

**Problematic Research
Practices and Inertia in
Scientific Psychology
History, Sources, and
Recommended Solutions**

**Edited by James T. Lamiell
and Kathleen L. Slaney**

2020

 **Routledge**
Taylor & Francis Group
LONDON AND NEW YORK

2 On the systemic misuse of statistical methods within mainstream psychology

James T. Lammiell

The concern in this chapter is with the troubling inertia among mainstream psychological investigators in the face of repeated critiques of the long-standing practice of regarding statistical knowledge of populations as if it warrants claims to scientific knowledge about the psychological functioning of individuals within those populations. The problematic nature of this common interpretive (mal)practice has been pointed out in various ways in publications dating at least to the middle of the 20th century. Scholar David Bakan (1921–2004) discussed the problem in terms of a widespread conflation by researchers of knowledge of the *general* with knowledge of the *aggregate* (cf. Bakan, 1955, 1966). As he correctly stated, “the use of methods which are appropriate to [gaining] the [former] type [of knowledge] in the establishment and confirmation of the [latter] leads to error” (Bakan, 1955, p. 211, brackets added). Bakan’s admonitions were widely ignored.

Subsequently, the well-known research methodologist, Fred N. Kerlinger (1910–1991), discussed in a 1979 text what he termed a “troublesome paradox” in much psychological research, created by treating group-level statistical findings as informative about the individual-level ‘doings’ of theoretical and/or practical interest (Kerlinger, 1979, p. 275). Although Kerlinger (1979) did not cite Bakan (1955, 1966), the conceptual problem Kerlinger identified is the direct consequence of the very conflation Bakan had discussed. Like Bakan, however, Kerlinger, too, was ignored—even, as it turns out, by himself, as he paid no further attention to the ‘conceptual paradox’ in the research methods text he published seven years later (Kerlinger, 1986).

Revisiting in 1981 what is essentially the same problem, the present author drew attention to mainstream ‘personality’ psychologists’ fallacious practice of interpreting of correlational indices of the reliability and validity of personality tests as empirical grounds for empirically evaluating the theoretical assumption of temporal and

trans-situational consistency in the manifestation by individuals of their respective traits (Lammiell, 1981). That publication was followed by numerous works, published periodically over the next four decades (cf. Lammiell, 1987, 1997, 2003, 2015, 2016), elaborating, from a variety of angles, on the problematic nature of established interpretive practices in both correlational and experimental studies (e.g., Lammiell, 2019). Yet, false interpretations of aggregate statistical knowledge have continued—indeed, have proliferated—while the critiques of those interpretations have been, for the most part, simply ignored (see, e.g., the commentary by Rom Harré (1927–2019) on the imperiousness of mainstream personality investigators to the various critiques just cited; Harré, 2006).

The conceptual problems here are fundamentally epistemic: fallacious interpretations of aggregate statistical findings result in unjustified claims to scientific knowledge about individuals. This is bad science. Beyond that, however: when those ill-founded knowledge claims are exercised in the applied domain in order to justify interventions in the lives of individuals, the problems take on a socio-ethical dimension as well (cf. Lammiell, 2019, chapter 6). It is to be hoped, therefore, that the present discussion, though necessarily brief, will prompt readers who remain oblivious to these critical discussions or skeptical of their merits to finally consider or reconsider, thoughtfully and in duly critical fashion, a more detailed examination of the relevant issues that can be found elsewhere (Lammiell, 2019). In that work, the various facets of the problem are explored in greater detail and depth than is possible here.

The present chapter offers a highly condensed treatment of the nature and historical development of the problem. This is followed by a discussion of the inadequacy of certain attempted rebuttals of the critique that have been reiterated in recent publications (Banicki, 2018; Proctor and Xiong, 2018).

