Test Review: Motor-Free Visual Perception Test
Sara Jean McCane
Journal of Psychoeducational Assessment 2006 24: 265
DOI: 10.1177/0734282906286339

The online version of this article can be found at:
http://jpa.sagepub.com/content/24/3/265
DOI: 10.1177/0734282906286339

The Motor-Free Visual Perception Test: Third edition (MVPT-3; Colarusso & Hammill, 2003) purports to measure overall visual perceptual ability. Task responses require no motor ability, eliminating the effect of motor performance on the overall visual perception score. The test authors suggested that this MVPT-3 characteristic allows for its use in the assessment of visual perceptual ability in motorically impaired individuals.

This individually administered test is designed for children, adolescents, and adults ranging from 4 through 95 plus years. Various professionals, including occupational therapists, teachers, school psychologists, and optometrists, can administer the MVPT-3. Examiners are not required to have specific education in assessment but must be trained and supervised in assessment and test interpretation. Only individuals familiar with both psychometric properties and test score limitations should conduct interpretations.

The purpose of this instrument is to provide an alternative measure of visual perception. Typical visual perception measures include motoric responses and tasks (e.g., copying or tracing) that ultimately assess both visual and motor abilities. The MVPT-3 seeks to produce an isolated measure of visual perceptual ability independent of motoric capability. Examiners use this instrument for a variety of purposes, including the determination of visual perceptual differences across several diagnostic classifications. Occupational therapists also administer this test to determine driver recertification eligibility following strokes or head injuries.

The authors stated that there is no one theoretical model regarding perceptual processes. Research in this area suggests that overall visual perceptual ability relies on five interrelated processes, including spatial relationships, visual discrimination, figure-ground, visual closure, and visual memory. Skills in spatial relationships require the ability to orient oneself in space. This ability also involves the accurate perception of objects in relation to oneself and other objects. Visual discrimination requires the ability to discriminate salient object features. The ability to discriminate an object from its background is classified as figure-ground. Visual closure involves gestalt-like ability to perceive a whole figure when fragments are missing. Lastly, visual memory involves the ability to recognize a stimulus following a brief interval. Research shows that these processes do not occur in isolation and should not be measured individually. The MVPT-3 authors designed their test based on this research. Items were developed to closely resemble each of these areas. As suggested, individual scores are not produced for each task type. The interrelatedness of these tasks allows only for an overall visual processing score to be produced.

Materials for the test include a manual that describes administration and scoring procedures, test development, and technical adequacy. Examiners use a record form to record examinees’ responses, and all test items are presented horizontally via a spiral-bound test plates easel.

The test plates easel contains 65 items representing the five visual perceptual areas previously described. These five task types do not represent different subscales or subtests and do not yield individual scores. Items with similar instructions are grouped together in order of increasing difficulty. Each item consists of a black-and-white line drawing stimulus, along
with four multiple-choice options. For most tasks, the stimulus and response choices appear on the same easel page, allowing examinees to view them concurrently. The stimulus drawing appears at the top of the page above a row of four solution choices. Memory-loaded items present the stimulus and choice items separately. The stimulus page is presented for 5 seconds, removed, and the options page is presented. Item responses do not require visual-motor skills. Examinees can respond either by saying the letter of or pointing to the selected option. The examiner records each response on the record form.

**Administration and Scoring**

According to the manual, the MVPT-3 takes approximately 20 to 30 minutes to administer and approximately 10 minutes to score. Authors suggested testing be conducted early in the day (when possible) in a distraction-free environment. This setting allows for better attention and concentration. Prior to assessment, the examiner should determine the examinee’s date of birth and compute his or her chronological age. The examinee’s chronological age determines the individual’s starting place within the test.

The same test format is used for all ages, but item starting points vary. Children ages 4 through 10 begin with the first example item and complete Items 1-40. Older examinees begin with the third example item and complete Items 14-65. No basal or ceiling rules exist. Examinees must complete all items corresponding to their age group, and no additional items are administered.

