Available online at www.sciencedirect.com
ScienceDirect
Personality and Individual Differences 44 (2008) 1652–1662
PERSONALITY AND INDIVIDUAL DIFFERENCES
www.elsevier.com/locate/paid

Schizotypy, taxometrics, and disconfirming theories in soft science
Comment on Rawlings, Williams, Haslam, and Claridge

Theodore P. Beauchaine [a,*], Mark F. Lenzenweger [b], Niels G. Waller [c]

[a] University of Washington, P.O. Box 351525, Seattle, WA 98195-1525, United States
[b] State University of New York at Binghamton, United States
[c] University of Minnesota, United States

Received 24 July 2007; received in revised form 17 November 2007; accepted 22 November 2007
Available online 3 January 2008

Abstract

Among the most consistently replicated findings in literature addressing the latent structure of psychopathology is evidence for a discrete latent class of individuals who are vulnerable to schizophrenia (schizotypes). Rawlings, Williams, Haslam, and Claridge (2008) challenge these findings by subjecting schizotypy scale scores to taxometric analysis, using a data simulation technique to accommodate variable skew. The authors conclude that schizotypy reflects a latent dimension, and that evidence for discrete latent structure from previous studies is likely an artifact of skewed variables. In this comment, we discuss (a) the philosophical implications of disconfirming well replicated findings in soft science with a single study, (b) important considerations when defining the schizotypy construct, (c) intricacies in executing and interpreting a taxometric analysis, and (d) problems with drawing strong conclusions from null results. Considerable evidence suggests that schizotypy is a discrete latent class, a conclusion that is unlikely the result of skewed variables.

Keywords: Schizotypy; Schizophrenia; Taxometrics; Philosophy of science

* Corresponding author. Tel.: +1 206 543 8844.
E-mail address: tbeaucha@u.washington.edu (T.P. Beauchaine).

0191-8869/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.paid.2007.11.015

1. Introduction

Few findings are indisputable in psychological science. Unlike the 'hard' sciences including physics, chemistry, and geology, psychology has almost no laws or axioms, and few networks of widely agreed upon facts. With the exception of certain subdisciplines such as psychobiology, little has changed in this regard since Kuhn (1962) described psychology as pre-paradigmatic nearly five decades ago.

Among the biggest differences between the hard and soft sciences concerns the role that prediction plays in disconfirming theories. In the hard sciences, precise experiments that disagree with predictions can overturn grand theories.[1] This is how Einstein's theory of special relativity supplanted Newton's universal law of gravitation. Highly precise observations of light deflecting through a strong gravitational field agreed more with Einstein's predictions than with Newton's.

In contrast, single experiments rarely if ever disconfirm theories in psychological science (Meehl, 1967). One reason for this is that psychologists are often interested in constructs such as motivation, personality, and psychopathology—the causes of which cannot be observed directly. These constructs must be inferred from behaviors, which are affected by many intervening influences. This results in low measurement precision compared with variables of interest to, for example, physicists (e.g., force, distance, mass). Due in part to reduced measurement precision, psychological theories generate few specific predictions, and are judged more by their consistency with broad networks of observations across multiple levels of analysis (e.g., biological, psychological, social). Thus, social scientists derive support for theories primarily from replication.

One of the most widely replicated findings in literature addressing the latent structure of psychopathology is evidence for a discrete latent class of individuals who carry a liability for developing schizophrenia—often referred to as schizotypy (Lenzenweger, 2006; Meehl, 1962, 1990). Although discussion of the theoretical bases for asserting a discontinuity in schizophrenia liability is beyond the scope of this comment, descriptions can be found in Meehl (1962, 1990), Lenzenweger (2006), and Lenzenweger, McLachlan, and Rubin (2007).

