Journal for Person-Oriented Research
2025, 11(3), 183-187.
Published by the *Scandinavian Society for Person-Oriented Research*
Freely available at https://journals.lub.lu.se/jpor
https://doi.org/10.17505/jpor.2025.28330

Commentary on Lundh's (2025) Review of *Psychology's Misuse of Statistics and Persistent Dismissal of its Critics*

James T. Lamiell

Georgetown University, Washington, DC, USA

E-mail: lamiellj@georgetown.edu

To cite this article: Lamiell, J. T. (2025). Commentary on Lundh's (2025) Review of *Psychology's Misuse of Statistics and Persistent Dismissal of its Critics*. *Journal for Person-Oriented Research*, 11(3), 183-187. https://doi.org/10.17505/jpor.2025.28330

Abstract

This document presents a commentary on Lundh's (2025) review of *Psychology's Misuse of Statistics and Persistent Dismissal of its Critics* (Lamiell, 2019). While the article acknowledges the validity of several points made by Lundh, it also questions the soundness of his critiques in a variety of respects.

I am grateful to Prof. Dr. Lars-Gunnar Lundh for devoting the time and effort necessary to author a review of my book, *Psychology's Misuse of Statistics and Persistent Dismissal of its Critics* (referred to below as *Misuse*; Lamiell, 2019). I'm additionally grateful to Lundh for inviting me to comment on his review in the pages of this journal.

Lundh (2025) has accurately identified my central concern in *Misuse* as the mistakenness that results when statistical knowledge that is defined only for aggregates of individuals is interpreted, discussed, and understood as if it were also knowledge about the individuals within those aggregates. However counterintuitive the point may be, such treatments are valid only under empirical circumstances that are imaginable but rarely—if ever—realized in practice. Consequently, those treatments routinely amount to bad science, and under certain practical/applied circumstances can lead to bad and even ethically questionable professional practices.

Early in his remarks, Lundh questions my "bad science" argument on the grounds that statistical knowledge defined only for aggregates can, after all, say something of potential significance about the relationship between psychological variables (see Lundh, 2025, p. 121). However, Lundh then quickly concedes the validity of the "bad science" argument "if empirical findings about the relationship between variables are interpreted as facts about individuals" (p. 121). Since that very conditional was stipulated in my argument from the start (it can be found in the last sentence in the first paragraph on p. 1 of *Misuse*), Lundh's expression of skepticism in this connection, however fleeting, struck me as gratuitous.

Yet, reading further into Lundh's text, I could see that deeper and more consequential misunderstandings are at play there. For example: Lundh argues that one could easily construct a scattergram of a set of data points marking individual differences on some two variables, call them 'X' and 'Y,' in order to visualize "how the correlation between those two variables looks for each individual in the study." Alas, that suggestion collides headlong with the conceptual reality that for each individual in the study the correlation in question doesn't look like anything at all, because the correlation in question is not defined for each—or even any—individual in the study.

If one were to deconstruct Lundh's hypothetical X,Y correlation in order to plot, for each individual, the datum that had been entered into the calculation of that correlation, each such plot would display precisely one data point. That point would be located at the graphic intersection of two lines, real or imagined: one extending vertically from the plot's x-axis at the point representing an individual's standing on Variable X, and the other extending horizontally from the plot's y-axis at the point representing that same individual's standing on Variable Y. None of these plots for individuals would display any scatter at all.

Contrary to Lundh's claim, the graphical illustration that I discussed on pp. 27-29 of *Misuse* was not introduced for the purpose of revealing "otherwise hidden and contradictory facts about percentages of individuals" (Lundh, 2025, p. 121). The point of the illustration was to show that one cannot validly generalize about *individuals* (not 'percentages of individuals') on the basis of knowledge about some empirical state of affairs (such as an r – or an r^2 – value) that has been found to exist only for a population.

A major flaw in Lundh's argument is his mistaken belief that a claim to knowledge about a percentage of individuals is a knowledge claim about individuals. To the contrary: a knowledge claim about some percentage of individuals is a factual claim about the size of some subset of a larger whole. The larger whole is a population, and unless the claimed percentage of that larger whole amounts to only one individual, that claimed percentage is itself a population.