Administration instructions for each item are printed on the back of the preceding item plate. Administration requires standardized instructions that must be stated verbatim. No variations are allowed. To ensure the examinee’s understanding, example items are presented for each new set of instructions. Examiners must ensure that the examinee understands these directions before proceeding to the next item. Each item is presented via a test plates easel and requires the examinee to select an answer from a row of four choices. The examinee indicates an answer either by saying the letter of or pointing to the selected option. The examiner records each response on the record form and calculates a total raw score.

The MVPT-3 yields a single raw score that represents the examinee’s overall visual perceptual ability. The raw score for this test is determined by subtracting the number of errors made from the number of the last item attempted. The manual contains norm tables to be used in converting raw scores to standard scores, percentile ranks, and age equivalents. These three types of derived scores allow for the comparison of an examinee’s performance to that of a normative group of same-aged peers. Normal curve equivalents (NCEs), T-scores, and stanines are also available. The manual lacks grade-based norms needed for determining grade equivalents, and no cluster scores are available.

MVPT-3 standard scores range from 55 to 145, with a distribution mean of 100 and a standard deviation of 15. Standard score confidence intervals can be determined at either the 85% or 90% level. Bracken and McCallum (1998) suggested that standard scores extend at least two standard deviations above and below the mean. The MVPT-3 standard score distributions extend three standard deviations above and below their means for all age groups, satisfying the aforementioned criterion.
Technical Adequacy

Test Construction

*Standardization.* The authors recruited 2,005 examinees for participation in the norming and standardization process. Overall, the authors selected 1,856 participants meeting specified demographic criteria to compose the large norm sample. The remaining participants not included in the norm sample participated in other validity studies. Participants resided in 118 cities across a total of 34 states. Public, private, and parochial school students were included. The authors selected student participants at random from regular education classrooms. Non-head-injured adult participants were also selected at random from hospitals and institutions.

Although researchers encouraged testing representative of a variety of disabilities, the manual failed to note the disability status for hospitalized or institutionalized adult participants. The omission of these classifications prevents reviewers from determining whether this adult sample is representative of the national population. Furthermore, adult participants were only obtained from hospitals and institutions. This sampling procedure may have led to an underrepresentation of nondisabled adults. The authors did include information regarding the disability status of children participants ages 6 through 18. These groups appear to be representative of the U.S. population based on the 2000 census.

The authors claimed the standardization sample to be representative of the U.S. population’s demographics, including regional, ethnic/racial, gender, residential, and disability characteristics. The manual further noted that some age groups were either under- or overrepresented, and a statistical weighting procedure had to be used to correct for this misrepresentation. Unfortunately, no information is provided as to which groups were misrepresented or as to how well each group represented the national population on the individual demographic characteristics. This information would be helpful to examiners in determining the appropriateness of this test for use with specific age groups possessing particular demographic characteristics.

*Item characteristics.* Item analyses conducted with previous editions of this test, the MVPT and MVPT-R, demonstrated that little item difficulty or total score variance existed for examinees older than 10. Because the test is often used with older populations, more difficult items were needed. For the MVPT-3, authors added new items in each of the five visual perceptual task areas. The manual is unclear about how many items were added for each area and whether these additions were balanced across task types. Several of these items were eliminated in the norming process due to inadequate discrimination or other problematic features.

The final version of the MVPT-3 contained a total of 65 scored items. Of these, 29 of the 31 newly developed items were retained, with 25 serving as scored items and 4 serving as example items. Authors reconducted data analyses following final eliminations. For ages 4 through 10, Items 1 through 40 were administered yielding a median item difficulty index of .78 for this age range. Median item difficulties by age ranged from .48 to .92, increasing from age 4 to age 10. Difficulty levels optimally should range from .15 to .85, with .50 yielding maximum discrimination (Anastasi & Urbina, 1997). A few extremely easy items \((p = .00 \text{ to } .15)\) and a few extremely difficult items \((p = .85 \text{ to } 1.00)\) should also be included for very limited and very capable examinees. The manual does not provide a table of individual item dif-
difficulties but does list the item difficulty range for each age group. The 4-year-old age group yields a difficulty range of .15 to .72 and a median of .48. It does appear as if some of the more difficult items may be too complicated for this age range, but there are no rules that allow these children to stop. Instead, all items must be completed, possibly frustrating younger examinees. For ages 7, 8, 9, and 10, median item difficulties are higher (.78, .87, .91, and .92, respectively) than the suggested .75 cutoff, suggesting that many items may be too easy for children in this age range.