The influx and proliferation of aggregate statistical methods in psychology: A brief review of some key historical developments

The aggregate statistical methods that have come to thoroughly dominate research practices within the mainstream of scientific psychology played no role in the experimental psychology famously pioneered by Wilhelm Wundt (1832–1920) in Leipzig in 1879. That discipline was oriented toward the discovery of lawful regularities in various aspects of psychological functioning (e.g., sensations, perceptions, judgments, memories, and other psychological processes) that would be *general* in the sense of *common to all of the individuals investigated*. Precisely because “*each individual* [was to be regarded as] *a particular in which the*

general is manifest" (Bakan, 1966, p. 433, emphasis in original, brackets added), experiments had to be conducted in such a way that the findings would be completely specifiable for each investigated individual. In other words, the original general-experimental psychology had to be prosecuted on what would be called, in modern parlance, an "N = 1" basis. This fact is reflected in the frequent references one can find in the literature of the day to the general-experimental-*individual* psychology (see, e.g., Stern, 1900; Wundt, 1912; cf. Lamie, 2019).

The differential psychology that William Stern (1871–1938) founded at the turn of the 20th century was introduced by him as a *complement to*—not a substitute for—the general-experimental-individual psychology (Stern, 1900). Rather than focusing on common-to-all regularities in the psychological doings of individuals, the focus in this newly proposed sub-discipline would be on systematic *differences between* individuals in their psychological doings. The pursuit of that knowledge objective would require, by its very nature, an 'N = many' approach, and investigators would therefore have extensive use for just those aggregate statistical methods that were eschewed within the original general-experimental-individual psychology.

The epistemic implications of the shift from studies of *individual* psychological doings to studies of *individual differences* in psychological doings was not lost on William Stern. To the contrary, and though the matter was not a point of emphasis in his 1900 book, it was explicitly recognized in his second book on differential psychology, titled (in translation) *Methodological Foundations of Differential Psychology* and published in 1911 (Stern, 1911). In this latter work, Stern made clear his recognition of the logical fact that knowledge of the statistical properties of *variables* in terms of which individuals have been empirically differentiated from one another, i.e., knowledge of such parameters as the means, variances, and co-variances of those variables within populations, is not knowledge of the *individuals* who have been differentiated in terms of those variables. It is most unfortunate that this critically crucial insight of Stern's was obscured in the writings of other influential differential psychologists of his time.

For example, E. L. Thorndike (1874–1979) wrote in his book titled *Individuality* (coincidentally published in the same year as Stern's *Methodological Foundations of Differential Psychology*), that the correlation between two variables measuring between-person differences in two traits indicates "the extent to which the amount of one trait possessed by an individual is bound up with the amount he possesses of some other trait" (Thorndike, 1911, p. 22, emphasis added). This claim is simply false (see Lamie, 1987, pp. 90–100; Lamie, 2019, pp. 28–30).

Similarly, Hugo Münsterberg (1863–1916) wrote in his widely read paean to applied psychology titled *Psychology and Industrial Efficiency* (Münsterberg, 1913) that, given knowledge of the correlation between variables marking individual differences in distinct features of the psychological function of attention, "the manifestation of one feature ... allows us to *presuppose without further tests* that certain other features may be expected in the particular individual" (Münsterberg, 1913, p. 136, emphasis added). This claim, likewise, is false.

Altogether contrary to the stance taken by Stern (1911), the view projected in these quotations of Thorndike (1911) and of Münsterberg (1913) is that knowledge of variables in terms of which individuals are differentiated just is, at one and the same time, knowledge of the individuals who have been differentiated in terms of those variables. Moreover, it is this view, and not Stern's, that would be adopted by two of the most prominent of the next generation of differential psychologists.

The voluminous writings of Anne Anastasi (1908–2001) in differential psychology, beginning with the first edition of her highly influential textbook on the subject (Anastasi, 1937) and extending over many decades, give no recognition to the variables-individuals distinction drawn by Stern (1911). As a consequence, and quite unlike Stern, Anastasi (1937) viewed differential psychology not as a disciplinary complement to the general-experimental-individual psychology but rather as an investigative framework in which

... the fundamental questions are *no different* from those of general psychology. [For] it is apparent that if we can explain why individuals react differently from one another, we shall understand why each individual reacts as he does.

