

Who is on First?

David M. Frederick, University of Colorado at Boulder
William H. Kaempfer, University of Colorado at Boulder
Richard Wobbekind, University of Colorado at Boulder

Research/statistical methodology
Abstract 2009-068

May 29, 2009
8:30 AM

25 minute oral
(Lexington B)

To do any research in social relations, one must be able to measure the constructs under study (Kidder 1981). The procedures used measuring constructs or characteristics of people must all meet the same requirements of measurement. They must be reliable and valid. Reliability can be broadly defined as the degree to which measures are free from error. Validity refers to the degree to which measures are accurate representations of the characteristic they are intended to measure. One of the major topics to dominate research in the business of baseball has been the examination of the extent of racial discrimination (Kahn 1992). Studies have examined discrimination with regard to salary, length of career, all-star voting, umpire performance, and even the market for baseball memorabilia (e.g., Parsons, Sulaeman, Yates and Hamermesh 2007; Depken and Ford 2005; Hanssen 2001; Gabriel, Johnson and Stanton 1999; Anderson and LaCroix 1991; Nardinelli and Simon 1990).

In spite of the numbers of studies, researchers have nevertheless failed to reach a uniform conclusion as to the presence or absence of discrimination in baseball. Two difficulties arise when measuring the extent of discrimination in baseball. One is the problem of measuring productivity. The other is in measuring race and ethnicity. An advantage of studying the concept of discrimination in baseball is that there is more publicly available data on performance and compensation of players than any other sport. Thus, the problem of measuring productivity is lessened. Yet, while great effort has been expended in attempting to obtain unbiased estimates of productivity in baseball discrimination studies, little effort appears to have been expended in attempting to obtain unbiased estimates of race and ethnicity. Few, if any research papers, even mention in a footnote, much less in the main body of text, how race and ethnicity was determined. Improperly classifying individuals based on race and ethnicity will produce invalid results, no matter how much care has gone into measuring productivity. This study has two goals. The first is to document the variance inherent in trying to categorize baseball players by race and ethnicity. The conflicting and inconclusive results of prior studies suggest the nontrivial nature of this classification task.

Undertaking experiments using pictures and/or names of players, we test for uniformity of response on race classification. These results suggest the degree of variance that might be present in the race variables of previous studies and may help to reinterpret previous results. The second goal is to initiate the development a more formal approach for understanding the concept that is being examined. Specifically, is it the cultural identity of the players that is being studied or is it the perceived identity? This study should be viewed as the first step in a series of research projects that will move from the artificial to the more realistic view of race and ethnicity. Ultimately, the results will lead to a better understanding of the extent of discrimination in baseball, other sports, and non-sports endeavors.