What Works Clearinghouse™

ies INSTITUTE OF EDUCATION SCIENCES

April 2014

WWC Review of the Report "Benefits of Practicing 4 = 2 + 2: Nontraditional Problem Formats Facilitate Children's Understanding of Mathematical Equivalence"[1]

The findings from this review do not reflect the full body of research evidence on using nontraditional problem formats for addition practice.

What is this study about?

The study examined the effects of addition practice using nontraditional problem formats on students' understanding of mathematical equivalence. In nontraditional problem formats, operations appear on the right side of the equal sign (e.g., __ = 4 + 3) rather than the traditional method of placing the operations on the left side of the equal sign (e.g., 4 + 3 = __). The concept of mathematical equivalence, an understanding that the equal sign means the two sides of an equation are interchangeable, is important for success in mathematics.

Study authors randomly assigned 95 7- and 8-year-old students to one of three conditions:[2]

- *Nontraditional practice:* Students in this intervention condition received supplemental addition practice using nontraditional problem formats, where the addition operation occurred on the right side of the equation (e.g., __ = 4 + 3). The students received four one-on-one tutoring sessions to practice addition facts with games and flashcards, supplemented with two brief homework assignments. In total, this represented approximately 100 minutes of supplemental practice of addition facts.
- *Traditional practice:* Students in this comparison condition received the same amount and types of extra math practice in one-on-one sessions as the nontraditional practice group. However, this group only used traditional problem formats (e.g., 4 + 3 = __).
- *No extra practice:* Students in this comparison condition did not receive any additional practice time beyond what they received at school and home (that is, they did not receive the 100 additional minutes of practice that the other two groups received).

Study authors examined the impact of nontraditional practice, compared to traditional practice and no extra practice, on three measures of mathematical equivalence understanding administered immediately after the intervention: (a) equation solving, (b) equation encoding, and (c) defining the equal sign. The authors also administered the equation solving measure a second time as a 2-week follow-up, with feedback provided. Because the authors were concerned that nontraditional practice could potentially compromise computational fluency (speed, flexibility, and accuracy in solving computation problems) as negative side effects, the study also assessed the impact on five assessments of computational fluency (four of which were based on a single administration of a computer-based assessment of arithmetic skill); however, the analyses of computational fluency were presented as secondary outcomes and were not the focus of the study (see Appendix D for more information).[3]

Features of Nontraditional Addition Problem Formats

In nontraditional addition problem formats, operations appear on the right side of the equal sign (e.g., __ = 4 + 3) rather than the traditional method of placing the operations on the left side of the equal sign (e.g., 4 + 3 = __).

The widespread use of traditional addition problem formats creates a pattern where operations occur on the left side of an equation and the solution appears on the right side. This may lead students to interpret the equal sign as a signal to "do something" rather than an indication that the two sides of the equation are equal, making solving later math equivalence problems (e.g., 3 + 4 = __ + 2) more difficult. Exposing students to practice using nontraditional formats may improve their understanding of mathematical equivalence, facilitating the learning of more advanced concepts later on.

What did the study find?

The study authors reported, and the WWC confirmed, that using nontraditional addition problem formats had a statistically significant positive impact on three of four measures of students' understanding of mathematical equivalence, relative to using traditional addition problem formats. Specifically, there were significant positive impacts shown for the measures of equation solving, defining the equal sign, and equation solving with feedback.

WWC Rating

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

The comparison of the nontraditional practice group to the traditional practice group is a well-executed randomized controlled trial with low levels of sample attrition. The comparison of the nontraditional practice group to the no-extra-practice group does not meet WWC group design standards because it is a randomized controlled trial with high differential attrition and baseline equivalence was not established for the two groups.

Appendix A: Study details

McNeil, N. M., Fyfe, E. R., Petersen, L. A., Dunwiddie, A. E., & Brletic-Shipley, H. (2011). Benefits of practicing 4 = 2 + 2: Nontraditional problem formats facilitate children's understanding of mathematical equivalence. *Child Development, 82*(5), 1620–1633.

Setting The study took place in a midwestern city in the United States.

Study sample A total of 95 7- and 8-year-old students from an economically diverse set of public and private elementary schools were randomly assigned to one of three groups: (a) nontraditional practice, (b) traditional practice, and (c) no extra practice. The final analysis sample, after attrition, consisted of 90 students, 30 students in each group.