(Anastasi, 1937, p. vi, emphasis and brackets added)

In another widely read differential psychology text, the first edition of which appeared ten years after Anastasi's, Leona Tyler (1906–1993) confined her discussion of the historical roots of differential psychology to Stern's 1900 text (Tyler, 1947). As noted above, however, the variables-individuals distinction was not explicitly discussed by Stern in that text. By ignoring Stern's 1911 text, therefore, Tyler effectively blinded herself to that distinction, and so, if only by default, embraced the view of differential psychology shared by Münsterberg, Thorndike and, later, Anastasi.²

This, then, was the conceptual soil in the one of psychology's 'two disciplines' that Cronbach (1957) characterized as the "Holy Roman Empire" (p. 671) of 'correlational' (i.e., differential) psychology. In

that soil, the view took firm root that statistical studies of individual differences can secure scientifically factual knowledge of the psychological functioning of individuals. Meanwhile, in the other of scientific psychology's 'two disciplines,' characterized by Cronbach (1957, p. 671) as the "Tight Little Island" of experimental psychology, developments were underway that would link it, epistemically, to correlational psychology. Knowledge claims in the two sub-disciplines would then be expressible in a common language.

Correlational studies of variables marking individual differences seemingly made differential psychology much better suited than the $N = 1$ framework for experimentation: established by Wundt (refer above) for addressing practical problems arising outside of psychology's 'brass instrument laboratories, e.g., in schools, business and industry, health care, and the military. As Münsterberg (1913) argued, a viable applied psychology would have to accommodate the reality—a reality avowedly outside the purview of the general-experimental-individual psychology—that "there are gifted and ungifted, intelligent and stupid, sensitive and obtuse, quick and slow, energetic and weak individuals" (Münsterberg, 1913, p. 10). However, a wholesale adoption of differential psychology's strictly correlational methods would undermine the discipline's suitability for discovering causal relationships (Danziger, 1990). That capacity was a strength of Leipzig-model experimentation, and was a feature widely considered essential to maintaining regard for psychology, both by psychologists themselves and by scientists in other disciplines, as a basic explanatory science and not merely as a kind of trade guild devoted to the production of statistical knowledge that could be put to practical use (cf. Wundt, 1913, 2013).

It was the availability of a form of experimentation that had been utilized for decades by medical researchers (Porter, 1986), treatment group experimentation, that seemed to obviate the need for psychologists to choose between the serviceability of correlational methods for the practical purposes of an applied psychology and the epistemic power of experimental methods for the scientific purposes of a basic psychology (Danziger, 1987, 1990). In the simplest form of such experimentation, groups of subjects are formed by assigning individuals at random to one of the two or more treatment conditions defining an independent variable (IV). A measure is then taken of the performance of each subject on some dependent variable (DV) used to represent ('operationally define') some or another psychological function. The findings of the experiment are then revealed by the outcome of a statistical comparison of the average DV values compiled by the subjects within each of the treatment groups. Where the difference(s) between those average DV

values are greater than would be expected on the basis of random variation alone, it is concluded that the difference(s) are not the result of random variation alone, and it is inferred that the difference(s) reflect(s), at least in part, the causal effects of the differential treatments imposed upon the respective groups of subjects.

As experimental psychologists became ever more enamored with the treatment group approach to experimentation, the sub-discipline as a whole became ever better aligned with the work of the differential psychologists, as the work in both sub-disciplines was being devoted to statistical analyses of variables in terms of which individuals were being differentiated. The differential psychologists were using tests of one sort or another to statistically investigate correlational patterns in between-person differences that had arisen outside the laboratory. The experimental psychologists, for their part, were statistically investigating what were seen as the causal effects of between-person differences that they themselves were creating inside the laboratory. The statistical concepts and methods of investigation were the same across the two sub-disciplines (cf. Cohen, 1968; Kertinger and Pedhazur, 1974), and, in both, the view prevailed that knowledge of the variables in terms of which individuals are differentiated—experimentally or otherwise—is *de facto*, knowledge of the individuals who have been differentiated in terms of the variables. In the words of Danziger (1987), the ascendant interpretive custom was one according to which

... the *statistical* structure of the *data* based on the responses of many individuals is assumed to conform to the [*psychological*] structure of the relevant *processes* operating on the *individual* level.

(Danziger, 1987, p. 45, emphasis and brackets added)

Banicki: "Is the conceptual gap really so unbridgeable?"

