



**The Socio-cultural Context of
Research in STEM Education
Or
Power, The Power of the Norm
and Research**

Patricia Campbell, PhD
Campbell-Kibler Associates, Inc.
www.Campbell-Kibler.com
www.FairerScience.org

Researchers and Objectivity



Researchers are not immune to the "disease and superstition of American racism" (Thomas and Sillen, 1972).

Scientific objectivity quickly becomes subjective under the influence of intense feelings. It is not easy... to prevent our likes, aversions, hopes and fears from coloring our conclusions (Nagel, 1961).

As soon as any strong passion intervenes to warp the expert's judgment, he [sic] becomes unreliable, whatever scientific equipment he may possess (Russell, 1959).

The history of scientific views on race serves as a mirror of social movement; reflecting good times and bad; periods of the belief in equality and of rampant racism. Changes in research findings on intellectual inferiority reflect changes in society, with biological determinism rising in times of political retrenchment (Gould, 1981).

A Little Bit of History



- A 1885 Psychological Review article concluded that whites' slower reaction time (compared to Blacks and Native Americans) was proof that whites were the superior group (Gossett, 1963).
- A 1966 study of self-esteem found high Black self-esteem; the authors concluded, without evidence, that it was "in reality a defense mechanism against discrimination" (McDill et al., 1966).
- Acknowledging that "intellect can be defined and measured in such a manner as to make either sex appear superior." The 1942 revision of the Stanford Binet IQ test had as its aim to "produce a scale which will yield comparable I.Q.s for the sexes." (McNemar, 1942).

What and How Questions Are Asked



Significant diversity reduces the power of any dominant group to say no that is not worth looking at.

In their 1964 research, Bettelheim and Janowitz asked “Are Black people as intelligent as white people? not “Are white people as intelligent as Black people?” or even “Are our intelligence tests equally valid for whites and Blacks??

In 2007 Clewell and Campbell’s finding that overall achievement in some inner city schools serving poor, minority students is comparable to schools serving middle income white students is greeted with surprise .

How Studies Are Done



Significant diversity reduces the power of any dominant group to say that is not the right approach.

Designs

“The Right Design for the Question” vs “Hierarchy of Study Designs for Evaluating a STEM Educational Intervention.”

Samples

“Appropriate levels of disaggregation” vs “White as normative.”

Measures

“Better a bad measure of the right thing than a good measure of the wrong thing” vs “Since the 1950s verbal IQ tests in English have been found not to be good measures of Spanish dominant or bilingual children (Altus, 1953); they continue to be used.”

Thinking About Design: The Right Design for the Question



Study Type	Design	Advantages	Disadvantages
Ethno- graphy	Observer examination of group behaviors and patterns	Explores complex effects over time	Resource intensive Story telling approach may limit audience Potential observer bias
Case Study	Exploration of a case (or multiple cases) over time	Provides an in-depth view	Limited generalizability
Content Analysis	Systematic identification of properties of large amounts of textual information	Allows for quantitative and qualitative analysis	Tends too often to simply consist of word counts Can disregard the context that produced the text
Mixed Methods Study	Use of more than one of the above designs	Can counteract the disadvantages of any one design	None

Hierarchy of Study Designs For Evaluating a STEM Educational Intervention



Generally the strongest study design for evaluating an intervention's effectiveness. Uniquely, it enables one to determine to a high degree of confidence whether the intervention itself, as opposed to other factors, causes the observed outcomes.

**Well-designed
Randomized
Controlled Trial**

A second-best alternative when a randomized controlled trial is not feasible. The evidence suggests that if the intervention and comparison groups are very closely matched in key characteristics (e.g., pre-intervention educational achievement, demographics), the study in many cases yields the correct overall conclusion about whether the intervention is effective, ineffective, or harmful. However, its estimate of the size of the intervention's effect is often inaccurate, possibly resulting in misleading conclusions about the intervention's policy or practical significance.

**Well-matched
Comparison-Group Study**

Other designs, such as Pre-Post Study, and Comparison-Group Study without careful matching

Can be useful in generating hypotheses about what works that merit confirmation in more rigorous studies, but should not be relied upon to inform policy decisions, as they often produce erroneous conclusions about an intervention's effectiveness. This is true even when statistical techniques (such as regression adjustment) are used to correct for factors other than the intervention that may affect the

Guidelines To Reduce Bias In Research: A Blast From The Past



Does the research question devalue or stigmatize [groups being studied]
(Psychology)

Is an observed difference assumed to reflect a problem or pathology of
[groups being studied]? (Psychology)

Do not automatically assume instruments used successful in prior
studies of one ethnic group can yield valid information when applied
to other ethnic groups. (Social Work)

Are conclusions directly referenced to the results of the study?
(Education)

Are nonstereotypic/nontraditional explanations of results included?
(Education)

Cultural Context & Educational Evaluation



Cultural awareness of the environment from which the participants are drawn must be emphasized.

Evaluations must recognize that the culture of learners influences how they respond to the assessment process and assessment items.

Non-minority evaluators should be trained to evaluate programs that target minority learners.

Cultural awareness of the environment from which the participants are drawn must be emphasized.

Test results must be reported with context data.

Disaggregation of program data should include, as appropriate, factors such as, but not limited to, race, gender, socioeconomic status and opportunity to learn.

Working With and Changing the Socio-cultural Context of Research



Diversify the mainstream:

- volunteer to review for NSF, USED, professional organizations and journals
- when submitting papers for presentations to race/ethnicity/gender SIGS and divisions, suggest other possible co-sponsors.

Work with professional organizations to update and enforce guidelines to reduce bias in research.

Adapt what evaluators are doing to educational research.

Together, figure out other things to do.