Overall, 53% of the final sample were male; 61% were White, 29% were African American, 9% were Hispanic or Latino, and 1% were Asian. Twenty-nine percent of the sample received free or reduced-price lunch. Study authors reported that all of the students were at or near grade level in mathematics performance. Grade level was not specified.

Intervention group The intervention group received nontraditional addition practice that involved presenting students with problems with the operations on the right side of the equal sign (e.g., __ = 4 + 3) rather than the traditional method of placing the operations on the left side of the equal sign (e.g., 4 + 3 = __). Students participated in four one-on-one practice sessions with a tutor. During the first three sessions, the students practiced addition by playing games with the tutor, answering problems on flashcards, and playing computer games. In between practice sessions, the students completed brief homework assignments for additional practice. Students were also given short worksheets of addition facts between the first and second sessions and between the second and third sessions. During the practice sessions, all addition problems were displayed with the operations on the right side of the equal sign.

The students received approximately 100 minutes of addition practice (including homework) as a result of participating in the intervention. During the fourth session, the students were given the assessments by a different experimenter than the one who had served as the tutor in the first three sessions.

Comparison group

There were two comparison conditions examined in this study:

a) *Traditional practice:* This group received traditional instruction in addition problems that involved presenting students with problems with the operations on the left side of the equal sign (e.g., 4 + 3 = __) rather than the nontraditional format described above. During the practice sessions, all addition problems were displayed with the operations on the left side of the equal sign. The activities and dosage were otherwise the same as they were for the intervention group.

b) *No extra practice:* This group did not receive any additional practice time beyond what they received at school and at home (that is, they did not receive the 100 additional minutes of practice that the intervention group and the other comparison group [traditional practice] received).

Analyses that contrasted outcomes of the nontraditional practice and no-extra-practice groups did not meet standards (there was high differential attrition, and the groups were not shown to be equivalent at baseline for all outcomes assessed). Analyses that contrasted the nontraditional and traditional practice groups do meet WWC standards, and therefore, this WWC report only presents results based on this contrast.[4]

Outcomes and measurement

The study authors assessed participating students on four measures of mathematical equivalence understanding: equation solving, equation encoding, defining the equal sign, and a 2-week follow-up measure of equation solving with feedback. The first three measures were also combined into a composite measure of mathematical equivalence understanding. The authors also assessed students on computational fluency outcomes (speed, flexibility, and accuracy in solving computation) to determine whether the intervention had negative side effects on performance in this domain. For definitions and more detailed descriptions of these outcome measures, see Appendix B.

Support for implementation

The practice sessions were implemented by tutors. The authors did not describe the background or the training that was given to the tutors who administered the intervention.

Reason for review

This study was identified for review by the WWC because it was supported by a grant to the University of Notre Dame (Principal Investigator: Nicole M. McNeil) from the Institute of Education Sciences (IES).

