with this view. For example, Cervone et al. (2020) clearly demonstrated awareness of the ergodic issue by recognising that individuals within a sample cannot be assumed to be similar enough to justify aggregation:

From the perspective of a measure of overall academic self-efficacy, the five students are the same; they obtained identical academic self-efficacy scores. However, Figure 3 reveals how they differed. As shown, some students felt confident in speaking up in class but not in discussing goals with an academic advisor, whereas others were confident at discussing goals but not in speaking up in class. Some were not confident in approaching professors. Others lacked self-efficacy for getting support from family members. In general, despite their identical general academic self-efficacy scores, the students were quite diverse; their contextualized self-efficacy responses filled most of the two-dimensional space of the profile graph. (p.1606)

Such intraindividual variability, which is integral to personality functioning..., can be understood if researchers are willing to invest effort in multicomponent assessments that are sensitive to the potentially idiosyncratic aspects of students' beliefs and life contexts. (p.1611)

A similar expectation about awareness of the ergodic issue is reasonable in the field of clinical psychology, where evidence for the effectiveness of treatments for psychological issues is ultimately of relevance to the treatment of individuals. Some studies in the 2020 volume of JCCP were sensitive to this issue. For example, Woods et al. (2020) captured the problem well with the following conclusion:

Clinical researchers are shifting emphasis from studying heterogeneous clinical syndromes to identifying and investigating trans-diagnostic features of psy-

chopathology ... While this is promising, the reigning nomothetic paradigm of psychopathology research often prioritizes interindividual differences over intraindividual processes. Here we argue that psychopathology is best understood as contextualized dynamic processes, and these manifest within an individual in a complex system over time and circumstances. The current study adopted this perspective in studying dynamic affective and interpersonal processes in social situations in a sample with a range of pathology and in a subsample whose members all shared a BPD diagnosis. We showed that the structure of each individual's processes was unique with some evidence for shared processes across individuals, particularly within the BPD group. (p.251)

Despite the different prevalence rates of the ergodic fallacy in the different research areas, the propensity to exhibit this behaviour was still high in all three research areas. The field of psychology is not so much the problem as the statistical assumptions researchers use and how this contributes to the ergodic fallacy. Molenaar and Campbell (2009) explain that psychological processes are person-specific and therefore analysis should be based on intraindividual variation to account for non-ergodicity in people. While researchers continue to use traditional methods, they cannot apply their results to the individuals in their samples no matter how thoroughly they discuss individual differences in their introductions and discussions (Rose et al., 2013). One potential solution to the ergodicity problem is to undertake pervasiveness analyses. Speelman and McGann (2020) point to a simple pervasiveness measure where aggregated data can be interpreted alongside the proportion of individuals who demonstrate the effect being studied. Many of the papers assessed in this study could have provided a more precise theoretical argument had they reported the number of people that demonstrated the effect being reported. Indeed this method has been used to assess a recent replication effort (Moore et al., 2023). Grice (2015) argues for an analysis of patterns which are visual in nature first then computational. Similar to the recommendations of Speelman and McGann (2020), a visual inspection of the raw data in a scatterplot allows researchers to make inferences about the direction of their effects and whether their summative statistics match the individuals in the sample. We cannot assume that the researchers whose papers we examined here did not consider their data in this way, but at least it was reported rarely. However, by providing this data graphically in an article, researchers can demonstrate, with evidence that they have found an effect present in all or most individuals. Finally, Rose, Rouhani and Fischer (2013) have recommended an analyse-then-aggregate approach. They suggest starting with the individual patterns and focusing on individual variability first before focusing on the aggregated data. This focus provides insight into how individuals vary systematically across contexts. None of these methods suggest that current practice needs to be discarded. And whether the field adopts all or only some of these recommendations, at least researchers will be bet-

ter equipped to reach conclusions about the complexity and variability of behaviour for many individuals.

Conclusion

The sample of papers considered here provide clear evidence of an assumption that may be implicit in the way researchers in psychology design their studies. The vast majority of papers included conclusions in the Abstracts and/or Discussion sections that implied the results found with aggregated group data also applied to the individuals in those groups and/or applied to individuals in the population. This practice reflects the ergodic fallacy, which is assuming samples are ergodic systems when they are not. The problem with adopting the ergodic fallacy is that, if group-based results do not apply to most individuals, theories of general principles derived to explain these results are not valid. Furthermore, application of these results to the field are limited because interventions developed on the basis of group aggregated data probably will not be effective when applied to individuals. The current research also indicated that the fields of cognitive psychology, educational psychology, and clinical psychology differ in terms of the extent to which they assume their experimental samples are ergodic, possibly for reasons related to the purposes underlying research in those areas.

It is important to note that assuming a sample is ergodic may not always be invalid, but without some check on this, we cannot be confident in conclusions that are based on this assumption. Thus researchers should be aware of this, and all, assumptions that underly their research, and the implications of these for the conclusions that are reached.

Contributions

Contributed to conception and design: CS, MMc
 Contributed to acquisition of data: LP, BJR
 Contributed to analysis and interpretation of data: CS, LP, BJR
 Drafted and/or revised the article: CS, LP, BJR, MMc
 Approved the submitted version for publication: CS, MMc

Competing Interests

The authors have no competing interests to declare.

Data Accessibility Statement

All data reported here are available as Supplementary Material (Tables S1, S2 & S3, all in docx format) and also at osf.io/v8p6b (all in pdf format).

Submitted: July 03, 2023 PST, Accepted: January 13, 2024 PST



This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-4.0). View this license's legal deed at <http://creativecommons.org/licenses/by/4.0> and legal code at <http://creativecommons.org/licenses/by/4.0/legalcode> for more information.

References

- Cervone, D., Mercurio, L., & Lilley, C. (2020). The individual STEM student in context: Idiographic methods for understanding self-knowledge and intraindividual patterns of self-efficacy appraisal. *Journal of Educational Psychology, 112*(8), 1597–1613. <https://doi.org/10.1037/edu0000454>
- Danziger, K. (1994). *Constructing the subject: Historical origins of psychological research*. Cambridge University Press.
- Fisher, A. J., Medaglia, J. D., & Jeronimus, B. F. (2018). Lack of group-to-individual generalizability is a threat to human subjects research. *Proceedings of the National Academy of Sciences, 115*(27), E6106–E6115. <https://doi.org/10.1073/pnas.1711978115>
- Grice, J. W. (2015). From means and variances to persons and patterns. *Frontiers in Psychology, 6*. <https://doi.org/10.3389/fpsyg.2015.01007>
- Griffiths, S., Rhodes, G., Jeffrey, L., Palermo, R., & Neumann, M. F. (2018). The average facial expression of a crowd influences impressions of individual expressions. *Journal of Experimental Psychology: Human Perception and Performance, 44*(2), 311–319. <https://doi.org/10.1037/xhp0000446>
- Hutchinson, K. A., Moffit, C. C., Hart, K., Hood, A. V. B., Watson, J. M., & Marchak, F. M. (2020). Measuring task set preparation versus mind wandering using pupillometry. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 46*(2), 280–295. <https://doi.org/10.1037/xlm0000720>
- Molenaar, P. (2008). On the implications of the classical ergodic theorems: Analysis of developmental processes has to focus on intra-individual variation. *Developmental Psychobiology, 50*(1), 60–69. <https://doi.org/10.1002/dev.20262>
- Molenaar, P. (2013). On the necessity to use person-specific data analysis approaches in psychology. *European Journal of Developmental Psychology, 10*(1), 29–39. <https://doi.org/10.1080/17405629.2012.747435>
- Molenaar, P., & Campbell, C. G. (2009). The new person-specific paradigm in psychology. *Current Directions in Psychological Science, 18*(2), 112–117. <https://doi.org/10.1111/j.1467-8721.2009.0161>
- Moore, S., Speelman, C. P., & McGann, M. (2023). Pervasiveness of effects in sample-based experimental psychology: A re-examination of replication data from nine famous psychology experiments. *New Ideas in Psychology, 68*, 100978. <https://doi.org/10.1016/j.newideapsych.2022.100978>
- Rose, L. T., Rouhani, P., & Fischer, K. W. (2013). The science of the individual. *Mind, Brain, and Education, 7*(3), 152–158. <https://doi.org/10.1111/mbe.12021>
- Speelman, C. P., & McGann, M. (2020). Statements about the pervasiveness of behaviour require data about the pervasiveness of behaviour. *Frontiers in Psychology, 11*, 1–16. <https://doi.org/10.3389/fpsyg.2020.594675>
- Woods, W. C., Arizmendi, C., Gates, K. M., Stepp, S. D., Pilkonis, P. A., & Wright, A. G. C. (2020). Personalized models of psychopathology as contextualized dynamic processes: An example from individuals with borderline personality disorder. *Journal of Consulting and Clinical Psychology, 88*(3), 240–254. <https://doi.org/10.1037/ccp0000472>

Supplementary Materials

Table S1

Download: https://collabra.scholasticahq.com/article/92888-most-psychological-researchers-assume-their-samples-are-ergodic-evidence-from-a-year-of-articles-in-three-major-journals/attachment/194319.docx?auth_token=vqdnBEzvY7Q3HVgjznrC

Table S2

Download: https://collabra.scholasticahq.com/article/92888-most-psychological-researchers-assume-their-samples-are-ergodic-evidence-from-a-year-of-articles-in-three-major-journals/attachment/194320.docx?auth_token=vqdnBEzvY7Q3HVgjznrC

Table S3

Download: https://collabra.scholasticahq.com/article/92888-most-psychological-researchers-assume-their-samples-are-ergodic-evidence-from-a-year-of-articles-in-three-major-journals/attachment/194317.docx?auth_token=vqdnBEzvY7Q3HVgjznrC

Peer Review History

Download: https://collabra.scholasticahq.com/article/92888-most-psychological-researchers-assume-their-samples-are-ergodic-evidence-from-a-year-of-articles-in-three-major-journals/attachment/194318.docx?auth_token=vqdnBEzvY7Q3HVgjznrC