Lundh's mistakenness in this matter is compounded just a few lines further on in his text. There he refers to the results of a large statistical meta-analysis of the findings of RCT studies reported by Cuijpers et al. (2023), showing that "42% of the patients in cognitive-behaviour therapy (CBT) responded to treatment, while the response rate was only 19% in the control groups." Lundh labels the Cuijpers et al. report as "a report about individuals, not a report about the relations between variables (p. 121)." For reasons just explained, the Cuijpers et al. (2023) report is not a report about individuals, and an appreciation for this logical fact can help one to see that the claim by Cuijpers et al. (2023) to knowledge about "42% of patients exposed to CBT as opposed to 19% of control group patients" is a claim about a *variable* defined for—and only for—a *population*. The variable in this case is 'CBT treatment vs. no treatment control,' and the population is the one implicitly sampled by the collections of participants in the studies that Cuijpers et al. (2023) meta-analyzed.

The relevant lesson here concerning the nature of claims to knowledge about percentages of populations is further conveyed by the text on pp. 116-118 of *Misuse*, culminating in the citation from an 1865 publication by the French physiologist Claude Bernard (1813-1878). Readers of Lundh's review of *Misuse* and of the present commentary on that review are urged to (re-)examine very carefully the pages from *Misuse* just cited.

Lundh validly points out that, contrary to what is implied by the wording of my claim on p. 158 of *Misuse*, the findings of RCT studies are not always based on the comparison of treatment group averages. They might at times entail the consideration of some other aggregate indices, such as treatment group percentages. In any case, however, the ultimate scientific warrant for a knowledge claim issuing from an RCT study is what is revealed by a statistical analysis of records of trial outcomes that have been compiled by aggregating those records across the multiplicity of patients within each of the entities being compared—entities, it is well to note here, that are traditionally and appropriately referred to as 'treatment *groups*.' Whether the aggregation arithmetic entails tallying the frequency with which various categorical outcomes occur, or computing averages on some outcome measure(s) is not epistemically crucial to the argument developed in *Misuse*. What is epistemically crucial to that argument is that the indices being statistically compared are aggregate, group-level indices.

In *Misuse*, I sought to head off confusion on some important points by incorporating into Chapter 4 (see pp. 80-82) a distinction between the research *results* recorded for individual subjects while a study is in progress, and the research *findings* that are ultimately reported to the scientific community as the empirical outcomes of group-level investigations—including but not limited to RCT studies and other treatment group experiments. As explained in that discussion, this distinction was unnecessary in the original version of experimental psychology, dominated as it was by the 'Leipzig model' of investigation. Under the terms of that model, psychological research was conducted on an '$N = 1$' basis, making an experiment's *findings* one and the same with the *results* respectively obtained for individual subjects. However, this ceased to be the case under the terms of the 'treatment group' model of investigation that eventually supplanted the Leipzig model. Under the terms of the treatment group model, which we may refer to as an '$N =$ many' model, investigative outcomes for individual subjects, i.e., 'results,' are still recorded, as they must be else there will later be nothing to aggregate into the indices that will eventually define group-level outcomes, but it is via the statistical analysis of the group-level outcomes, and *not* via the consideration of individual-level results, that an experiment's *findings* are ultimately scientifically arbitrated.

Lundh is justified in pointing out that an investigator/investigative team will normally have access not only to a study's aggregate-level statistical findings but also to the individual-level empirical results on which those findings are based. In principle, therefore, it is, as Lundh claims, normally possible to shift back-and-forth between the consideration of aggregate-level statistical findings and the individual-level empirical results.

Clearly, however, publications offering readers such perspectival shifts are not now and long have not been the norm in the contemporary mainstream psychological literature. Instead, publications reporting empirical research outcomes regularly focus primarily or even exclusively on the findings revealed by aggregate statistical analyses. This is a direct reflection of the thoroughly dominant but sadly mistaken view that the domain of aggregate statistical research findings is the appropriate domain for scientifically testing claims about the generality across individuals in certain aspects of their psychological functioning.

It is in this connection where the genuine merit of the work reported by Grice (2015) and discussed in Chapter 7 of *Misuse* is to be found. Grice (2015) showed that if an hypothesis about the effect of a given experimental treatment is tested on a case-by-case basis, as should routinely be the case in psychological science, and as his Observation Oriented Method (OOM) makes possible again in reviving the spirit of the original Leipzig model for psychological experimentation, empirical outcomes can vary widely across individual cases. Grice (2015) found this to be the case even as he varied the stringency of the standard that was set in order for a given individual-level test outcome to be counted as an instance of hypothesis confirmation. Grice's work thus vividly illustrates the *inappropriateness* of relying on the outcomes of aggregate-level statistical analyses for generalizing across individuals about the effect(s) of experimental treatments.

