Checklist for Identifying Scientifically Based Research

Consumers of research must be able to evaluate the research they read. Unfortunately, much of the research in education and related fields is contradictory, and an effort to find the empirical answer to a question of practice can be quite frustrating. Even in peer-refereed journals, much of the research is not good research. Many times, sample sizes are small and poorly selected, inappropriate statistics are used or mistakes are made in calculating and reporting statistics, or research designs are weak. For all of these reasons, much research is of marginal quality, and it becomes necessary for consumers to evaluate research results. No research is perfect. When we evaluate research we need to weigh its strengths and weaknesses and decide whether the results are trustworthy, whether the treatment described really resulted in the observed behavioral change. The purpose of this checklist is to assist individuals and groups when they evaluate the results of research studies.

Do Research Conclusions Align with the Problem Studied?

_____ Yes _____ No

Caveats: Some researchers collect data or study a problem in ways that don't allow them to answer their research questions. Check the researchers conclusions against their introductions and methods section to ensure consistency.

Comments:

Are Data Analyses Rigorous?

_____ Yes _____ No

Caveats: Many times, standardized tests are used to judge whether or not an intervention was successful. However, the standardized tests may be used for purposes other than what they are designed for.

Comments:

Is the sample size large enough to ensure generalizability of the results to the overall population?

_____ Yes _____ No

Caveats: Research is often conducted using relatively small samples of students, or samples that do not include special populations, such as students with special needs, minority ethnic background, or low-SES families. Be wary if the research has not been replicated with larger groups or with the types of special populations you are working with.

Comments

Is the evidence statistically significant?

_____ Yes _____ No

Caveats: Statistical significances means that when program outcomes are compared by a proper test (e.g., t-test, analysis of variance), the difference in results is due to factors other than chance. P values of .05 or less are typically considered significant. View nonsignificant results skeptically.

Comments:

Was a control group included in the study?

_____ Yes _____ No

Caveats: Researchers should use empirical methods that draw upon observations or experiments, ideally in classrooms or school-based settings. One of the best indicators of systematic research methods is the used of a control group. The control group should be as similar as possible to the treatment group. Beware of "before and after" studies where groups are compared before and after receiving an intervention. If there was not a control group included, it is not possible to detect whether the differences between the before and after measures are due to the intervention or to outside factors.

Comments:

Were groups randomly assigned?

_____ Yes _____ No

Caveats: Participants in the study should be randomly assigned into either a treatment group- that receives the new program or teaching method- or a control group that does not.

Comments:

Was the study reported in a refereed journal?

_____ Yes _____ No

Caveats: Many national education journals write articles that may or ma y not cite research. These articles are largely based on opinion. Much of educational research is paid for by publishers, software companies, or other business that have a financial stake in the results. It is important that these studies be reviewed by outside parties through publication in an academic journal or through a similar peer-review process. Not all journals and periodicals are equal! Only journals that use a peer-review process by researchers who are highly regarded in the filed should be used.

Comments:

Does the research evidence have external validity?

_____ Yes _____ No

Caveats: External validity means that the results of the study can be reasonably generalized to real-life, educational settings. Are the people participating in the research study similar to those we might encounter in education? Are treatments used similar to those we would use in education? The more comparable the research context to the application context (i.e., schools), and the larger and more inclusive the sample, the more meaningful the evidence is likely to be.

Comments:

Were the measurement tools used in the study reliable and valid?

_____ Yes _____ No

Caveats: Does the study adequately define the measures used to determine effectiveness of the intervention? The measurements and observations used must assure that each evaluator is making observations the same way (i.e., operational definition of behavior), data based on different measurements and observations can be compared, and there is reason to believe that similar studies by the same or different researchers would yield consistent results. A good test is whether the results of the study have been replicated in other studies.

Comments:

Are the results valid?

_____ Yes _____ No

Caveats: Were the researchers' methods and observations accurate? Do the results seem supportable based on the methods the researchers used? Are the conclusions consistent with the results are do they go well beyond them?

Comments:

Did the research include strengths and weaknesses of the study?

_____ Yes _____ No

Caveats: The article should include limitations of the research along with the major strengths and weaknesses of the study.

Comments:

Has the intervention or theory been tested in real classrooms?

_____ Yes _____ No

Caveats: Beware of theories that have not been tested using experimental methods in real classrooms. Examples of such theories include Piaget and Developmental Psychology, Gardner's theory of multiple intelligences, interdisciplinary/integrated curriculum, and reading recovery.

Comments: