



MONTCLAIR STATE
UNIVERSITY

Montclair State University
**Montclair State University Digital
Commons**

New Jersey Test of Reasoning Skills

IAPC Curriculum

1983

New Jersey Test of Reasoning Skills - Review

Virginia Shipman

Follow this and additional works at: https://digitalcommons.montclair.edu/iapc_nj_test_reasoningskills



Part of the [Social and Philosophical Foundations of Education Commons](#)

Recommended Citation

Shipman, Virginia, "New Jersey Test of Reasoning Skills - Review" (1983). *New Jersey Test of Reasoning Skills*. 3.

https://digitalcommons.montclair.edu/iapc_nj_test_reasoningskills/3

This Article is brought to you for free and open access by the IAPC Curriculum at Montclair State University Digital Commons. It has been accepted for inclusion in New Jersey Test of Reasoning Skills by an authorized administrator of Montclair State University Digital Commons. For more information, please contact digitalcommons@montclair.edu.

The link information below provides a persistent link to the article you've requested.

Persistent link to this record: Following the link below will bring you to the start of the article or citation.

Cut and Paste: To place article links in an external web document, simply copy and paste the HTML below, starting with "<a href"

To continue, in Internet Explorer, select **FILE** then **SAVE AS** from your browser's toolbar above. Be sure to save as a plain text file (.txt) or a 'Web Page, HTML only' file (.html). In FireFox, select **FILE** then **SAVE FILE AS** from your browser's toolbar above. In Chrome, select right click (with your mouse) on this page and select **SAVE AS**

Record: 1

Title: New Jersey Test of Reasoning Skills--Form B

Authors: Shipman, Virginia

Publication Date: 1983-1985.

Publisher Information: Institute for the Advancement of Philosophy for Children, Montclair State University, Upper Montclair, NJ, 07043, iapc@mail.montclair.edu, <http://cehs.montclair.edu.ezproxy.montclair.edu:2048/academic/iapc/>.

Purpose: Assesses elementary reasoning and inquiry skills.

Test Category: Intelligence and General Aptitude.

Population: Reading level grade 5 and over.

Scores: 22 skill areas: Converting Statements, Translating into Logical Form, Inclusion/Exclusion, Recognizing Improper Questions, Avoiding Jumping to Conclusions, Analogical Reasoning, Detecting Underlying Assumptions, Eliminating Alternatives, Inductive Reasoning, Reasoning with Relationships, Detecting Ambiguities, Discerning Causal Relationships, Identifying Good Reasons, Recognizing Symmetrical Relationships, Syllogistic Reasoning (Categorical), Distinguishing Differences of Kind and Degree, Recognizing Transitive Relationships, Recognizing Dubious Authority, Reasoning with 4-Possibilities Matrix, Contradicting Statements, Whole-Part and Part-Whole Reasoning, Syllogistic Reasoning (Conditional)..

Administration: Group.

Time: (30-45) minutes.

Price Data: 2001: \$2.40 per test booklet.

Cross References: See T4:1723 (1 reference); for reviews by Arthur S. Ellen and Rosemary E. Sutton, see 11:259 (1 reference).

Status: In March 2016, the Buros Center for Testing learned this test is out of print..

Reviewers: Ellen, Arthur S.; Sutton, Rosemary E..

Yearbook Volume: 11.

Yearbook Reference: J. J. Kramer & J. C. Conoley (Eds.), The eleventh mental measurements yearbook. 1992.

Published Test Description: New Jersey Test of Reasoning Skills--Form B. Purpose: Assesses elementary reasoning and inquiry skills. Population: Reading level grade 5 and over. Publication Dates: 1983-1985. Scores: 22 skill areas: Converting Statements, Translating into Logical Form,

Inclusion/Exclusion, Recognizing Improper Questions, Avoiding Jumping to Conclusions, Analogical Reasoning, Detecting Underlying Assumptions, Eliminating Alternatives, Inductive Reasoning, Reasoning with Relationships, Detecting Ambiguities, Discerning Causal Relationships, Identifying Good Reasons, Recognizing Symmetrical Relationships, Syllogistic Reasoning (Categorical), Distinguishing Differences of Kind and Degree, Recognizing Transitive Relationships, Recognizing Dubious Authority, Reasoning with 4-Possibilities Matrix, Contradicting Statements, Whole-Part and Part-Whole Reasoning, Syllogistic Reasoning (Conditional). Administration: Group. Manual: No manual. Price Data, 1987: \$2.40 per 12-month test booklet rental including scoring and analysis service for up to 4 answer sheets. Time: (30-45) minutes. Author: Virginia Shipman. Publisher: Institute for the Advancement of Philosophy for Children.