Schizotypy is a latent construct that is not directly observable and is therefore inferred from manifest indicators. Members of the latent schizotypy class exhibit incipient signs of vulnerability, including (a) behaviors such as social withdrawal and flat affect (Tyrka, Cannon, et al., 1995; Tyrka, Haslam, & Cannon, 1995); (b) unusual sensory experiences such as perceptual aberration, magical ideation, and referential thinking (Lenzenweger, 1999; Lenzenweger & Korfine, 1992); (c) specific patterns of item responses on objective psychological tests (Golden & Meehl, 1979); and (d) endophenotypic indicators of risk such as compromised neuromotor performance and eye-tracking dysfunction (Erlenmeyer-Kimling, Golden, & Cornblatt, 1989; Grove, Clementz, Iacono, & Katsanis, 1992; Lenzenweger et al., 2007). At each of these levels of analysis, schizotypy has emerged as a discrete latent class in previous research.

Few questions in psychological science receive this level of scrutiny, and few findings are so well replicated. Evidence for the schizotypy taxon has emerged from independent research groups using data collected across many levels of analysis. This provides strong support for a discrete

[1] We are not suggesting that unexpected results from precise experiments always overturn established theories, even in the hard sciences. In some cases, auxiliary conjectures are challenged and modified—with few or no implications for central tenets of the theory. Complete discussion of these issues is beyond the scope of this comment.

model of schizophrenia liability (Meehl, 1990). According to Meehl's (1962) original theory, genetic vulnerability to schizophrenia—which he referred to as *schizotaxia*—was transmitted through a single gene that was considered a necessary but insufficient cause of schizophrenia, resulting in the disorder for only a subset of susceptible people. It is now recognized that schizophrenia is either (a) a polygenically-determined disorder, or (b) a disorder caused by one of several genetic polymorphisms (for more detailed discussions of the genetics of schizophrenia, see Gottesman & Gould, 2003; McClellan, Susser, & King, 2007). Yet regardless of the specific genetic mechanisms, Meehl's assertion that schizotypy comprises a discrete latent class is supported by the studies cited above.

Schizotypy afflicts roughly 10% of the population, a base rate identified repeatedly in taxometrics studies using representative data collection strategies (e.g., Lenzenweger, 1999; Lenzenweger & Korfine, 1992). Moreover, as expected from taxonic models of schizotypy (Lenzenweger, 2006; Meehl, 1962, 1990), this base rate increases substantially—to nearly 50%—among the offspring of a parent with schizophrenia (Erlenmeyer-Kimling et al., 1989; Tyrka, Cannon, et al., 1995).

Rawlings, Williams, Haslam, and Claridge (2008) acknowledge the consistency of previous findings. However, they note that dimensional models of schizotypy also exist. In such models, schizotypy is viewed as a component of normal personality (Claridge, 1997) rather than an expression of schizophrenia liability. Claridge's dimensional view derives from Eysenck's "psychoticism" construct, which is largely an amalgam of aggression and impulsivity. This conceptualization of schizotypy is somewhat different than traditional definitions, and may therefore have less relevance for schizophrenia.

Given somewhat different definitions of schizotypy, dimensional variation in Claridge's construct is not necessarily incompatible with taxonic models of schizophrenia liability. It is possible that schizotypy as defined by Claridge varies dimensionally within the population whereas schizotypy defined more traditionally captures a discrete latent class at high risk for schizophrenia. Yet rather than considering this possibility, Rawlings et al. suggest that limitations to the taxometric method are responsible for spurious taxonic results in previous studies. In their article, they present taxometric analyses that are consistent with a dimensional model of schizotypy, and they suggest that the current consensus of schizotypy as a discrete latent class be re-evaluated.

In our view, this call to re-evaluate the consensus regarding schizotypy as a discrete latent trait based on the results of one study—which we argue suffers from some important methodological and analytic limitations—strikes us as premature, perhaps reflecting an alternative interpretation of schizotypy as normal personality variation. We are concerned that Rawlings et al. are mistaken in their conclusion, for reasons articulated below.

2. Sampling and assessment

2.1. Sampling

The manner in which data are collected is of great relevance to taxometric investigations. Meehl (1992) was clear about the effects that environmental factors could have on shaping data subjected to taxometric analysis and, importantly, on the results of the analyses themselves. Unfortunately, Rawlings et al. provide little information on their data collection method.

