

WWC Review of the Report “Conceptualizing Astronomical Scale: Virtual Simulations on Handheld Tablet Computers Reverse Misconceptions”¹

The findings from this review do not reflect the full body of research evidence on the true-to-scale (TTS) display mode in the *Solar Walk* software application for the iPad.

What is this study about?

The study examined the effects of using the true-to-scale (TTS) display mode versus the orrery display mode in the iPad’s *Solar Walk* software application on learning concepts in astronomy. The *Solar Walk* application is intended to help students understand the order of planets and their orbits. The TTS display mode shows planets in accurate scale, while the orrery display mode exaggerates the sizes of planets relative to their orbits. The study examined the effects of each mode on students’ knowledge of the Earth’s place in the solar system.

Study authors recruited 152 students enrolled in math and science courses in a public high school in Massachusetts. Fourteen science classes were randomly assigned to two groups: group A, consisting of five classes with 63 students; and group B, consisting of nine classes with 89 students. Both groups used *Solar Walk* on iPads in their regular science classes for about 20 minutes a day for 2 days.

On the first day, group A used the TTS display mode, while group B used the orrery display mode. On the second day, the experimental conditions were reversed, with group A using the orrery display mode and group B using the TTS display mode.

Study authors assessed the effectiveness of the TTS display mode versus the orrery display mode by comparing student performance on a 10-minute multiple-choice science test administered immedi-

ately after students completed the *Solar Walk* activities each day. Two scores were analyzed from this assessment: the percent correct for general questions about the solar system (general-type items), and the percent correct for questions related to scale in the solar system (scale-type items).

Features of the iPad’s *Solar Walk* Application

Solar Walk (developed by Vito Technology), version 1.9.4, is a software application that enables users to navigate a simulated, realistic-looking 3D solar environment using a “pinch-to-zoom” interface on the iPad. There are two display modes: true-to-scale (TTS) and orrery. In the TTS display mode, planetary bodies are shown in accurate scale. The orrery display mode exaggerates the size of planetary bodies relative to the size of their orbits. *Solar Walk* is intended to help users learn about the structure of the solar system, including the order of the planets and the shape and size of their orbits.

What did the study find?

The authors provided the WWC with scores of the intervention and comparison groups after the first day of implementation, referred to as the 1-day analysis.² WWC calculations indicated no statistically significant differences between students who used *Solar Walk* in the TTS display mode and students who used it in the orrery display mode on either the general-type questions or the scale-type questions.³ However, the WWC characterized the effect of the TTS display mode as having a substantively important positive effect because the mean effect calculated by the WWC across both outcomes is equal to 0.25 standard deviations.

The results from the 2-day analysis, which used scores from the first and second days of implementation, do not meet WWC standards; therefore, they are not presented in this WWC report.

WWC Rating

The research described in this report meets WWC group design standards without reservations

The study is a well-executed randomized controlled trial that meets WWC standards for assessing impacts from a comparison of students who used the TTS display mode to students who used the orrery display mode after the first day of implementation. The authors' comparison of students after the second day of implementation does not meet WWC group design standards because all students had used both display modes. Therefore, differences in achievement cannot be attributed to the TTS display mode or the orrery mode alone.

Appendix A: Study details

Schneps, M. H., Ruel, J., Sonnert, G., Dussault, M., Griffin, M., & Sadler, P. M. (2014). Conceptualizing astronomical scale: Virtual simulations on handheld tablet computers reverse misconceptions. *Computers & Education, 70*, 269–280.

Setting	The study was conducted in a public high school in eastern Massachusetts. Seventy-five percent of the school's 855 students were White, 11% qualified for a lunch subsidy, and 6% did not speak English as their first language.
Study sample	The study authors randomly assigned and analyzed 152 high school students enrolled in math and science courses. Five classes with 63 students were in the intervention group, and nine classes with 89 students were in the comparison group. Among the students participating in the study, 40% were male, 13% received a lunch subsidy, and their average age was 15.2 years.
Intervention group	Students in classes assigned to the intervention group completed a 20-minute self-administered instructional protocol of <i>Solar Walk</i> activities on iPads during their science class. <i>Solar Walk</i> is a software application that enables users to navigate a simulated, realistic-looking 3D solar environment using a "pinch-to-zoom" interface. There are two display modes: TTS and orrery. In the TTS display mode, planetary bodies are shown in accurate scale; in the orrery display mode, the size of the planetary bodies is exaggerated according to the size of their orbits. Students in the intervention group used the TTS display mode on the first day of the intervention and the orrery display mode on the second day.
Comparison group	Students in classes assigned to the comparison condition completed the same <i>Solar Walk</i> activities in their science classes as the intervention group, but used the orrery display mode on the first day of the intervention and the TTS display mode on the second day.
Outcomes and measurement	Students were tested on astronomical concepts regarding the Earth's place in the solar system. A 16-item test instrument was developed by the researchers using the middle-school inventory of multiple-choice questions from the Astronomy and Space Science Concept Inventory (ASSCI). The test was designed to take about 10 minutes. The outcomes assessed were two subtest scores: the percentage correct out of nine general solar system items (general-type score) and the percentage correct out of seven items related to scale in the solar system (scale-type score). Students were assessed before and after the science classes on each day of implementation. The 1-day analysis used scores from the first day of implementation, and the 2-day analysis used scores from both days of implementation. Scores from the first day of implementation were provided by the study authors to the WWC in response to a WWC request for additional information. For a more detailed description of these outcome measures, see Appendix B.
Support for implementation	A protocol was written so that students would work individually; teachers provided only clarifying information during the experiment.
Reason for review	This study was identified for review by receiving media attention.