Commenting critically on a recent article by the present author pointing out—yet again—the fundamental epistemic problem created by the conceptual gap between knowledge of variables defined only for populations and knowledge of the individuals differentiated from one another in terms of those variables (Lamie, 2018a), Banicki (2018) subtitled a section of his text with the very question (which he posed rhetorically) in the heading above, and proceeded to defend his conviction that the correct answer to that question must be 'no.' He conceded the logical validity of the claim that the conceptual gap exists, but then dismissed the claim as "trivial" (p. 259) thanks to the long-standing convention of

framing knowledge claims about individuals *probabilistically*. The same argument has been made in another recent article by Proctor and Xiong (2018; cf. Lamiell, 2018b, 2018c).³

The conceptual essence of probabilistic thinking as it is applied in this context is that knowing the proportion, P , of individuals within a population of size N who exhibit some phenomenon X under specified circumstances, one can validly infer that the probability is likewise P that, under those circumstances, an individual member of the designated population will exhibit X . In just this way, it is thought, one anchors the conceptual bridge that delivers one from knowledge of aggregates of individuals to grounds for valid claims to knowledge of the individuals within the aggregates. As an illustrative case, Banicki (2018) cited the findings of an experiment reported by Isen and Levin (1972), showing, as Banicki (2018) put it, that “people who had found a dime in a phone box were about 20 times more likely to help a stranger than those who had not been so lucky” (Banicki, 2018, p. 259).

In critically appraising this claim, it must be asked of Banicki’s statement about “people” who found/did not find a dime in a phone box: *Which people?* Is the reference to each of ‘this,’ ‘that,’ and ‘the other’ person who found/did not find a dime? Is the reference to only some of them? In the latter case, which ones? And: how to know? Or is the reference actually to no particular person at all, but instead just to ‘people’ in the amorphous aggregate? Locutions of the sort used by Banicki (2018) are ubiquitous in the literature of mainstream psychology, and, routinely, the questions just posed are simply left to go begging.

Of immediate relevance to those questions is a basic logical reality discussed by the British polymath John Venn (1834–1923) over 130 years ago in his 1888 treatise of the concept of probability (Venn, 1888). Brought to bear on the Isen and Levin (1972) experiment cited by Banicki (2018), that basic logical reality was that any given subject in that experiment either *would* or *would not* exhibit the target helping behavior. Just as surely as in a coin flip exercise like the one discussed for illustrative purposes by Venn (1888), where it is known in advance that any one flip of a fair coin will yield *either* a ‘head’ or a ‘tail,’ the ‘would help/would not help’ binary was the reality *for each and every one* of the individual subjects in the Isen and Levin (1972) experiment, and this was true regardless of a given subject’s experimental condition.

Banicki’s (2018) ‘20 times more likely’ claim is based on an empirical reality that Isen and Levin (1972) documented by tallying up their ‘helped/did not help’ observations across subjects in each of the two groups into which they had been experimentally divided. One of those groups was composed of helpers who had found a dime; the

other group was composed of helpers who had not. On the question of whether ‘Smith,’ who was observed to help, had been assigned to the ‘found a dime’ group or the ‘did not find a dime’ group, the experimental findings reported by Isen and Levin (1972) and cited by Banicki (2018) are *completely silent*, and this is true for *every single subject* in the experiment. The ‘20 times more likely’ reality was nothing that was or possibly could have been established for any individual subject. All of this just *means* that the findings of the Isen and Levin (1972) experiment documented an empirical reality that was observed of, quite literally, *no one*. Knowledge of no one cannot possibly be knowledge of *any* living, breathing *someone*.