We do know that their sample of 1073 adults included student and health professionals in courses at the universities of Oxford and Liverpool, student and community members of the Oxford subject pool—and in the case of the largest subset of 408 participants—respondents to a media appeal for a study of out-of-the-body experiences. There are two potentially competing biases in this recruitment strategy. First, student and community members from Oxford are likely selected for attributes—including achievement and social class—that filter out individuals who score high on most psychopathological traits including, we presume, schizotypy. Second, out-of-body experiences reflect dissociation, which may itself mark a latent taxon that is distinct from schizotypy (Waller, Putnam, & Carlson, 1996; Waller & Ross, 1997). Because schizotypes do experience dissociation, including such items is not necessarily problematic. However, dissociation is also a core symptom of other psychiatric disorders (e.g., borderline pathology, dissociative identity disorder). Thus, if the advertisements included information about the objectives of the study—which the authors fail to specify—individuals with other psychiatric syndromes may have been over-sampled. This latter sampling bias could be especially insidious for efforts to identify a schizotypy taxon. Admixing samples with different characteristics produces significant problems for taxometric analyses (see Beauchaine, 2003).

In addition, it is likely that the different recruitment streams resulted in non-representative sampling (with associated volunteer biases), and that each stream yielded sub-samples of males and females with sub-sample-specific means and *SD*s (see below). The authors provide no information on the response rate (which obscures any effort to evaluate the representativeness of their sampling), nor do they discuss important differences across the various recruitment streams (e.g., Oxford undergraduates vs. "out-of-body" experience believers). Thus, one is left to wonder how these samples differ, and the extent to which convenience sampling yielded data suitable for taxometric analysis. Ideally, representative samples drawn from well defined populations should be used (Meehl, 1995).

2.2. Assessment

The Chapman Scales (Chapman, Chapman, & Kwapil, 1995) are excellent measures with established validity and reliability (see Waller, in press, for a discussion of reliability in taxonic samples). However, the manner in which Rawlings et al. collected data for the schizotypy assessments requires careful scrutiny. Based on their description, they constructed an omnibus screening questionnaire that consisted of 420 schizotypy-relevant items, drawn from 18 different scales. Subsequent analyses focused on just 4 of these scales.

Sex differences often emerge from the Chapman scales, so it is routine to standardize samples separately by sex prior to using scale data for participant selection or taxometric analysis (Lenzenweger, 1999). Analyses of schizotypy scales that are not standardized by sex run the nontrivial risk of inducing sex-linked artifact into the data. In the Rawlings et al. study, this problem may be compounded by the use of alternative data streams (see above). In other words, it appears that they combined their data across both sex and recruitment streams without regard to standardization. It is impossible to determine what the joint effects of these data reduction decisions might be. One possibility is that sex differences, if large enough, could induce a spurious latent taxon. In the current study, this is clearly not the case. Alternatively, the joint effects of sampling biases and sex

differences could produce a sex × recruitment stream artifact that was large enough to swamp detection of a schizotypy taxon.

3. Executing a taxometric analysis

Before we discuss specifics of the Rawlings et al. analyses, it is important to acknowledge how complicated the taxometric method can be. Comprehensive coverage of the approach is therefore well beyond the scope of this comment. We refer interested readers elsewhere for such descriptions (Beauchaine, in press; Waller & Meehl, 1998).

3.1. Precision

Two of the most fundamental issues for any method concern the validity and precision of data that are subjected to analysis. Variables that measure a construct badly can produce misleading outcomes, and variables that measure a construct imprecisely can lead to underpowered analyses that cannot detect real effects.

Precision is often confused with validity in psychological science, yet the terms have different meanings. Validity refers to the ability of a set of (obtained) scores to (a) predict a specific criterion or class of criteria or (b) measure an underlying construct adequately. In contrast, precision refers to the amount of error contained in a measure. Although rating scale data are often sufficiently valid indicators of the constructs they are intended to assess, these scales yield some of the least precise measures available for statistical analysis. This situation, depicted in Fig. 1, can be problematic for any method, including taxometrics, which requires large effect sizes to uncover discrete latent structure (Meehl, 1995).