Appendix B: Outcome measures for the science achievement domain

Science achievement	
<i>General-type items</i>	A 16-item instrument that assessed astronomical concepts regarding the Earth's place in the solar system was developed by the researchers using the middle-school inventory of multiple-choice questions from the ASSCI. This outcome is the score on nine items from the 16-item ASSCI-based test related to general solar system concepts.
<i>Scale-type items</i>	This outcome is the score on seven items from the 16-item ASSCI-based test related to scale in the solar system.

Appendix C: Study findings for the science achievement domain

Domain and outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			p-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Science achievement								
<i>General-type items</i>	High school classes, first day of implementation	14 classes/152 students	35.62 (20.82)	33.58 (17.88)	2.04	0.11	+4	.71
<i>Scale-type items</i>	High school classes, first day of implementation	14 classes/152 students	38.63 (18.60)	31.99 (15.25)	6.64	0.40	+15	.17
Domain average for science achievement						0.25	+10	Not statistically significant

Table Notes: For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on student outcomes, representing the average change expected for all students who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average student’s percentile rank that can be expected if the student is given the intervention. The WWC-computed average effect size is a simple average rounded to two decimal places; the average improvement index is calculated from the average effect size. The statistical significance of the study’s domain average was determined by the WWC.

Study Notes: The findings presented here, based on outcomes measured at the end of the first day’s implementation, were not shown in the study report; they were obtained from the authors in response to a request for more information. The analysis reported in the study combines student gain scores on the general-type and scale-type items. Gains were computed as posttest minus pretest scores on a given day of implementation. The analyses reported in the article are not included in this WWC report because they use test scores from both days of implementation, and the 2-day analysis does not meet WWC standards.

A correction for clustering was needed and resulted in the WWC-computed p-values shown above. Therefore, the WWC does not find the result for either outcome to be statistically significant. The study report did not provide p-values for the 1-day results. The WWC calculated the intervention group mean using a difference-in-differences approach (see WWC Handbook) by adding the impact of the program (i.e., difference in mean gains between the intervention and comparison groups) to the unadjusted comparison group posttest means. Please see the WWC Procedures and Standards Handbook (version 3.0) for more information.

This study is characterized as having a substantively important positive effect because the mean effect reported is positive and not statistically significant but is substantively important. For more information, please refer to the WWC Standards and Procedures Handbook (version 3.0), p. 24.

Endnotes

¹ Single study reviews examine evidence published in a study (supplemented, if necessary, by information obtained directly from the authors) to assess whether the study design meets WWC design standards. The review reports the WWC's assessment of whether the study meets WWC design standards and summarizes the study findings following WWC conventions for reporting evidence on effectiveness. This study was reviewed using the WWC Evidence Review Protocol for Science Interventions (version 2.0). The WWC rating applies only to the study outcomes that were eligible for review under this topic area. The reported analyses in this single study review are only for those eligible outcomes that either met WWC design standards without reservations or met WWC design standards with reservations, and do not necessarily apply to all results presented in the study.

² The WWC requests additional information from authors when necessary to complete the study review. As described on page C.1 of the WWC Handbook (version 3.0), the information obtained from the authors is documented in the study's master review guide and available by request from the WWC.

³ Two outcomes included in the study are not described in this WWC report: student gain scores over both implementation days for the general-type and scale-type items. As noted in Appendix B, they are not included because the 2-day analysis does not meet WWC standards. Instead, the WWC used the pretest and posttest means and a difference-in-differences adjustment to compute effect sizes for percent correct on general- and scale-type items for the 1-day analysis.

Recommended Citation

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2014, September). *WWC review of the report: Conceptualizing astronomical scale: Virtual simulations on handheld tablet computers reverse misconceptions*. 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 intervention 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 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.
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 intervention 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 intervention and comparison groups.
Single-case design (SCD)	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 are 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% ($p < .05$).
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 \(version 3.0\)](#) for additional details.