Unfortunately, mainstream research psychologists routinely take the ill-begotten liberty of discussing what their research reveals about ‘people’ in such a way as to imply an understanding of the term ‘people’ not as a reference to a *set* of individuals as a singular entity, but instead as a reference to a plurality of individuals each of whom is to be regarded as a distinct entity. Implicitly invoking just such an understanding, Banicki (2018) has effectively taken the findings of the Isen and Levin (1972) experiment as scientific evidence that a —i.e., some given individual person who has found a dime in a phone box would be “about 20 times more likely” to help a stranger than a —i.e., some other given individual person who has not found a dime. Perhaps prepared to go even a step further, Banicki (2018) might possibly see the findings of the Isen and Levin (1972) experiment as evidence that ‘an individual person who has found a dime would be “about 20 times more likely” to help a stranger than that same individual would be were no comparable serendipitous discovery made. From a scientific standpoint, the obvious problem with either of these individual-level interpretations is that neither conforms to what the Isen and Levin (1972) experiment actually showed. With reference to any given individual, then, Banicki’s (2018) ‘20 times more likely’ statement cannot stand as a claim to *knowledge* about that individual. It could stand as an expression of Banicki’s *subjective belief* about that individual, but it would have to be understood that that belief might or might not be confirmed by actual knowledge of what the individual does (or has done), and there is nothing in the findings reported by Isen and Levin (1972) that establishes otherwise.

The larger lesson here is that no research findings of imperfect statistical relationships between variables marking between-person differences, whether as correlates of differences that have been assessed nonexperimentally by means of tests, or as effects of differences that have been created experimentally by alternative treatment conditions, can ever warrant claims to probabilistic knowledge about ‘people’ if

'people' is understood as a reference to distinct individual entities—i.e., *persons*—and not simply as a reference to unitary aggregates of such entities. This is why the only correct answer to the question posed rhetorically by Banicki (2018; refer above) is: 'Yes, the conceptual gap between knowledge of aggregates and knowledge of individuals within those aggregates really is unbridgeable.' Matters can be made to seem otherwise only by abandoning the frequentist understanding of probabilistic knowledge as tied inextricably to the consideration of a multiplicity of observations and embracing instead the subjectivist understanding of probabilistic statements as expressions of belief. Unfortunately, mainstream thinking in psychology has long systematically conflated these two understandings of probability (Cowles, 1989).

Nor should it be acceptable within a genuine science to maintain this indifference simply because doing so "appears to be working" (Banicki, 2018, 258). In the face of that notion, the question begs: *Working in what sense, and for whom?*

On the 'practical utility' defense of traditional mainstream interpretive practices

We can consider in this connection a hypothetical scenario constructed around the use of psychological tests as instruments for the pre-employment screening of job applicants (cf. Hogan, Hogan, and Roberts, 1996). In that scenario (cf. Lamieff, 2019), specialists in psychological testing who have been retained as consultants by a trucking firm have recommended that an individual, 'Lesley,' not receive further consideration for a job as truck driver. This recommendation has been based on (a) 'Lesley's' high score, X, on a test for the personality trait 'sociability,' and (b) statistical evidence, compiled through studies of large numbers of individuals, that highly sociable people have, on average, more accidents and hence are costlier to trucking firms, perhaps as a consequence of heightened distractibility, than are their less sociable counterparts. Clearly, exploiting knowledge of that aggregate statistical relationship in guiding employment decisions could facilitate the efforts of the trucking firm to minimize its operating costs in the long run. Doing this would also gain, for the consultants, compensation for acquiring and maintaining the technical expertise necessary in order to secure and then apply the requisite knowledge. In short, the practice of guiding employment decisions about individuals by the knowledge of statistical relationships discovered in studies of aggregates might, indeed, be said to be "working" to the financial advantage of both the employer and the consultants.

However: what about Lesley? She has been eliminated from further consideration for employment as a truck driver due to a high 'costliness' score predicted for her on the basis of (a) her high score on a test of the personality trait 'sociability' and (b) a population-level correlation between sociability scores and 'costliness.' Suppose that Lesley, or someone on her behalf, would inquire of the consultants about the accuracy of that prediction? If the consultants' answer to Lesley's question would maintain fidelity to the logical limits of the statistical knowledge on which the prediction about her has been based, the consultants would concede that although they could properly claim knowledge of *average* predictive inaccuracy across a large number of cases, they could not properly claim any knowledge at all of what the individual predictive inaccuracy might prove to be in Lesley's (or any other) particular case.⁴

