What Works Clearinghouse

INSTITUTE OF EDUCATION SCIENCES

Science February 2012

ThinkerTools

No studies of *ThinkerTools* that fall within the scope of the Science review protocol meet What Works Clearinghouse (WWC) evidence standards. The lack of studies meeting WWC evidence standards means that, at this time, the WWC is unable to draw any conclusions based on research about the effectiveness or ineffectiveness of *ThinkerTools* on middle school students. Additional research is needed to determine the effectiveness or ineffectiveness of this intervention.

Program Description[1]

ThinkerTools[2] is a computer-based program that aims to develop students' understanding of physics and scientific modeling. The program is composed of two curricula for middle school students, *ThinkerTools Inquiry* and *Model-Enhanced ThinkerTools*. *ThinkerTools Inquiry* allows students to explore the physics of motion and then asks them to apply that knowledge to solve real-world problems. In the *Model-Enhanced ThinkerTools* curriculum, students create computer models that express their own theories of force and motion.

During the program, students navigate through four environments of increasing complexity known as microworlds: motion in one dimension, motion in two dimensions, understanding continuous forces, and analyzing trajectories. Students can run simulations of moving objects and observe the effects of various forces such as impulse, gravity, and friction. Measurement tools enable students to calculate distance, time, and velocity. The social interaction in the *ThinkerTools* classroom is similar to that of an actual scientific community, using a cycle of inquiry. Research begins with a whole-class forum and is carried out in groups. These groups make predictions, do experiments, and share their findings.

Research[3]

The WWC identified four studies of *ThinkerTools* for middle school students that were published or released between 1990 and 2011.

- Two studies are within the scope of the Science review protocol but do not meet WWC evidence standards. These quasi-experimental studies did not establish that the comparison group was comparable to the treatment group prior to the start of the intervention.
- Two studies are out of the scope of the Science review protocol because they have an ineligible study design. For the purposes of this report, the studies did not use a valid comparison group, as authors compared *Model-Enhanced ThinkerTools* to the prior curriculum version, *ThinkerTools Inquiry*.

References

Studies that do not meet WWC evidence standards

White, B. Y. (1993). ThinkerTools: Causal models, conceptual change, and science education. *Cognition and Instruction, 10*(1), 1–100. The study does not meet WWC evidence standards because it uses a quasi-experimental design in which the analytic intervention and comparison groups are not shown to be equivalent.

White, B. Y., & Frederiksen, J. R. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and Instruction,* 16(1), 3–118. Study 2: QED. The study does not meet WWC evidence standards because it uses a quasi-experimental design in which the analytic intervention and comparison groups are not shown to be equivalent.

Studies that are ineligible for review using the Science Evidence Review Protocol

Schwarz, C. V., & White, B. Y. (2005). Metamodeling knowledge: Developing students' understanding of scientific modeling. *Cognition and Instruction, 23*(2), 165–205. The study is ineligible for review because it does not use a comparison group design or single-case design.

Additional source:

Schwarz, C. V. (1998). Developing students' understanding of scientific modeling (seventh-grade, science students, model-enhanced ThinkerTools curriculum). *Dissertation Abstracts International, 60*(03A), 521-694.

White, B. Y., & Frederiksen, J. R. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and Instruction, 16*(1), 3–118. Study 1: RCT. The study is ineligible for review because it does not use a comparison group design or single-case design.

Additional source:

White, B. Y., & Frederiksen, J. R. (2000). Technological tools and instructional approaches for making scientific inquiry accessible to all. In M. J. Jacobson & R. B. Kozma (Eds.), *Innovations in science and mathematics education* (pp. 321–359). Mahwah, NJ: Lawrence Erlbaum Associates.

Endnotes

[1] The descriptive information for this program was obtained from publicly available sources: the program's website (http://thinkertools.org/, downloaded October 2011) and White (1993). The WWC requests developers to review the program description sections for accuracy from their perspective. The program description was provided to the developer in October 2011; however the WWC received no response. Further verification of the accuracy of the descriptive information for this program is beyond the scope of this review. The literature search reflects documents publicly available by June 2011.

[2] This review focuses on *ThinkerTools Inquiry* and *Model-Enhanced ThinkerTools*, but it uses the general term *ThinkerTools* to encompass both programs.

[3] The studies in this report were reviewed using WWC Evidence Standards, Version 2.1, as described in the review protocol for the Science topic area. The evidence presented in this report is based on available research. Findings and conclusions may change as new research becomes available.

Recommended Citation

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2012, February). *Science intervention report: ThinkerTools*. Retrieved from http://whatworks.ed.gov.

Glossary of Terms

Attrition Attrition occurs when an outcome variable is not available for all participants initially assigned to the intervention and comparison groups. The WWC considers the total attrition rate and the difference in attrition rates across groups within a study.

Clustering adjustment If treatment assignment is made at a cluster level and the analysis is conducted at the student level, the WWC will adjust the statistical significance to account for this mismatch, if necessary.

Confounding factor A confounding factor is a component of a study that is completely aligned with one of the study conditions, making it impossible to separate how much of the observed effect was due to the intervention and how much was due to the factor.

Design The design of a study is the method by which intervention and comparison groups were assigned.

Domain A domain is a group of closely related outcomes.

Effect size The effect size is a measure of the magnitude of an effect. The WWC uses a standardized measure to facilitate comparisons across studies and outcomes.

Eligibility A study is eligible for review and inclusion in this report if it falls within the scope of the review protocol and uses either an experimental or matched comparison group design.

Equivalence A demonstration that the analysis sample groups are similar on observed characteristics defined in the review area protocol.

Extent of evidence An indication of how much evidence supports the findings. The criteria for the extent of evidence levels are given in the WWC Procedures and Standards Handbook (version 2.1).

Improvement index Along a percentile distribution of students, the improvement index represents the gain or loss of the average student due to the intervention. As the average student starts at the 50th percentile, the measure ranges from –50 to +50.

Multiple comparison adjustment When a study includes multiple outcomes or comparison groups, the WWC will adjust the statistical significance to account for the multiple comparisons, if necessary.

Quasi-experimental design (QED) A quasi-experimental design (QED) is a research design in which subjects are assigned to treatment and comparison groups through a process that is not random.

Randomized controlled trial (RCT) A randomized controlled trial (RCT) is an experiment in which investigators randomly assign eligible participants into treatment and comparison groups.

Rating of effectiveness The WWC rates the effects of an intervention in each domain based on the quality of the research design and the magnitude, statistical significance, and consistency in findings. The criteria for the ratings of effectiveness are given in the WWC Procedures and Standards Handbook (version 2.1).

Single-case design A research approach in which an outcome variable is measured repeatedly within and across different conditions that are defined by the presence or absence of an intervention.

Standard deviation The standard deviation of a measure shows how much variation exists across observations in the sample. A low standard deviation indicates that the observations in the sample tend to be very close to the mean; a high standard deviation indicates that the observations in the sample tend to be spread out over a large range of values.

Statistical significance Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups. The WWC labels a finding statistically significant if the likelihood that the difference is due to chance is less than 5% ($p < 0.05$).

Substantively important A substantively important finding is one that has an effect size of 0.25 or greater, regardless of statistical significance.

Please see the WWC Procedures and Standards Handbook (version 2.1) for additional details.