Lundh sees in Grice's method a procedure 'analogous' to the practice, common in RCT studies in the domain of psychotherapy, of recording a tally for each patient to indicate whether or not any clinically significant improvement had been observed over the course of the study. Such as it is, however, the analogy that Lundh has seen is strictly superficial.

The outcome tallies recorded for individual patients in RCT studies are what I refer to in *Misuse* (and above) as individual-level research *results*. They may be all that is of interest from a purely clinical standpoint, but things are different from the standpoint of validating scientifically a claim about the 'general' effectiveness of some therapeutic approach relative to others. For that purpose, traditional mainstream thinking requires the aggregation of research results across the *N* individual patients who had been exposed to a given therapeutic treatment, and that the outcome of a statistical comparison of the aggregated tallies across the different treatment groups then be presented to the scientific community as the RCT study's *findings*.

The primary function of each one of the respective tallies Grice recorded for individual subjects in the experiments he reported in his 2015 article was to test an hypothesis about the effect of an experimental treatment (rejection or no-rejection) on the popularity rating that a given subject would assign to the apparent agent of that subject's treatment. It is important to appreciate that (a) in the service of this hypothesis-testing function, no *aggregation* of the tallies was necessary, or would have been appropriate, within the OOM framework, and that (b) this is emphatically *not* analogous to, but in fact starkly contrasts with the use to which treatment outcome tallies are put in RCT studies, as just described.

Now: Grice's agenda in the work he reported in his 2015 article went beyond hypothesis testing in accordance with OOM procedures. His additional objective was to contrast the conclusion that would be indicated by the outcomes of such individual-level testing with the conclusion that would be judged warranted if the results of each of his three studies were analyzed by conventional *t*-tests. In order to achieve that objective, Grice had to proceed in a way that conformed to the procedures conventionally employed in RCT studies. This fact might well be responsible for the appearance of the procedural similarities that Lundh has perceived.

Note carefully, however, that there is nothing in the logic of OOM that demands the particular investigative agenda that Grice was pursuing in the work he reported in his 2015 article. Indeed, if and when a critical mass of mainstream researchers can finally be persuaded to abandon long-dominant but thoroughly inappropriate practices in favor of procedures that actually are logically suited to the project of advancing our scientific knowledge about the psychological doings of individuals, then the specific investigative agenda that Grice adopted—with very good reason—for the demonstrative purposes of his 2015 article will no longer be called for. At that point, should the field ever reach it, (a) the inappropriate methods will have gradually disappeared from the psychological research literature, (b) Grice and other practitioners of OOM will be free to focus on testing their theoretically-based hypotheses at the individual level, no longer hindered and distracted—as they are now—by continued widespread adherence to methods that were never appropriate for that agenda to begin with, and (c) the superficial similarity between OOM and RCT methods that Lundh now sees will have disappeared. Had Lundh more fully appreciated these points, perhaps he would have found in *Misuse* more of the guidance he claims to have been seeking on "how we should proceed into the future" (p. 124).

Regarding the concerns I expressed in *Misuse* about the socio-ethics of certain possible practical applications of the findings of RCT studies, Lundh states that my argument implies that

> anything goes, including treatments for which there is no evidence from population level research. ... Perhaps this can be seen as a *reductio ad absurdum* of his argument in this area. (Lundh, 2025, p. 123, emphases in original.)

In this context, Lundh's claim that I advocate resistance to treatment "*recommendations*"—his emphasis—is egregiously inaccurate. I view the inaccuracy as 'egregious' because the passage from *Misuse* to which Lundh then appeals to support his characterization of my thinking clearly indicates that my advocacy is of resistance to treatment *prescriptions*—not "recommendations"—and in the relevant section of text in *Misuse*, I further explain my rationale for this view.

There I state my concern that *pre*scriptions would entail *pro*scriptions, meaning that the alternatives made available to individual clients and their respective therapists would be regulated not by objective knowledge known to be applicable to that individual client but, instead, by the subjective belief(s) of one or more individuals empowered, e.g., by insurance companies or government agencies, to make pre- and proscriptive availability determinations (see also endnote 3 in Chapter 6 of *Misuse*, called out on p. 131 and printed on p. 142). My advocacy of resistance to such treatment pre- and proscriptive determinations is certainly not advocacy of resistance to treatment *recommendations* that would be *guided*—not regulated—by the findings of RCT studies, nor is it by any means an endorsement of 'anything goes'—including exorcism!—in therapeutic treatments. I stand by my argument in *Misuse* that any regulatory circumscription of treatment alternatives for individuals based on the aggregate findings of RCT studies would be ethically questionable. Ultimately, the reason for this is that, as Lundh himself acknowledges further on in his discussion, "at the level of the individual it is impossible to know [from the findings of RCT studies] if a given patient will benefit from [some particular] treatment (p. 122, brackets added)." Additional documentation of my views on this matter, published two years prior to *Misuse*, can be found in Lamiell and Martin (2017, pp. 223-224).

In the penultimate major section of his review of *Misuse*, Lundh questions the validity of the distinction I have drawn between *psychological science* and *psycho-demography*. I have characterized the overriding knowledge objective of the former as that of advancing scientifically our understanding of the psychological doings of persons. By contrast, I argue, psycho-demography serves to expand and systematize our knowledge of population-level patterns in psychological phenomena. Without question, psycho-demographic knowledge can be valuable in various ways, including that of complementing psychological science, and this is a point that I discuss at some length in Chapter 7 of *Misuse*. Nevertheless, it remains the case that psycho-demographic knowledge is, by its very nature and quite literally, knowledge of no one, and I do not see how knowledge of no one can qualify as knowledge of the distinctive sort being sought in a psychological science.

In challenging my stance on this point, Lundh re-introduces his argument (cf. Lundh, 2023) in favor of recognizing what I call 'psycho-demography' as one of three branches of psychological science. He would call that branch 'population psychology,' and place it alongside 'person psychology' and 'mechanism psychology.' My critical appraisal of Lundh's proposal is already part of the archival record (cf. Lamiell, 2024), and I will not revisit it here except to note that it questions, as I do here, the very notion of a 'population psychology.'

In the heading of the concluding sub-section of his review of *Misuse*, Lundh asks rhetorically "Is psychology only about individuals?" (p. 124). Unfortunately, the brief text which follows in pursuit of this question is conceptually problematic at multiple points.

First, Lundh expresses uncertainty about how to reconcile my commitment to the view that "psychology can only be about individuals, never about populations of individuals" (Lundh's wording) with the fact that I "speak positively about the nomothetic search for *general laws* which apply to all people" (p. 124). The tension that Lundh finds here is immediately resolved by the realization that, contrary to many decades of misunderstandings about the nature of nomothetic knowledge of persons, the quest for such knowledge properly proceeds not apart from but necessarily *through* the study of individual persons.

Genuinely nomothetic knowledge of persons cannot be achieved simply through studies of populations. This is because such studies cannot validly be regarded as yielding knowledge of what is true for *any* particular individual, let alone for *each of many* individuals, and the latter is what nomothetic knowledge of *persons in general* is properly understood to be. Nor does it matter if, historically, "it has not been easy to find general laws that apply to all individuals" (Lundh, 2025, p. 124). Regardless of the relative difficulty of acquiring knowledge of what is true of persons *in general*, as opposed to knowledge of what is true of persons *on average* (or in some other merely aggregate way), the latter variety of knowledge simply does not qualify—has never *truly* qualified—as nomothetic knowledge of persons (cf. Lamiell, 1998, 2003).

Lundh wonders "why it should be psychology only when we identify [phenomena] that apply to *all* people, but not when we study [phenomena] where there is individual *variation*?" (p. 124, emphases in original; brackets added). Nowhere in *Misuse* do I argue that *only* inquiry culminating in knowledge that applies to all people—i.e., inquiry that would effectively produce nomothetic knowledge—could qualify as psychological. The argument is simply that *psychological* inquiry must concern some aspect(s) of the psychological doings of *persons*. Knowledge that eventually proves applicable only to some persons, or even to only one person (and so might be termed partly or fully *idiographic* knowledge) can certainly qualify as psychological knowledge.

What, in my view, cannot qualify as psychological knowledge is aggregate statistical knowledge of no one. It should be noted, however, that this view does not preclude the very possibility of a 'social' psychology. Indeed, in work published more recently than *Misuse*, I have sought to provide at least a glimpse of what a 'critically *inter*-personalistic' perspective in various social contexts might entail (see Lamiell, 2024). Much further work in this direction is needed.