For ages 11 and older, Items 14 through 65 are administered yielding a .81 median item difficulty index. Median item difficulties by age group ranged from .67 to .89, suggesting that this test may still not be as challenging as needed for this older population. Overall, median item difficulty indexes suggest that items were hardest for the youngest examinees and easiest for older examinees, as expected.

Authors may need to consider developing more conventional floors and ceilings that rely on both age and performance rather than age alone. Median difficulty indexes for ages 7, 8, 9, and 10 suggest that some of the items assigned for these ages may be overly simplistic. By not allowing these children to continue to their peak performance, there is no way to distinguish between a person who may have completed all 65 items correctly and a person who may not have succeeded past Item 40. Both examinees would obtain a raw score of 40 based on the current discontinue rule. The authors justify this format by citing data from earlier test editions. They claim that sufficient score variability existed between scores in the 4 to 10 age group, eliminating the need to extend the test for this group. Actual data however are not presented, and no further rationale is offered regarding discontinue rules. Similarly, the authors suggest that performance on previous editions indicated little variance among test scores and item difficulties for examinees older than 10 years of age. These data led to the development and addition of more difficult items. On the MVPT-3, examinees older than 10 begin with Item 14 and end with Item 65. No rationale is provided for this designated beginning point. Examinees in this age group are not administered Items 1 through 13 regardless of performance. Lower functioning examinees could potentially miss Items 14, 15, 16, and so on but still be required to continue to item 65. This situation would likely frustrate the examinees. Furthermore, as the examinee could miss all administered items but never be administered Items 1 through 13, true levels of functioning would not be determined.

The authors further computed item discrimination indexes using biserial correlations. These analyses yielded median discrimination indexes ranging from .32 to .57. Discrimination index medians for ages 4 through 10 and 11 plus years are .47 and .54, respectively. All median item discrimination indexes for age levels within these groups are also positive, suggesting that MVPT-3 items generally discriminated between high and low performers.

Adequate item gradients require that each raw score point increment leads to a standard score increment equal to or less than one third of the test’s standard deviation (Bracken, 1987). A number of item gradients found within all age groups fail to meet this criterion.

Reliability. To determine internal consistency, Cronbach’s coefficient alpha was computed for each age group, ranging from .69 to .90. Sattler (2001) suggested that reliabilities equal to or greater than .90 are needed for diagnostic and decision-making purposes. Reliabilities of .80 and .60 reliabilities are typically recommended for screening and research purposes, respectively. Alpha reliability coefficients for ages 4, 5, and 7 equal .69, .76, and .73, respectively. These values meet the research criterion but fail to meet the screening criterion. Therefore, the
examiners should be cautious when considering the test for other than research purposes for these three age groups. Reliability coefficients for all other age groups exceed .80 and are acceptable for a screening instrument. However, the 14- through 18-year age group is the only age group that satisfies the diagnostic criterion ($r = .90$). The authors claim that the MVPT-3 should be used only as a screening instrument for 4-year-old children but can be used for diagnostic purposes at all other ages. However, even more cautious interpretation is urged.

The authors conducted a test-retest study using a 34-day retest interval. Results for 103 examinees were examined across two age groups, 4 through 10 years and 11 through 84 plus years. There were 28 examinees in the former group, 75 in the latter. Correlations for the two age groups equaled .87 and .92, respectively, suggesting that the test is relatively stable over time.