One means of reducing measurement error is to include variables from levels of analysis other than rating scales. For example, objective measures of motor dysfunction (e.g., the Maher line-drawing task; Lenzenweger & Maher, 2002) are more precise than a neurologist's ratings on a 1–5 scale. Yet despite recommendations to use more precise measures including endophenotypes when performing taxometric analyses (e.g. Beauchaine, 2003), most researchers continue to rely exclusively on rating scale data (both self- and informant-report). This is an important consideration when evaluating the Rawlings et al. findings because they compare their results—which are derived from rating scales—with results reported by other researchers who used more precise measures of schizotypy (e.g., eye-tracking dysfunction, sustained attention, neuromotor performance; see Erlenmeyer-Kimling et al., 1989; Lenzenweger et al., 2007). Rating scales are also susceptible to response biases and halo effects, which can alter the latent structure of data (Beauchaine & Waters, 2003). Without acknowledging the limitations of rating scale data, this direct comparison implies that precision is irrelevant in taxometrics. Yet refutation of a consistently replicated finding by an experiment of lower measurement precision is almost unthinkable in the hard sciences.

3.2. Rejection of low base rate responses

Even more troubling, Rawlings et al. omitted all items on the Combined Schizotypal Traits Questionnaire (Bentall, Claridge, & Slade, 1989) that were endorsed by fewer than 10% of partic-

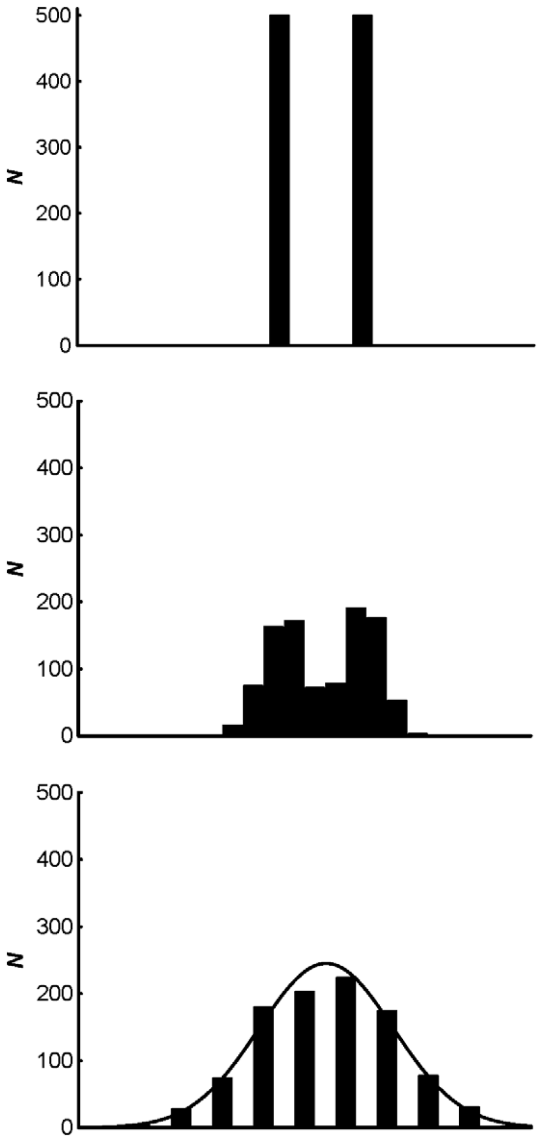

Fig. 1. The effect of measurement imprecision (error) on the distribution of a discretely distributed variable. Top panel: perfect measurement of a dichotomous trait of equal proportions (50:50) within an overall sample of $n = 1000$. This near-perfect precision would be associated with genotyping a trait with a single locus. Middle panel: extremely precise measurement of behavioral symptoms. Because symptoms are far removed from their genetic and neurobiological substrates, individual differences are always observed. Bottom panel: additional 50% increase in variance associated with imprecise rating scale measurement (see Hoyt & Kerns, 1999). The bottom panel is unimodal despite the admixture of two discrete groups. All figures have equivalent *x*-axis scaling. From Beauchaine (in press).

ipants. This practice runs counter to the test construction approach adopted by the Chapmans, where the guiding principle was identification of a relatively rare entity (the schizotype; see Chapman et al., 1995). As noted above, the base rate of schizotypy in the general population is very

near .10. By eliminating items that were endorsed by fewer than 10% of respondents, the authors may have discarded the most valid indicators of the schizotypy taxon. The importance of this point cannot be overstated. Omitting these items is likely to have reduced the authors' chances of identifying a latent taxon considerably. In our view, this data censoring decision alone may have invalidated the study.