Accession Number: test.721

Persistent link to this record <http://ezproxy.montclair.edu:2048/login?>

(Permalink): [url=http://search.ebscohost.com.ezproxy.montclair.edu:2048/login.aspx?direct=true&db=mmt&AN=test.721&site=ehost-live&scope=site](http://search.ebscohost.com.ezproxy.montclair.edu:2048/login.aspx?direct=true&db=mmt&AN=test.721&site=ehost-live&scope=site)

Cut and Paste: New Jersey Test of Reasoning Skills--Form B

Database: Mental Measurements Yearbook with Tests in Print

New Jersey Test of Reasoning Skills--Form B

Review of the New Jersey Test of Reasoning Skills--Form B by ARTHUR S. ELLEN, Assistant Professor of Psychology, Pace University, New York, NY:

The New Jersey Test of Reasoning Skills--Form B (NJTRS) was designed to evaluate the Philosophy for Children program, an innovative curriculum that asks students to reflect on their thinking through carefully planned class discussion. The NJTRS is based upon a 22-category taxonomy of children's elementary reasoning. Elementary reasoning is here defined as logic learned by children while they acquire language. It differs from higher level thinking, which occurs when students apply elementary skills to more advanced disciplines or solve problems using more than one elementary skill. True to this definition, the questions on the test tap basic logical operations by using an elementary-school reading level (Flesch reading grade level of 4.5 and a Fogg Index of 5.0).

The test consists of 50 multiple-choice questions, each in the form of a short dialogue with three possible answers. Directions are clearly given in the test booklet along with one practice item. The test, according to the publisher, should take about 45 minutes and a 1-hour time limit is suggested. An optical scan form goes back to the publisher, who will return for each class: (a) the Kuder-Richardson reliability index, (b) the mean and standard deviation, (c) the percent correct for each item, (d) a report for each of the pupils telling them the number correct in each of the 22 skill areas of the test, and (e) a chart of available test averages by grade from the publisher's data base.

The NJTRS appears to be used mainly with middle-school students, although it has been given experimentally to students from first grade to college (M. Lipman, personal communication, December 14, 1989). For students under fourth grade, the publisher suggests reading the test aloud. However, this may not work because many younger students will not remember or comprehend the questions and answers. As the test has been used to evaluate a year-long curriculum project, pre- and post-testing are possible with the same form if testing is

completed during one school year.

Technical information is not provided in a manual, but instead the publisher furnishes four sources of information: (a) a portion of a 1982-83 curriculum evaluation report that used an earlier 55-item version of the test; (b) a 1983 three-page brochure from Montclair State College that briefly describes the test; (c) a sheet that reports correlations of the NJTRS with college-level measures of achievement; and (d) a sheet that contains grades 4 through 8 test means as of February 1986.

Although the current test has 50 items, the bulk of the technical information apparently comes from a 55-item version used in the curriculum evaluation project. In 10 communities, that project sampled 2,346 fifth- through seventh-grade students in 74 experimental and 42 control classes. From these data the publishers derived item statistics, test reliability, and some of the rationale for the test's validity.

Item statistics included item difficulties and point-biserial correlations for item to total test score. Although not reported for each item, a good range of item difficulties with a preponderance of "moderately difficult items" was found. Seven items with low point-biserial correlations (.20 or less) were found, and perhaps these results helped to drop items from the 55-item version to make the final 50-item test.

Test reliability in the form of coefficient alpha, an index of internal consistency, was obtained from a random and representative sample of classrooms in the curriculum project. From grades 5 to 7, the coefficient alphas consistently increased from .84 to .94. These reasonably high indices agree with the reliability index of .83 found for a previous version of the test. Unfortunately, test stability was not examined by generating test-retest correlations.

One claim to validity is based upon the test's sensitivity to experimental intervention found in the year-long curriculum evaluation project. From the fall to spring of that school year, the matched-control classes, on the average, increased their test scores approximately 3 points. In contrast, experimental classes, on the average, increased their total test scores a little more than 6 points. These significant differences led the publishers to conclude that the NJTRS responds to changes caused by the Philosophy for Children program.

The publishers argue for at least four additional kinds of validity: content, construct, developmental, and concurrent. Content validity exists because the NJTRS is supposed to sample adequately the elementary reasoning skills taxonomy. Construct validity is based upon the unreported test development research performed at the Educational Testing Service between 1976 and 1978.

Developmental validity usually means test scores exponentially increase with either age or grade. But from grades 4 through 8, the NJTRS's means do not exponentially increase; instead, they remain relatively flat, a test plateau. To explain this, the publishers hypothesize that students develop elementary reasoning skills at about grade 4. However, the interaction between a child's age and item format (O'Brien & Shapiro, 1968; Roberge, 1970) might partially explain the test's ostensible plateau.

Concurrent validity comes from two studies that correlated the NJTRS with measures of college achievement. The first study, with over 600 college freshmen, found moderate correlations with the New Jersey College Basic Skills Placement Test, an exam consisting of five academic achievement subtests in the areas of reading, math, and writing. The second study, with 150 college students, reported moderate correlations between the NJTRS and the Scholastic Aptitude Test math and verbal subtests, and small but significant correlations to college grade-point average. Both studies support the interrelationship between the NJTRS and achievement with college-level students, but not with middle-school students, the intended test takers.

A major shortcoming of the NJTRS is the absence of a comprehensive test manual. A manual minimally must include such missing technical information as: (a) definitions and examples for the taxonomy of logical reasoning; (b) which items correspond to particular skills in the taxonomy, a test plan; (c) what research served as the basis for construct validity; (d) which test statistics were derived with the 55-item evaluation instrument and which came from the 50-item test; and (e) item difficulties and point-biserial correlations for each item.

Another limit of the NJTRS is its lack of subtests. Although the test is intended to evaluate a complex curriculum project, there is only one global test score. This score, a composite of the 22 elementary reasoning skills, will not inform a teacher or program evaluator which skills a pupil has learned. It should be noted, however, that constructing a test with reliable subtests for this many skills would require a much longer test and

evidence that the subtest constructs exist.

In sum, the NJTRS provides an internally consistent composite measure of a unique taxonomy of elementary reasoning skills using a clever item format. However, the publishers could make a better case for test use by providing a thorough and comprehensive report of test information.

REVIEWER'S REFERENCES

O'Brien, T. C., & Shapiro, B. J. (1968). The development of logical thinking in children. *American Educational Research Journal*, 5, 531-542.

Roberge, J. J. (1970). A study of children's abilities to reason with basic principles of deductive reasoning. *American Educational Research Journal*, 7, 583-596.

Review of the New Jersey Test of Reasoning Skills--Form B by ROSEMARY E. SUTTON, Associate Professor of Education, Cleveland State University, Cleveland, OH:

The New Jersey Test of Reasoning Skills was developed in the early 1980s to evaluate the Philosophy for Children Program (PCP). The purpose of the PCP program, developed by Dr. Matthew Lipman at Montclair State College, is to improve students' reasoning skills through classroom discussion that emphasizes generating ideas, discovering resemblances and differences, and finding reasons. Although the Philosophy for Children Program is intended for kindergarten through high school students, the New Jersey Test of Reasoning Skills was developed for use with students in the fifth through seventh grades.

The 50-item multiple-choice test represents 22 skill areas of inductive and deductive reasoning. The language is very simple and the test can be used for as low as fourth grade level. Because so many skill areas are covered, and no information is given about which questions cover which skill areas, this test can be used only to provide information about general critical thinking ability. Thus, this test may be used to evaluate a program such as the PCP, but cannot be used to diagnose specific strengths and weaknesses.

RELIABILITY. The technical information supplied was not in the form of a technical manual, but as part of a final report on the experimental Philosophy for Children Program intervention. Details about reliability indices are given only for an earlier 55-item version of this test. Cronbach's alpha for fifth grade classes ranged from .84 to .87, for sixth grade classes the range was .86 to .89, and for seventh grade classes the range was .91 to .94. The published version of the test has only 50 items. I assume five items with low point-biserial correlations (item to total test score) discussed in the technical information were eliminated, but there is no way to determine which items were eliminated or the new reliability coefficients. These reliability coefficients are high and are unlikely to be altered significantly with the elimination of five items, but the correct figures should have been supplied.

VALIDITY. Content validity was established by producing a taxonomy of the skills needed to perform the operations in the discipline of logic used in childhood and by developing questions from this taxonomy. No information is provided about why some logical skills were selected for inclusion in the test and others were not. Correlations between this test and measures of academic performance are also provided to support validity. These correlations were high and statistically significant for the majority of the measures (e.g., SAT Math, .59; SAT Verbal, .57; Reading Comprehension subtest of the New Jersey College Basic Skills Placement Test, .82). However, all of these data were gathered from samples of college students, even though this test was developed and pilot tested on middle-school children.

NORMS. Norms for 1986 were provided for fourth through eighth grade students. These norms are based on large samples, but no demographic information was provided. Earlier norms were based on 10 diverse subsamples including suburban, rural, and inner-city children. The recent norms show little change in the average number of right answers from fourth grade to eighth grade (31.1 for fourth grade and 34.1 for eighth grade, with standard deviation for both groups approximately 10).

Evaluating critical thinking and reasoning is a very difficult task and there are very few tests appropriate for this age group. While this test appears to be of some value for its original purpose of evaluating the Philosophy for Children Program, I do not recommend it for other purposes. The technical information provided is too limited. If

the publishers develop an appropriate technical manual with accurate reliability indices, details about normative samples, and more information about validity, this test may be worth consideration.

*** Copyright © 2014. The Board of Regents of the University of Nebraska and the Buros Center for Testing. All rights reserved. Any unauthorized use is strictly prohibited. Buros Center for Testing, Buros Institute, Mental Measurements Yearbook, and Tests in Print are all trademarks of the Board of Regents of the University of Nebraska and may not be used without express written consent.