In this light, it would be clear to Lesley that she, as an individual, had been eliminated from consideration for employment on the basis of a *prediction* about her future performance the accuracy of which was completely unknown! This would certainly give Lesley ample reason to conclude that the practice of basing employment decisions about individuals on the basis of aggregate statistical knowledge, a practice that might, indeed, be said to 'be working' very well for a business and for the consultants retained by that business, was not 'working' particularly well for her.⁵

Whether or not the interpretive practices of psychologists (and many other social scientists) may be regarded as 'working' despite their conceptual flaws depends greatly on whose interests are being taken into consideration. They will often prove to 'be working' very well for those whose interests are not with the affected individuals to begin with but are instead served best by minimizing their own predictive errors in the long run—i.e., on average, for a set or cohort of individuals. However, the same practices will often not be judged to 'be working' so well by affected individuals, whose interests are not encompassed by actuarial considerations (cf. Hanson, 1993).

From psychological science to psycho-demography and ... back again?

As explained above, there is no viable conceptual bridge from aggregate statistical knowledge of variables with respect to which individuals have been differentiated to knowledge of any given individual among the differentiated, and this is true whether individuals are being differentiated by test scores reflecting differences that have arisen outside

the investigative doings of researchers or by treatments that have created the differences inside the laboratory. Contrary to the hyperbolic claim of Hofstee (2007), this argument is not equivalent to "denying the right of existence" (p. 253) to research on between persons differences, whether correlational, experimental, or both. The argument does, however, force recognition of the fact that such research effectively constitutes a species of *demography* (Lamiell, 2018b; cf. Proctor and Xiong, 2018; Lamiell, 2018c). This is so because the knowledge generated by such studies is knowledge of *populations*, whether of real populations with a recognized existence quite apart from the projects of social scientists, or of hypothetical populations suggested by experimental treatments. The knowledge generated by such studies is not knowledge of *any* individual within those populations, but is, in fact and quite literally, knowledge of *no one*.

I have suggested that research of this sort be designated 'psycho-demography.' This term recognizes that the variables selected or created for investigation often reflect between-person differences in psychological attributes or processes, but still identifies the work as essentially demographic in nature. By no means is this designation intended as a challenge to the scientific merit or practical usefulness of well-executed studies of between-person differences in psychological attributes. For purposes of guiding public policy decisions, for example, psycho-demographic knowledge is often exactly the sort of knowledge that is needed. Moreover, the findings issuing from psycho-demographic studies can often prompt hypotheses about psychological processes that could be tested subsequently in studies that, unlike psycho-demographic studies themselves, are formally suited to the task of gaining scientific knowledge of the psychological doings of individuals (see chapter by Grice, Hunjens, and Johnson, this volume).⁵

In short, psycho-demography can serve hugely important practical functions in its own right, and it can also serve as a *complement* to a scientific psychology oriented to the advancement of our scientific understanding of the psychological doings of individuals. Widespread recognition of this would effectively re-align the two disciplines relative to one another as they originally were when William Stern founded differential psychology as a complement to the general-experimental-individual psychology pioneered by Wilhelm Wundt and others. This development, were it to transpire, would nurture a broader and deeper understanding among mainstream psychologists (and, perhaps, among social scientists more widely) of the true nature and limits of population-level statistical knowledge. In the process, conceptual space would be re-opened for a genuine psychology, one that is actually capable of

advancing our scientific understanding of the psychological doings of individuals.