Appendix B: Outcome measures for each domain

Mathematics equivalence understanding	
Equation solving	Students were asked to solve and explain four problems of the form "3 + 5 + 6 = 3 + __" or "1 + 5 = __ + 2." Students received a point for solving an equation correctly or if they used a correct strategy to solve the problem, with the total score equaling the sum of the points from all four problems.
Equation encoding	Students were asked to rewrite four problems of the form "7 + 1 = __ + 6" or "3 + 5 + 4 = __ + 4" after viewing each for 5 seconds. Students received a point for encoding an equation correctly, with the total score equaling the number of equations correctly encoded.
Defining the equal sign	Students responded to a series of questions about the equal sign; for example, "What is the name of this math symbol?" or "What does this math symbol mean?" For this binary measure, students received a point if they defined the equal sign relationally as a symbol of mathematical equivalence; for example, by saying "two amounts are the same" (McNeil, Fyfe, Petersen, Dunwiddie, & Brletic-Shipley, 2011, p. 1626).
Composite score	The authors constructed a composite of the three measures above. Students received one point if they scored above average on equation solving, one point for scoring above average on equation encoding, and one point for providing a relational definition of the equal sign, for a total of 0 to 3 points.
Equation solving follow-up, with feedback	Two weeks after the intervention, students in nontraditional and traditional practice groups were asked to solve the same four equivalence problems that they had been given previously, although this time, the examiner provided feedback after each question. For example, if the student answered the problem incorrectly, the tutor might say, "No, that's not the number that goes in the blank. The correct number is x because a plus b is equal to x plus y," substituting real numbers for the variables (McNeil et al., 2011, p. 1626).
Computational fluency	
Math computation section of the Iowa Test of Basic Skills, Level 8	As a measure of computational fluency, students were administered the math computation section of the Iowa Test of Basic Skills. This section assesses student ability to conduct arithmetic operations (addition, subtraction, multiplication, or division). The results were scored as the percent correct on this exam.
Difficult addition problems, reaction time	Students were asked to solve single-digit addition problems displayed on a computer screen. Results were scored based on the speed with which the students answered the three most difficult addition problems.
Difficult addition problems, percent correct	Students were asked to solve single-digit addition problems displayed on a computer screen. Results were scored based on the percent of correct answers given for the three most difficult addition problems.
Matched addition problems, reaction time	Students were asked to solve single-digit addition problems displayed on a computer screen. Results were scored based on the speed with which the students answered three easier questions matched to the most difficult problems based on one of the addends.
Matched addition problems, percent correct	Students were asked to solve single-digit addition problems displayed on a computer screen. Results were scored based on the percent of correct answers given for three easier questions matched to the most difficult problems based on one of the addends.

Appendix C: Study findings for the mathematics equivalence understanding domain

Domain and outcome measure	Study sample	Sample size	Mean (standard deviation): Intervention group	Mean (standard deviation): Comparison group	WWC calculations: Mean difference	WWC calculations: Effect size	WWC calculations: Improvement index	*p*-value
Mathematics equivalence understanding								
Equation solving	Nontraditional practice vs. traditional practice	60 students	1.43 (1.72)	0.33 (1.06)	1.10	0.76	+28	nr
Equation encoding	Nontraditional practice vs. traditional practice	60 students	1.47 (1.41)	0.87 (1.22)	0.60	0.45	+17	nr
Defining the equal sign	Nontraditional practice vs. traditional practice	60 students	0.23 (na)	0.03 (na)	0.20	1.36	+41	nr
Equation solving follow-up, with feedback	Nontraditional practice vs. traditional practice	60 students	2.33 (1.73)	1.30 (1.53)	1.03	0.62	+23	0.02
Domain average for mathematics equivalence understanding						**0.80**	**+29**	**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. na = not applicable. nr = not reported.

Study Notes: The sample sizes presented here were reported by the authors in response to an author query. The study did not present *p*-values for the first three contrasts shown in this table, because the inferential tests shown in the article assess the significance of whether there are differences across any of the three groups in the study, not the two groups being examined in this table. The WWC calculated the statistical significance of these contrasts and found that the *p*-values for the equation solving, equation encoding, and defining the equal sign outcomes to be < 0.01, 0.09, and < 0.01, respectively. For the equation solving follow-up with feedback outcome, the study did present the *p*-value for the contrast shown in the table, therefore, that study-reported *p*-value is presented in the table. A correction for multiple comparisons was needed but did not affect whether any of the contrasts were found to be statistically significant.

This study is characterized as having a statistically significant positive effect on mathematics equivalence understanding because the effect for at least one measure within the domain is positive and statistically significant, and no effects are negative and statistically significant, accounting for multiple comparisons. For more information, please refer to the WWC Standards and Procedures Handbook, version 3.0, page 26.

Appendix D: Supplemental findings by 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	
Mathematics equivalence understanding								
Composite score	Nontraditional practice vs. traditional practice	60 students	1.17 (1.09)	0.37 (0.72)	0.80	0.85	+30	nr
Composite score	Nontraditional practice vs. combined traditional practice and no additional practice	90 students	1.17 (1.09)	0.47 (0.81)	0.70	0.76	+28	0.00
Computational fluency								
Math computation section of Iowa Test of Basic Skills, Level 8	Nontraditional practice vs. traditional practice	60 students	52.70 (24.56)	52.93 (29.65)	–0.23	–0.01	0	nr
Difficult addition problems, reaction time	Nontraditional practice vs. traditional practice	60 students	7.64 (4.08)	6.98 (3.86)	0.66	0.16	+7	nr
Difficult addition problems, percent correct	Nontraditional practice vs. traditional practice	60 students	92 (14)	90 (25)	2	0.10	+4	nr
Matched addition problems, reaction time	Nontraditional practice vs. traditional practice	60 students	5.36 (2.86)	5.45 (3.40)	–0.09	–0.03	–1	nr
Matched addition problems, percent correct	Nontraditional practice vs. traditional practice	60 students	96 (12)	93 (16)	3	0.21	+8	nr

Table Notes: The supplemental findings presented in this table are additional findings that do not factor into the WWC's characterization of the study's evidence. 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. nr = not reported.

Study Notes: The sample sizes presented here were reported by the authors in response to an author query. The computational fluency outcomes are presented in this Appendix because the authors wanted to examine whether nontraditional practice had deleterious effects on computational fluency as negative side effects. As such, these analyses of computational fluency outcomes are presented in this Appendix for transparency. The composite score outcome (based on three measures presented in Appendix B) is an alternate, ad-hoc aggregation of the main analyses shown in Appendix C. The author did present one analysis of this outcome that compared the nontraditional practice condition with the combined traditional practice and no additional practice groups. This comparison met WWC group design standards without reservations, and the WWC included this analysis in this table for completeness, though this comparison assesses an unclear contrast in experiences: the effect of nontraditional practice (including 100 minutes of additional practice time) relative to traditional practice (where students had either zero or 100 additional minutes of practice).

The *p*-value for the comparison of the nontraditional practice with the combined traditional practice and no additional practice groups was reported in the article. The study did not present *p*-values for the remaining contrasts shown in this table, because the inferential tests shown in the article assess the significance of whether there are differences across any of the three groups in the study, not the two groups being examined in this table. The WWC calculated the statistical significance of these contrasts and found that the analysis of nontraditional practice vs. traditional practice on the composite score had a *p*-value of < 0.01; the analysis of the math computation section of the Iowa Test of Basic Skills, Level 8 had a *p*-value of 0.97; the analysis of the difficult addition problems, reaction time had a *p*-value of 0.53; the analysis of the difficult addition problems, percent correct had a *p*-value of 0.71; the analysis of the matched addition problems, reaction time had a *p*-value of 0.91; and the analysis of the matched addition problems, percent correct had a *p*-value of 0.42. A correction for multiple comparisons was needed but did not affect whether any of the contrasts were found to be statistically significant.

Endnotes

[1] Single study reviews examine evidence published in a study (supplemented, if necessary, by information obtained directly from the author[s]) to assess whether the study design meets WWC group design standards. The review reports the WWC's assessment of whether the study meets WWC group design standards and summarizes the study findings following WWC conventions for reporting evidence on effectiveness. This study was reviewed using the single study review protocol, version 2.0.

[2] The authors reported that 100 participants were initially recruited for the study, but that 10 students were excluded from the analysis sample. In response to an author query, the authors reported that five of these students were excluded prior to random assignment. The authors also reported the number of students randomly assigned to each condition and confirmed the number included in the analysis for each condition.

[3] After completing the preliminary analysis of the four primary measures of mathematics equivalence understanding, the authors created a composite measure to summarize findings from three measures of mathematical equivalence understanding. Because the composite measure was created as an ad-hoc analysis ("for a more efficient presentation" [McNeil et al., 2011, p. 1626]), the review team content expert and methodologist determined that this measure would be included only as a supplemental finding in Appendix D of this WWC report. The study authors also included an analysis of computational fluency to examine whether nontraditional practice had deleterious effects on this domain as a negative side effect of the intervention. We have included the results of this analysis of computational fluency outcomes in Appendix D of this WWC report.

[4] The authors also presented one analysis that compared the nontraditional practice condition with a combined traditional practice and no additional practice groups (this comparison met WWC group design standards without reservations), and this contrast is presented in Appendix D as a supplemental finding for completeness. This comparison assesses an unclear contrast in experiences: the effect of nontraditional practice (including 100 minutes of additional practice time) relative to traditional practice (where students had either zero or 100 additional minutes of practice).

Recommended Citation

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2014, April). *WWC review of the report: Benefits of practicing 4 = 2 + 2: Nontraditional problem formats facilitate children's understanding of mathematical equivalence*. 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% ($p < 0.05$).

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 3.0) for additional details.