

Rushton's Defenders and Their Hasty Rejection of the Null Hypothesis

Zack Z. Cernovsky

University of Western Ontario

Rosenthal and Rubin (1985) pointed out that in research on extreme situations (e.g., new treatments for terminally ill patients) any noticeable statistical trend in the desirable direction is valuable. It should be published even if it is of low magnitude and fails to meet our traditional criteria of statistical significance. Their approach is now being misused by those defending Rushton's (1988) "theory" about American Blacks (based on weak trends in excessively suspect data sets). Hasty and eager acceptance of weak, biased, and unrepresentative data as scientific evidence of genetically based and relatively immutable racial differences in human potential amounts to psychological warfare on oppressed racial groups. Similar defamation of vulnerable minorities by Nazi pseudoscientists led to the loss of millions of human lives in the past. Statistical theory classifies similar endeavors as a Type I error (a misleading rejection of the null hypothesis).

Rushton (1988, 1990) has repeatedly claimed that Blacks are genetically less intelligent and more prone to crime and mental disease than Whites. His "theory" is artificially supported by a blatantly biased review of literature and is based on obsolete methodology and a misrepresentation of statistical data (Cernovsky, 1992; Cernovsky & Litman, 1993a, 1993b; Flynn, 1989, 1990; Weizmann, Wiener, Wiesenthal, & Ziegler, 1991; Zuckerman & Brody, 1988). However, these criticisms have not prevented some writers within academic psychology from misinterpreting Rushton's work as a milestone of scientific progress and from imitating his procedures and extending his

AUTHOR'S NOTE: *This article is based on a paper presented at the 101st Annual Convention of the American Psychological Association in Toronto, Canada, August 20-24, 1993. Address correspondence to Dr. Z. Cernovsky, University of Western Ontario, St. Thomas Psychiatric Hospital, St. Thomas, Ontario, Canada N5P 3V9.*

JOURNAL OF BLACK PSYCHOLOGY, Vol. 20 No. 3, August 1994 325-333
© 1994 The Association of Black Psychologists

defamation of Blacks to similar misguided accounts of presumed psychological inferiority of native American Indians (Phelps, 1993) or women (Lynn, 1993).

Some of Rushton's procedures are statistically absurd. For example, Rushton used tape measures of human head circumference as an indicator of intelligence even though his own review of the literature showed that head size was too weakly related to scores on tests of intelligence to provide a viable substitute for IQ scores (an average Pearson r of .18). Rushton's overinterpretation of weak trends (e.g., Pearson r s < .20) in suspect data would, alone, provide a sufficient reason for labeling his work *pseudoscientific* (Cernovsky, 1992). In his search for suitable evidence, Rushton "selects" data sets consistent with his theory and fails to mention and include those in which Blacks were superior to Whites. This has been already documented with respect to his review of Tobias's (1970) data (see Cernovsky, 1992). Similar procedures (failure to report disconfirmatory data) are presently also used by other racists in academic circles, for example, by Arthur Jensen (see Kamin & Grant-Henry, 1987). These practices should not remain unchallenged, because they discredit psychology as a science and are tantamount to psychological warfare on oppressed racial groups.

In a paper presented at the 25th International Congress of Psychology in 1992, Cernovsky discussed Rushton's misinterpretations of extremely weak correlational trends and his misconceptions about "aggregation" techniques. This paper was subsequently published in the *Journal of Black Psychology* (see Cernovsky, 1993). In defense of Rushton, Thomas J. Bouchard of the University of Minnesota argued that during the question period following Cernovsky's paper at this congress, weak correlation coefficients should not be underestimated. He referred to Rosenthal's recent work on this issue. The present article discusses Rosenthal's theory and how it is commonly misinterpreted by Bouchard and others.

TYPE I AND TYPE II ERRORS

The underlying issue has a long history and is sketched in most textbooks of elementary statistics. For example, McCall (1986) explains that the purpose of statistical analyses of data is to make an educated guess about what exists in the population. The statistician's conclusions usually are, in fact, only a guess because normally only a small sample of persons from the population can be studied due to economic or other restraints. McCall points out that because the decision is a guess, it may be wrong. For instance, if the

significance level obtained in the study on therapeutic potential of a new drug is .05 and the investigators concluded that the drug has the desired expected effect, then there is a 5% chance that the researchers erroneously concluded that the drug is effective when it actually has no effect. In these 5% of studies, the null hypothesis (namely, the hypothesis that the drug has no effect) is erroneously rejected, and this mistake is known as the *Type I error*.

There could also be an error in the opposite direction. A new drug that actually has some therapeutic effects might erroneously be labeled as definitely noneffective after a few statistical studies with negative outcomes. This error (the null hypothesis is not rejected even though it should be rejected) is known as the *Type II error*.