Notes

- 1 Hofstee (2007) has argued that this critique of conventional mainstream interpretive practices "counterfactually denies that people form expectations and make estimates, activities that are well represented by statistical models" (p. 253). The critique does nothing of the sort. The phenomena to which Hofstee (2007) referred are *psychological* in nature; features of the day-to-day cognitive functioning of lay persons well-documented in research conducted over many years by, most prominently, psychologists Amos Tversky (1937–1996) and Daniel Kahneman (b. 1934). The concern in the critiques that Hofstee (2007) cavalierly recommends be "taken with a grain of salt" (p. 253) concern the illogic of the reasoning reflected in the knowledge claims of trained scientists. The difference between these two concerns is apparent in the realization that there would be no grounds for labeling as 'biases' certain features of the day-to-day cognitive functioning of lay persons (cf. Tversky and Kahneman, 1974) were there no standard of logically sound reasoning against which to evaluate that functioning!
- 2 In an unpublished essay written much later in her career, titled *Neglected Insights in Personality*, Tyler (1984) clearly acknowledged that she had egregiously mischaracterized Stern's views (cf. Lamiell, 2003, pp. 1–2).
- 3 For a critical discussion of still another recent endorsement of this view, the reader is referred to pp. 222–225 of Lamiell and Martin (2017).
- 4 Were it otherwise, the consultants would simply correct their predictions by the putatively known error built into them, thereby eliminating their errors and making all of their predictions perfectly accurate!
- 5 Nor could the consultants validly finesse the issue here by claiming to know the 'probability' to be p that Lesley's 'costliness' score, were she to be hired, would be found to lie somewhere within the specified upper and lower limits of a so-called 'confidence interval' along the scale of 'costliness.' In this respect, all that the consultants could validly claim to know is that out of N cases with 'sociability' scores identical to Lesley's, ($p * N$) of those cases would manifest 'costliness' scores falling within the designated interval and ((1 – p))* N would not. The question of which of those two groups would be found to contain Lesley's 'costliness' score is a question on which the consultants' aggregate statistical data are completely silent.
- 6 For a detailed illustration of these points, see the discussion by Lamiell (2019, pp. 151–156) of work by Johnson, Markowitz, Hill, and Phillips (2016).

References

- Anastasi, A. (1937). *Differential psychology: Individual and group differences in behavior*. New York: Macmillan.