Finally, Rawlings et al. note that most taxometric analyses of schizotypy have been conducted with MAXCOV. They then suggest that by using MAXEIG, their study represents an improvement over previous reports. However, Rawlings et al. subject only three variables to each of their MAXEIG analyses. In the absence of nuisance covariance, the mathematics of MAXEIG and MAXCOV are identical when three variables are used.

3.3. Simulations and variable skew

Extensive Monte Carlo simulations indicate that MAXCOV is effective in distinguishing discrete from dimensional latent structure with skew values as high as 2.0—even when the taxon base rate is as low as .05 (e.g., Beauchaine & Beauchaine, submitted for publication; Cleland & Haslam, 1996). Nevertheless, the effects of variable skew have received particular attention as a possible source of spurious taxa (Ruscio & Ruscio, 2002; Ruscio, Ruscio, & Keane, 2004).

Initially, Ruscio and Ruscio (2002) noted that skewed data can produce right-end peaked MAMBAC, MAXCOV, and MAXEIG plots that suggest spurious low base rate taxa. They used taxometric procedures to analyze self-report depression data collected from a sample of 2260 undergraduates. These analyses suggested a low base rate taxon, which the authors attributed to variable skew. In the same paper, they introduced a simulation procedure aimed at distinguishing between true low base rate taxa and spurious taxa derived from skewed variables.

Using this technique, both latent dimensional and latent taxonic data are simulated to match (as closely as possible) observed data parameters (e.g., univariate means, *SD*s, skew). Taxometric analyses are then conducted on all three datasets (observed, simulated dimensional, simulated taxonic), and inferences regarding latent structure are drawn based on whether results from the observed data are more similar to results from the simulated dimensional or the simulated taxonic data. If analyses from the simulated dimensional data are similar to analyses of the observed data, it is assumed that variable skew produced a spurious result.

In principle, this is a sensible approach to protecting against spurious low base rate taxa that are induced by variable skew. However, although the method has been used in several recent studies, its validity and operating characteristics have been scrutinized in only one published Monte Carlo report (Ruscio, Ruscio, & Meron, 2007), with no replications by independent labs. Thus, it is not fully clear whether the procedure is effective in preventing spurious findings, or whether it goes too far by invalidating legitimate taxonic inferences, as some have suggested (Beach, Amir, & Bau, 2005). Moreover, although offered originally as a means of accommodating variable skew (Ruscio & Ruscio, 2002), the simulation technique is now used as a matter of course in many studies. Yet until the method is evaluated by independent labs, it should probably be used cautiously, especially in discrediting well replicated findings (see Beauchaine, in press).

Although space constraints preclude a full discussion of data simulation techniques, additional limitations may also be relevant. As noted by a reviewer, when using simulation procedures one first fabricates nontaxonic data that mimic taxonic structure as closely as possible. These

simulated data are therefore not representative of the universe of datasets expected to occur if a taxon were not present. Rather, they are constructed to be as likely as possible to "fool" taxometric methods. This is one reason why simulation procedures may sometimes fail to detect existing taxa (see Beach et al., 2005).

In accordance with the original rationale for the Ruscio and Ruscio (2002) simulation technique, Rawlings et al. use it to accommodate skewed variables. Yet the method was probably unnecessary for this purpose. As noted above, taxometric procedures are effective in distinguishing discrete from dimensional data at skew values as high as 2.0. Although simulation procedures may be needed when skew exceeds this level, only the perceptual aberration variable was even moderately skewed (1.79). All other skew values fell between .78 and .87—acceptable for virtually any multivariate procedure.

Using the simulation technique is also unnecessary when the taxon base rate is moderate to high—a situation that does not yield skew-induced spurious taxa. In the Rawlings et al. study, MAXEIG analyses of the magical ideation, social anhedonia, and physical anhedonia scales yielded base rate estimates of .25, .27, and .28, respectively. With base rates this high, simulations to accommodate low base rate skewed indicators are not needed.