In general, most social scientists are well aware of the dangers of Type I error, that is, of overinterpreting weak trends in their correlational or experimental data, but are less concerned about Type II error because strong statistical trends (e.g., those associated with Pearson r s $> .60$) are only rarely missed even in studies on small samples. The efforts to avoid Type II error by adopting less stringent criteria of statistical significance (e.g., .10 instead of .05) and by overinterpreting weak trends result in an increase in Type I error. The Type I errors can be costly. This is especially true when false findings are intensively publicized in mass media as a viable and solidly established scientific basis for social policies on the government level, as this is done by J. P. Rushton and his associates. The credibility of Rushton's work in the eyes of the lay public is artificially bolstered and protected by his academic status (fellow of the American Psychological Association [APA], full professor of psychology at the University of Western Ontario, etc.) as well as by an intensive support, academic promotion, and eloquent praise of his methodology by politically prominent figures in contemporary psychology such as Hans Eysenck, Thomas Bouchard, and Arthur Jensen.

Bona fide researchers normally do not and should not seek to lower our scientific criteria to willfully mislead the public. Criteria of statistical significance are almost never lowered except for unusual situations in which Type II error could result in an unnecessary threat to human safety or a loss of human lives. As shown in the following paragraphs, these exceptional and extreme cases are of an intrinsic interest to statisticians in psychology.

CONTRIBUTION BY ROSENTHAL AND RUBIN

In an article dealing with similar extreme situations, Rosenthal and Rubin (1985) brilliantly argued that small statistical trends, even those failing to

reach our traditional criteria of statistical significance ($p = .05$ or $.01$), can frequently be invaluable. They provide the following example:

Suppose that, of 20 critically ill patients in a small, randomized experiment, 10 are assigned to a treatment condition and the other 10 are assigned to a control condition. If none of the control patients survive and 3 experimental patients survive, our results will not be significant at $p \leq .05$ by a $\chi^2(1)$ test or a Fisher exact test. However, we believe it is essential on scientific as well as ethical grounds that such results should be published. (p. 528)

Suppose that in the example provided by Rosenthal and Rubin only one patient in the treatment group but none in the control condition survived. The majority of us, if critically ill, would still opt for the treatment, that is, for the chance of survival. In Rosenthal's example, Type I error (use of a noneffective drug to "treat" a patient that is going to die very soon with or without any known treatment) has only minor negative consequences compared to Type II error (failure to use treatment that could save 30% of the critically ill patients). In contrast, Type I error, underlying Rushton's speculations, leads to the defamation of Blacks, waste of their human potential, a promotion of racial hatred, and to unnecessary loss of life in racist mob activities.

In this century, Type I error in psychology already has had devastating social consequences. For example, erroneous and highly biased reports, by prominent North American psychologists, of a high incidence of low intelligence in Jewish and some other European immigrants to the United States led to the government policy of extremely restrictive immigration quotas for these groups. In subsequent decades, these quotas greatly obstructed the escape routes of Jewish and East-European refugees from persecution before, during, and following World War II. Thus this psychological war on non-Anglo-Saxon immigrants by prominent American psychologists had been responsible in part for unnecessary death of millions of Jews and of East-European dissidents in German Nazi or Stalin's death camps. It is noteworthy that, in our decade, "psychologists" in high academic positions (e.g., Thomas Bouchard) are still willing to reject or compromise scientific standards for the sake of derogating and victimizing vulnerable oppressed groups.

The risk of Type II error (missing out on a new important finding), in the case of Rushton's theory, is nonexistent. The theory has not been developed from an empirical basis and is lacking both in conceptual originality and in logical structure. There is no carefully assembled unbiased evidence. Rushton's archaic methodology with an excessive reliance on outdated cranial measurements and on old and highly suspect data sets is also invali-

dated by his failure to report or include disconfirmatory data. His conclusion that the presumed racial differences are genetic has no even remotely adequate basis in the methodology and data. The issue is not, as in Rosenthal and Rubin's model, whether a weak trend in methodologically satisfactory data is sufficient evidence to establish a temporary treatment strategy to provide at least some hope (and psychological relief) for critically ill patients who would otherwise certainly die. Rosenthal and Rubin's model involves a presumably unbiased data set and presumably unbiased scientific procedures. Rushton's data are methodologically worthless even before the issue of significance criteria is discussed. His blatant personal bias greatly increases the probability of Type I errors.

TYPE I ERRORS AND SOCIAL POLICIES

Pseudoscientific defamation of vulnerable groups in Rushton's, Jensen's, and Bouchard's style, combined with their political activism, establishes a nefarious precedent and a model for future generations of psychologists. It teaches both our White and Black college students that it might be relatively safe and beneficial for an academic career (including upward mobility in APA circles) to derogate oppressed groups in pseudoscientific papers. If the current population trend continues, the children of today's academic racists can themselves become a minority and victims of racial hatred.