- Bakem, D. (1955). The general and the aggregate: A methodological distinction. *Perceptual and Motor Skills*, 5, 211-212.
- Bakem, D. (1966). The test of significance in psychological research. *Psychological Bulletin*, 66, 423-437.
- Banicki, K. (2018). Psychology, conceptual confusion, and disquieting situationalism: Response to Lamiell. *Theory and Psychology*, 28, 255-260. doi: 10.1177/0959354318759609.
- Cohen, J. (1968). Multiple regression as a general data analytic system. *Psychological Bulletin*, 70, 292-303.
- Cowles, M. (1989). *Statistics in psychology: A historical perspective*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Conbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, 12, 671-684.
- Danziger, K. (1987). Statistical method and the historical development of research practice in American psychology. In L. Krueger, G. Gigerenzer, & M. S. Morgan (Eds.), *The probabilistic revolution, Vol. 2: Ideas in the sciences* (pp. 35-47). Cambridge, MA: MIT Press.
- Danziger, K. (1990). *Constructing the subject: Historical origins of psychological research*. Cambridge: Cambridge University Press.
- Grice, James W., Hungjens, R., & Johnson, H. (this volume). Persistent disregard for the inadequacies of null hypothesis significance testing and the viable alternative of observation-oriented modeling. In Lamiell, J. T., & Shaney, K. L. (Eds.), *Scientific psychology's problematic research practices and inertia: History, sources and recommended solutions*. Abingdon-on-Thames: Routledge, forthcoming in 2021.
- Hanson, F. A. (1993). *Testing testing: Social consequences of the examined life*. Berkeley: University of California Press.
- Harré, R. (2006). *Key thinkers in psychology*. Thousand Oaks, CA: Sage Publications.
- Hogan, R., Hogan, J., & Roberts, B. W. (1996). Personality measurement and employment decisions: Questions and answers. *American Psychologist*, 51, 469-477.
- Hofstee, W. K. B. (2007). *Unbehagen in individual differences—A review. Journal of Individual Differences*, 28, 252-253. <https://doi.org/10.1027/1614-0001.28.4.252>.
- Isen, A. M., & Levin, P. F. (1972). Effect of feeling good on helping: Cookies and kindness. *Journal of Personality and Social Psychology*, 21, 384-388.
- Johnson, A. D., Markowitz, A. I., Hill, C. J., & Phillips, D. A. (2016). Variation in impacts of Tulsa pre-K on cognitive development in kindergarten: The role of instructional support. *Developmental Psychology*, 52, 2145-2158.
- Kerlinger, F. N. (1979). *Behavioral research: A conceptual approach*. New York: Holt, Rinehart, & Winston.
- Kerlinger, F. N. (1986). *Foundations of behavioral research*. 3rd edition. Chicago: Holt, Rinehart and Winston.
- Kerlinger, F. N., & Pedhazur, E. J. (1974). *Multiple regression in behavioral research*. New York: Holt, Rinehart, and Winston.
- Lamiell, J. T. (1981). Toward an idiographic psychology of personality. *American Psychologist*, 36, 276-289.
- Lamiell, J. T. (1987). *The psychology of personality: An epistemological inquiry*. New York: Columbia University Press.
- Lamiell, J. T. (1997). Individuals and the differences between them. In R. Hogan, J. A. Johnson & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 117-141). New York: Academic Press.
- Lamiell, J. T. (2003). *Beyond individual and group differences: Human individuality, scientific psychology, and William Stern's critical personalism*. Thousand Oaks, CA: Sage Publications.
- Lamiell, J. T. (2015). Statistical thinking in psychological research: In quest of clarity through historical inquiry and conceptual analysis. In Martin, J., Sugarman, J., & Shaney, K. L. (Eds.), *The Wiley handbook of theoretical and philosophical psychology: Methods, approaches, and new directions for social sciences* (pp. 200-215). Hoboken, NJ: John Wiley & Sons.
- Lamiell, J. T. (2016). On the concept of 'effects' in contemporary psychological experimentation: A case study in the need for conceptual clarity and discursive precision. In Harré, R., & Moghaddam, F. (Eds.), *Questioning causality: Scientific explorations of cause and consequence across social contexts* (pp. 83-102). Santa Barbara, CA: Praeger.
- Lamiell, J. T. (2018a). On the concepts of character and personality: Correctly interpreting the statistical evidences putatively relevant to the disposition-situation debate. *Theory and Psychology*, 28, 249-254. doi: 10.1177/095935431774837.
- Lamiell, J. T. (2018b). From psychology to psycho-demography: How the adoption of population-level statistical methods transformed psychological science. *American Journal of Psychology*, 131, 489-492.
- Lamiell, J. T. (2018c). Rejoinder to Proctor and Xiong. *American Journal of Psychology*, 131, 489-492.
- Lamiell, J. T. (2019). *Psychology's misuse of statistics and persistent dismissal of its critics*. London: Palgrave Macmillan.
- Lamiell, J. T., & Martin, J. (2017). The in-correctible science. In H. Macdonald, D. Goodman, & B. Decker (Eds.), *Dialogues at the edge of American psychological discourse: Critical and theoretical Perspectives in Psychology* (pp. 211-244). London: Palgrave-Macmillan.
- Münsterberg, H. (1913). *Psychology and industrial efficiency*. Boston and New York: Houghton Mifflin.
- Porter, T. M. (1986). *The rise of statistical thinking: 1820-1900*. Princeton, NJ: Princeton University Press.
- Proctor, R. W., & Xiong, A. (2018). Adoption of population-level statistical methods did transform psychological science, but for the better: Commentary on Lamiell (2018). *American Journal of Psychology*, 131, 483-487.
- Stern, W. (1900). *Über Psychologie der individuellen Differenzen (Haben zu einer "differentialen Psychologie")* [On the psychology of individual differences (Toward a "differential psychology")]. Leipzig: Barth.

- Stern, W. (1911). *Die Differentielle Psychologie in ihren methodischen Grundlagen* (Methodological foundations of differential psychology). Leipzig: Barth.
- Thorndike, E. L. (1911). *Individuality*. New York: Houghton Mifflin.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124-1131.
- Tyler, L. (1947). *The psychology of human differences*. New York: Appleton-Century-Crofts.
- Tyler, L. (1984). *Some neglected insights in personology*. Unpublished manuscript, Eugene, OR.
- Venn, J. (1888). *The logic of chance*. London and New York: Macmillan.
- Wandt, W. (1912). *Elemente der Völkerpsychologie*. Leipzig: Alfred Kröner Verlag.
- Wandt, W. (1913/2013). Psychology's struggle for existence (J. T. Larnieil, translator). *History of Psychology*, 16, 197-211. doi: 10.1037/a0032319.