Finally, findings from studies of samples with high base rates of schizotypy challenge the Rawlings et al. assertion that previous taxonic findings result from skewed indicators. Among the studies cited above indicating a schizotypy taxon, two were conducted with children of a parent with schizophrenia (Erlenmeyer-Kimling et al., 1989; Tyrka, Cannon, et al., 1995). These studies identified schizotypy taxa with base rates of .47 and .48. Such findings present a significant problem for the authors' conclusion that schizotypy is dimensional, because (a) the increased prevalence rate is exactly what one would expect among individuals selected for genetic vulnerability, and (b) the results cannot be attributed to an artifact-induced spurious low base rate taxon. Consistent with these findings, Grove et al. (1992) reported evidence for a major gene underlying eye-tracking dysfunction—a well-validated schizotypy endophenotype—in both individuals with schizophrenia and their relatives. Major gene effects are taxonic by definition.

4. Over-interpreting null results

It has become common to advance strong arguments for the dimensionality of constructs when no taxon is found. This is what Rawlings et al. do by concluding—despite significant evidence to the contrary from previous research—that schizotypy is dimensional. Such inferences are problematic from a philosophy of science standpoint because taxometric procedures search for disjunctions in the structure of data as evidence of discrete latent classes (Beauchaine, 2003, in press). They do not test the hypothesis that individual differences are dimensional. Rather, when taxometric procedures fail to identify a taxon, continuous latent structure is inferred based on the *absence* of evidence to the contrary.

In this regard, there is a parallel between conclusions drawn from taxometric analyses and conclusions drawn from null hypothesis significance testing (NHST). When there is inadequate separation of the null distribution (complement class) from the alternative distribution (taxon), no effect will be found even if one exists. Although the analogy between NHST and taxometrics is not complete (the null hypothesis is always false given perfect measurement precision (Meehl,

1967), whereas hypothesis of dimensional latent structure is not), in both cases an inference is made based on the *absence* of an observed effect. Concluding that a negative result from a taxometric procedure provides strong evidence for continuous latent structure is therefore similar to proving the null, and we should be cautious about doing so for the same philosophical reasons.

The formal philosophical problem with making such inferences is instantiated in the logical fallacy *denying the antecedent*. Consider the following three statements: (1) If I own a Ferrari, then I am rich. (2) I do not own a Ferrari. (3) Therefore I am not rich. In this example, even if we assume that the first two premises are *always* true, the conclusion does not necessarily follow because not all rich people own Ferraris. Nevertheless, in many circumstances this logical fallacy seems to be valid. In the case at hand, when we infer continuous latent structure from MAXEIG (or any other taxometric procedure), the following assumptions apply: (1) If there is a latent taxon, then a discontinuity in Eigenvalues will be observed; (2) there is no observed discontinuity in Eigenvalues; (3) therefore, there is no taxon. As in the above example, this conclusion does not necessarily follow from the preceding premises. There are many reasons why real taxa can go undetected (Beauchaine & Beauchaine, 2002), including but not limited to sampling biases, invalid indicators, and small effect sizes, to name but three. Thus, unless great care is taken toward obtaining representative samples, selecting valid and precise indicators, and properly implementing taxometric procedures, it is always possible that a taxon could be detected using more precise methods.

It is incumbent on researchers using NHST to specify the power of their test statistics. Yet in taxometrics, power has been ignored almost completely. Despite existing knowledge that taxometric procedures require large effects to detect latent taxa (Beauchaine & Beauchaine, 2002; Meehl, 1995), researchers often subject any and all variables to analysis, with no priority assigned to more precise measures that are likely to yield greater power. We should therefore be *very* cautious in concluding that a construct represents a continuum until we are sure the taxonic hypothesis has been put to a strong test by using precise indicators. This is rarely done, and it was not done by Rawlings et al.

5. Concluding remarks

Based on their taxometric analysis, Rawlings et al. assert that schizotypy is a dimensional construct rather than a discrete index of schizophrenia liability. This conclusion challenges one of the most widely replicated findings in research addressing the latent structure of psychopathology—that schizotypy is taxonic. The authors state that "findings from the present study imply that taxonic inferences about the latent structure of schizotypy may have been mistaken, despite being consistently replicated in previous research" (p. qq). They go on to suggest that variable skew is likely to have induced spurious taxa in prior studies.