Rosenthal and Rubin's example of life threatening situations can be reversed to illustrate its potential misuse. Instead of serving as an invaluable guide to promote hope and health, weak trends in the data can be misused to justify a systematic abuse of selected vulnerable groups. One day, the vulnerable group could be the Whites. Suppose that, in a predominantly Black hypothetical country of the future (e.g., the United States if the current population trend continues) where most university educated professionals are Black, a study on a random sample of 25 Blacks and 25 Whites has found that Blacks performed somewhat better in higher management positions and were less prone to crime than Whites. Let us assume that the results are significant at $p = .05$ by a t test. Given this statistically significant difference, would it be ethical to interpret it as an established fact on public media (in Rushton's style) urging the (predominantly Black) government to modify social policies on the basis of this new research on racial differences? Would it be ethical to assert that these differences are genetic and relatively immutable (as does Rushton, Bouchard, and Jensen on the basis of less adequate data)? Obviously, methodologically weak or blatantly biased data are too

prone to Type I error to serve as a basis for theories of genetic racial superiority of any group over any other group.

Or, suppose that in the same predominantly Black country and time a study on a randomly selected sample of 25 Black and 25 White psychologists found that Blacks performed somewhat better as psychotherapists and as teachers of psychology and were less prone to illicit sexual affairs with their clients than their 25 White counterparts. Again, let us assume that our results are significant at $p = .05$ by the t test. Given this statistically significant difference, would it be ethical to pretend that this is sufficiently solid evidence of immutable genetic differences, and/or to use similar data as a scientific justification of repressive administrative policies against individuals identified by their skin color? Of course, these examples of discrimination are only hypothetical. The reverse is presently true: Blacks are being treated unfairly on the basis of poor data.

RESEARCH METHODS AND TYPE I ERROR

Systematically biased sampling procedures increase the likelihood of Type I error. There is a host of data from reviews of past research that could be misused, by Rushton's and Jensen's procedures of discarding disconfirmatory data, to produce a pseudo-solid empirical evidence of racial superiority of almost any racial group over other groups. For example, Tobias's (1970) review of data on racial differences in cranial size and in the number of excess neurons includes tables of average values from a wide variety of data sets. Rushton reported only those data from Tobias that were less favorable for Blacks even though Tobias's article also contains statistical data that could be used by some writers as evidence for intellectual superiority of American Blacks over various Caucasian groups, including American Whites, the French, and the English.

Similarly, Rushton (1990) presented selected evidence from Interpol archives, based on worldwide statistics of criminal activities, for his theory of greater genetic proclivity of Blacks than Whites toward crime. The same data source has been subsequently found to contain impressive statistical data in the opposite direction (Cernovsky & Litman, 1993b). These data could, one day, be publicized by some to proclaim a "genetic" racial superiority of Blacks over Whites. The use of "selected" statistical evidence, especially if consisting of conspicuously weak trends, for grandiose theories of genetic superiority of any racial group over other groups is a trademark of charlatanism and deserves no place in academic circles.

EDITORIAL POLICIES AND TYPE I ERROR

Editorial policies that prevent or partly obstruct publication of negative or disconfirmatory findings have devastating effects on subsequent attempts to statistically summarize and pool results of different studies to evaluate the overall trend in the data as this is attempted in meta-analytic techniques. This is also true about nonnegative findings that are dissonant with "intuitive beliefs" of the editor (or of those that provide financial support for the journal). These circumstances increase the Type I error in a direction unfavorable for Blacks. Until recently, Black students were systematically prevented from entering U.S. universities, and those who persisted frequently paid for this with their lives. So, until recently, most research on racial differences has been carried out by White psychologists and published by White editors. As known with respect to Rushton and Jensen, many White investigators (and editors) hesitate (and often fail) to publish not only negative findings but also findings disadvantageous for their ethnic group. These distortions are further compounded by the impact of an experimenter's expectations on his or her "data." The investigator's errors are only partly a random phenomenon. Most of these errors happen to be in the direction of his or her hypothesis. According to statistical studies reported in Rosenthal (1991) on this phenomenon, about two thirds of the observational errors that are made by the investigator are in the direction of the hypothesis. These biased errors occasionally push a result over the magic .05 cliff.

The combination of these various sources of systematic racial bias greatly facilitates Rushton's task of reviewing past research to locate suitable evidence that Blacks and Whites statistically differ in a direction advantageous for his racial group. The onus is on all psychologists to actively prevent similar bias against any ethnic group and to help to correct its effects.