Lundh expresses concern about my alleged disinterest in substantive domains of inquiry where there is 'individual variation' (Lundh, 2025, p. 124), by which I take him to mean between-person variation (as opposed to within-person variation over time). Contrary to his contention, however, I do not see the phenomenon of between-person variation as, in and of itself, problematic for psychology. The question is: how should *psychological* inquiry proceed in the face of such variation?

A major implication of the argument developed in *Misuse* is that the research psychologist should take into account whatever is substantively the case with the person who is the subject of the researcher's immediate attention. That researcher should then be prepared to shift the substantive focus in whatever way necessary as attention shifts from one research subject to another. Proceeding thusly, successive investigative encounters with *persons* of, say, different mental acuities, or artistic sensitivities, or degrees of sociability (to name just three possible dimensions of individual variation) can proceed without ever abandoning the study of persons for the study of one or more *variable(s)* marking *between-person differences*. That is the key.

As I explain in Chapter 3 of *Misuse*, the latter course is the one historically taken a century ago by leading differential psychologists such as E. L. Thorndike (1874-1949) and Hugo Münsterberg (1863-1916). That course was then further—and influentially—pursued by such next-generation disciplinary luminaries as Anne Anastasi (1908-2001) and Leona Tyler (1906-1993), though late in her career Tyler did come to recognize the error of that way for one concerned with human individuality (Tyler, 1978; cf. *Misuse*, p. 65). In this connection, a major question raised by *Misuse* for contemporary psychological investigators is: Other than possibly as the focus of a subject's own thinking about self and/or others, just where in the study of *a person's* psychological doings would an investigator expect to empirically encounter the phenomenon of between-person *differences*?

In the light of Lundh's review of *Misuse*, I continue to recommend careful contemplation of this question, as it offers yet another conceptual avenue to the realization that the study of individuals and the study of individual differences are two quite distinct projects. I stand by my argument in *Misuse* that while the first-named of these projects lies squarely within the purview of a scientific psychology, the second-named really is more accurately understood as work generating knowledge of an essentially psycho-demographic nature.

Open access

References

Cuijpers, P., Miguel, C., Harrer, M., Plessen, C. Y., Ciharova, M., Ebert, D., & Karyotaki, E. (2023). Cognitive behavior therapy vs. control conditions, other psychotherapies, pharmacotherapies and combined treatment for depression: a comprehensive meta-analysis including 409 trials with 52,702 patients. *World Psychiatry, 22*(1), 10 5-115. https://doi.org/10.1002/wps.21069

Grice, J. W. (2015). From means and variances to persons and patterns. *Frontiers in Psychology, 6,* 1-12. https://doi.org/10.3389/fpsyg.201501007

Lamiell, J. T. (1998). "Nomothetic" and "idiographic": Contrasting Windelband's understanding with contemporary usage. *Theory and Psychology, 10,* 715-730.

Lamiell, J. T. (2003). *Beyond individual and group differences: Human individuality, scientific psychology, and William Stern's critical personalism*. Sage Publications.

Lamiell, J. T. (2019). *Psychology's misuse of statistics and persistent dismissal of its critics*. Springer-Nature. https://doi.org/10.1007/978-3-030-121131-0

Lamiell, J. T. (2024a). On persons, populations, and causal mechanisms. Some critical reflections on Lundh (2023). *Journal for Person-Oriented Research, 10*(1), 61-63. https://doi.org/10.17505/jpor.2024.xxxxx

Lamiell, J. T. (2024b). *Primer in critical personalism: A framework for reviving psychological inquiry and for grounding a socio-cultural ethos*. Routledge. https://doi.org/10.4324/9781003375166

Lamiell, J. T., & Martin, J. (2017). The incorrigible science: A conversation. In Macdonald, H., Goodman, D., & Becker, B. (Eds). *Dialogues at the edge of American psychological discourse*, pp. 211-244. Palgrave-Macmillan. https://doi.org/10.1057/978-1-1137-59096-1_8

Lundh, L. G. (2023). Person, population, mechanism. Three main branches of psychological science. *Journal for Person-Oriented Research, 9*(2), 75-92. https://doi.org/10.17505/jpor.2023.25814

Lundh, L. G. (2025). Book review: "Psychology's Misuse of Statistics and Persistent Dismissal of its Critics" by James T. Lamiell. *Journal for Person-Oriented Research, 11*(2), 120-125. https://doi.org/10.175/jpor.2025.28098

Tyler, L. E. (1978). *Individuality: Human possibilities and personal choice in the psychological development of men and women.* Jossey-Bass. https://doi.org/10.1007/978-3-030-65613-3_9