In this comment, we discussed the philosophical implications of disconfirming well replicated findings based on a single null result, and we presented a number of complexities—including alternative definitions of the schizotypy construct, recruitment of participants, construction of schizotypy scales, and execution of taxometric analyses—that likely account for the disparity between previous reports and the Rawlings et al. findings. It is important to re-emphasize that among the previous reports that Rawlings et al. challenge are studies that used more representative sampling techniques and more precise measures of schizotypy. Thus, this is not merely a failure to

replicate. Rather, it is a null finding derived from data and data analyses that in several important ways fall below standards set in previous reports. Papers that discredit the work of others should be held to at least the same standard as the work they call into question. Anything less would be unacceptable in the hard sciences. It is time that we adopt a similar criterion for refuting theories in psychology.

References

Beach, S. R. H., Amir, N., & Bau, J. (2005). Can sample specific simulations help detect low base rate taxonicity? *Psychological Assessment, 17*, 446–461.

Beauchaine, T. P. (2003). Taxometrics and developmental psychopathology. *Development and Psychopathology, 15*, 501–527.

Beauchaine, T. P. (in press). A brief taxometrics primer. *Journal of Clinical Child and Adolescent Psychology*.

Beauchaine, T. P., & Beauchaine, R. J. III, (2002). A comparison of maximum covariance and *k*-means cluster analysis in classifying cases into known taxon groups. *Psychological Methods, 7*, 245–261.

Beauchaine, T. P., & Beauchaine, R. J., III (submitted for publication). A comparison of measures of fit for inferring taxonic structure from maximum covariance analyses.

Beauchaine, T. P., & Waters, E. (2003). Pseudotaxonocity in MAMBAC and MAXCOV analyses of rating scale data: Turning continua into classes by manipulating observer's expectations. *Psychological Methods, 8*, 3–15.

Bentall, R. P., Claridge, G. S., & Slade, P. D. (1989). The multidimensional nature of schizotypal traits: A factor-analytic study with normal subjects. *British Journal of Clinical Psychology, 28*, 363–375.

Chapman, J. P., Chapman, L. J., & Kwapil, T. R. (1995). Scales for the measurement of schizotypy. In A. Raine, T. Lencz, & S. Mednick (Eds.), *Schizotypal personality* (pp. 79–106). New York: Cambridge University Press.

Claridge, G. (1997). Theoretical background and issues. In G. Claridge (Ed.), *Schizotypy: Implications for illness and health* (pp. 3–18). Oxford: Oxford University Press.

Cleland, C., & Haslam, N. (1996). Robustness of taxometric analyses with skewed indicators: I. A Monte Carlo study of the MAMBAC procedure. *Psychological Reports, 79*, 243–248.

Erlenmeyer-Kimling, L., Golden, R. R., & Cornblatt, B. A. (1989). A taxometric analysis of cognitive and neuromotor variables in children at risk for schizophrenia. *Journal of Abnormal Psychology, 98*, 203–208.

Golden, R. R., & Meehl, P. E. (1979). Detection of the schizoid taxon with MMPI indicators. *Journal of Abnormal Psychology, 88*, 217–233.

Gottesman, I. I., & Gould, T. D. (2003). The endophenotype concept in psychiatry: Etymology and strategic intentions. *American Journal of Psychiatry, 160*, 636–645.

Grove, W. M., Clementz, B. A., Iacono, W. G., & Katsanis, J. (1992). Smooth pursuit ocular motor dysfunction in schizophrenia: Evidence for a major gene. *American Journal of Psychiatry, 149*, 1362–1368.

Hoyt, W. T., & Kerns, M. D. (1999). Magnitude and moderators of bias in observer ratings: A meta-analysis. *Psychological Methods, 4*, 403–424.

Kuhn, T. (1962). *The structure of scientific revolutions*. Chicago: University of Chicago Press.