SAMPLE SIZE AND INTERPRETATION OF TRENDS IN THE DATA

Even with extremely weak data, Rushton may reach the .05 by resorting to his "aggregation procedure" (see Cernovsky, 1993). This procedure has been presented by Rushton as a panacea for social research. It consists in increasing sample size to the extent that an absurdly weak trend meets the .05 or even the .001 level. As explained in introductory textbooks to statistics, the level of statistical significance depends only partly on the magnitude of the trend in the data and can be easily reached, with very weak relationships,

when using samples with extremely large numbers of cases (see e.g., the discussion in McCall, 1986). Expert statisticians explicitly warn students about the dangers involved in interpreting excessively weak trends that reach our traditional criteria of significance only with very large samples (Witte, 1993).

In summary, the likelihood of Type I error in Rushton's work appears to be greatly increased by a combination of numerous factors. It is of interest for social psychology why the most blatantly archaic methodology, as in Rushton's use of head measures by tape as measures of intelligence, can still be promoted and defended, without much censure, by those who are entrusted by our profession with educating future scientists.

In the past, scientific revolutions have occasionally been spurred by original theories that were initially unsupported by adequate methodology and data. Their redeeming quality consisted in their new and innovative conceptual perspectives. Rushton's theoretical work does not fall in this category. Far from being innovative and original, Rushton's theory has faithfully followed old-fashioned racist stereotypes about innate racial differences in behavior. In the style of a highly prejudiced lay neo-Nazi, he has grandiosely avoided admitting the existence of numerous alternative explanations of his data. His work has been financially supported by the Pioneer Fund, that is, an institution known for its Nazi history and neo-Nazi propaganda. The social cost of academically legitimizing Rushton, his followers, and similar authors by artificially promoting them to prominent ranks in academic psychology and of uncritically publicizing their outmoded racist misconceptions as a "scientific theory" is potentially enormous and likely, in the long run, to lower the credibility of psychology as a science.

REFERENCES

- Cernovsky, Z. Z. (1992). J. P. Rushton on Negroids and Caucasoids: Statistical concepts and disconfirmatory evidence. *International Journal of Dynamic Assessment and Instruction*, 2, 55-67.
- Cernovsky, Z. Z. (1993). J. P. Rushton's aggregational errors in racial psychology. *Journal of Black Psychology*, 19, 282-289.
- Cernovsky, Z. Z., & Litman, L. C. (1993a). Re-analyses of J. P. Rushton's crime data. *Canadian Journal of Criminology*, 35, 31-36.
- Cernovsky, Z. Z., & Litman, L. C. (1993b, June). *Interpol crime statistics, race, and Rushton's racial conclusions*. Paper presented at the 19th International Congress of the International Academy of Law and Mental Health, Lisbon, Portugal.
- Flynn, J. R. (1989). Rushton, evolution, and race: An essay on intelligence and virtue. *The Psychologist: Bulletin of the British Psychological Society*, 9, 363-366.

- Flynn, J. R. (1990). Explanation, evaluation, and a rejoinder to Rushton. *The Psychologist: Bulletin of the British Psychological Society*, 5, 199-200.
- Kamin, L. J., & Grant-Henry, S. (1987). Reaction time, race, and racism. *Intelligence*, 11, 299-304.
- Lynn, R. (1993). Further evidence for the existence of race and sex differences in cranial capacity. *Social Behavior and Personality*, 21, 89-92.
- McCall, R. B. (1986). *Fundamental statistics for behavioral sciences* (4th ed.). New York: Harcourt Brace Jovanovich.
- Phelps, M. T. (1993). An examination of Lynn's evolutionary account of racial differences in intelligence. *Mankind Quarterly*, 33, 295-307.
- Rosenthal, R. (1991). *Meta-analytic procedures for social sciences*. Newbury Park, CA: Sage.
- Rosenthal, R., & Rubin, D. B. (1985). Statistical analysis: Summarizing evidence versus establishing facts. *Psychological Bulletin*, 97, 527-529.
- Rushton, J. P. (1988). Race differences in behavior: A review and evolutionary analysis. *Personality and Individual Differences*, 9, 1009-1024.
- Rushton, J. P. (1990). Race and crime: A reply to Roberts and Gabor. *Canadian Journal of Criminology*, 32, 315-334.
- Tobias, P. V. (1970). Brain size, grey matter, and race: Fact or fiction? *American Journal of Physical Anthropology*, 32, 3-25.
- Weizmann, F., Wiener, N. I., Wiesenhal, D. L., & Ziegler, M. (1991). Eggs, eggplants and eggheads: A rejoinder to Rushton. *Canadian Psychology*, 32, 43-50.
- Witte, R. S. (1993). *Statistics* (4th ed.). New York: Harcourt Brace Jovanovich.
- Zuckerman, M., & Brody, N. (1988). Oysters, rabbits, and people: A critique of "race differences in behavior" by J. P. Rushton. *Personality and Individual Differences*, 9, 1025-1033.