Lenzenweger, M. F. (1999). Deeper into the schizotypy taxon: On the robust nature of maximum covariance analysis. *Journal of Abnormal Psychology, 108*, 182–187.

Lenzenweger, M. F. (2006). Schizotaxia, schizotypy and schizophrenia: Paul E. Meehl's blueprint for experimental psychopathology and the genetics of schizophrenia. *Journal of Abnormal Psychology, 115*, 195–200.

Lenzenweger, M. F., & Korfine, L. (1992). Confirming the latent structure of schizotypy: A taxometric analysis. *Journal of Abnormal Psychology, 101*, 567–571.

Lenzenweger, M. F., & Maher, B. A. (2002). Psychometric schizotypy and motor performance. *Journal of Abnormal Psychology, 111*, 546–555.

Lenzenweger, M. F., McLachlan, G., & Rubin, D. B. (2007). Resolving the latent structure of schizophrenia endophenotypes using expectation-maximization-based finite mixture modeling. *Journal of Abnormal Psychology, 116*, 16–29.

McClellan, J. M., Susser, E., & King, M.-C. (2007). Schizophrenia: A common disease caused by multiple rare alleles. *British Journal of Psychiatry, 190*, 194–199.

Meehl, P. E. (1962). Schizotaxia, schizotypy, schizophrenia. *American Psychologist, 17*, 827–838.

Meehl, P. E. (1967). Theory-testing in psychology and physics: A methodological paradox. *Philosophy of Science, 34*, 103–115.

Meehl, P. E. (1990). Toward an integrated theory of schizotaxia, schizotypy, and schizophrenia. *Journal of Personality Disorders, 4*, 1–99.

Meehl, P. E. (1992). Factors and taxa, traits and types, differences of degree and differences in kind. *Journal of Personality, 60*, 117–174.

Meehl, P. E. (1995). Bootstraps taxometrics: Solving the classification problem in psychopathology. *American Psychologist, 50*, 266–275.

Rawlings, D., Williams, B., Haslam, N., & Claridge, G. (2008). Taxometric analysis supports a dimensional latent structure for schizotypy. *Personality and Individual Differences*, *44*, 1640–1651.

Ruscio, A. M., & Ruscio, J. (2002). The latent structure of analogue depression: Should the Beck Depression Inventory be used to classify groups? *Psychological Assessment, 14*, 135–145.

Ruscio, J., Ruscio, A. M., & Keane, T. M. (2004). Using taxometric analysis to distinguish a small latent taxon from a latent dimension with positively skewed indicators: The case of involuntary defeat syndrome. *Journal of Abnormal Psychology, 113*, 145–154.

Ruscio, J., Ruscio, A. M., & Meron, M. (2007). Applying the bootstrap to taxometrics: Generating empirical sampling distributions to help interpret results. *Multivariate Behavioral Research, 42*, 349–386.

Tyrka, A. R., Cannon, T. D., Haslam, N., Mednick, S. A., Schulsinger, F., Schulsinger, H., et al. (1995). The latent structure of schizotypy: I. Premorbid indicators of a taxon in individuals at risk for schizophrenia-spectrum disorders. *Journal of Abnormal Psychology, 104*, 173–183.

Tyrka, A., Haslam, N., & Cannon, T. D. (1995). Detection of a latent taxon of individuals at risk for schizophrenia spectrum disorders. In A. Raine, T. Lencz, & S. A. Mednick (Eds.), *Schizotypal personality* (pp. 168–191). New York: Cambridge University Press.

Waller, N. G. (in press). Commingled Samples: A neglected source of bias in reliability analysis. *Applied Psychological Measurement*.

Waller, N. G., & Meehl, P. E. (1998). *Multivariate taxometric procedures: Distinguishing types from continua*. Newbury Park, CA: Sage.

Waller, N. G., Putnam, F. W., & Carlson, E. B. (1996). Types of dissociation and dissociative types: A taxometric analysis of dissociative experiences. *Psychological Methods, 1*, 300–321.

Waller, N. G., & Ross, C. A. (1997). The prevalence and biometric structure of pathological dissociation in the general population: Taxometric and behavior genetic findings. *Journal of Abnormal Psychology, 106*, 499–510.