**Validity: Content.** Regarding content validity, previous research indicates that general visual perceptual ability involves five types of interrelated processes, including spatial relationships, visual discrimination, figure-ground, visual closure, and visual memory (Chalfant & Scheffelin, 1969). The authors developed items to represent these five areas, attempting to develop a complete item sample. Due to the briefness of this instrument and the interrelatedness of these skills, only a general visual processing score is produced. The representation of each item type may eliminate mono-operational bias as a score is not based on only one task type. Regrettably, the manual provides no information regarding how many items of each type are included on the test.

Furthermore, the content of the test was designed to measure examinees’ visual perceptual ability without requiring motoric responses. To accomplish this end, authors labeled each answer choice A, B, C, or D, allowing examinees to simply state their selection. Research conducted with the original MVPT demonstrated that children with severe physical disabilities performed better on the MVPT as compared to the Developmental Test of Visual-Motor Integration (VMI) and the Developmental Test of Visual Perception (DVTP; Zeitschel, Kalisch, & Colarusso, 1979). Similar score increases were not found for children not classified as possessing a physical disability. Both the VMI and DVTP require motoric responses. As severity of disability increased, the number of children failing these tests also increased. Although this study was conducted on the original MVPT, formatting has not changed for the most recent addition. These data support the content validity of this instrument.

The authors also examined item bias. The effects of gender (male/female), residence (urban/rural), and ethnicity (Black/White, Asian/White, and Hispanic/White) were examined. The groups compared differed on one of these demographic variables but possessed the same overall visual perceptual ability. Performance on each item was compared for differing groups to determine any biased content. Only three items in 2 of the 14 different comparisons appeared to function differently based on group membership. The authors examined these items and chose not to eliminate the items based on other psychometric data. From their analyses, authors conclude that the content of this test is not biased based on gender, place of residence, or ethnicity.

**Construct.** To examine construct validity, the authors evaluated information regarding chronological age, cognitive ability, academic achievement, and exceptional group differences. They hypothesized that visual perception scores increase with age, peak, and then begin to decrease steadily. To test this hypothesis, the authors examined score trends across
ages. Data revealed that MVPT-3 scores and age are positively correlated from ages 4 to 39, from which point the trend reversed, and the two variables became negatively correlated. For the 4 to 10 age group and the 11 to 39 age group, correlations equaled .72 and .37, respectively. The lower positive correlation for the 11 to 39 age group could be explained by smaller yet constant gains. As hypothesized, the authors found a correlation of −.46 for the 40 plus age group. These data support the hypothesis that visual perceptual skills and age are positively correlated at earlier ages and become negatively correlated later in life.

The authors further hypothesized that visual perception and cognitive ability are slightly related due to their common link to physiological maturity. This correlation was expected to be low. To demonstrate this relationship, the authors compared examinees’ MVPT-3 performance to available cognitive test scores. Scores on the MVPT and the Slosson Intelligence Test were correlated for 63 examinees and yielded a significant correlation of .31 ($p < .05$). Similarly, researchers compared MVPT scores with Pinter-Cunningham Primary Test scores for 42 examinees ($r = .32$, $p < .05$). MVPT-3 scores were also correlated with WISC-III Full Scale IQs, Performance IQs, and Verbal IQs for 21, 18, and 18 examinees, respectively, with no significant correlations. The authors concluded that these data support their hypothesis and provide evidence for construct validity.

To further assess construct validity, the authors compared MVPT-3 scores with academic achievement as measured by subtest scores on the Metropolitan Readiness Test, Stanford Achievement Tests, and Durell Analysis of Reading Difficulties. They expected only low correlations to result. Overall, the significant subtest correlations ranged from .33 to .51, with a median of .41. Performance on all subtests of the Durell Analysis of Reading Difficulties correlated significantly with MVPT-3 scores, ranging from .33 to .46. Total Metropolitan Readiness Test scores also significantly correlated with MVPT-3 performance ($r = .42$). Only the Arithmetic subtest drawn from the Stanford Achievement Tests resulted in significant correlations ($r = .42$). The magnitude of these relationships supports the hypothesis that visual perceptual ability as measured by the MVPT-3 and academic achievement has only a low to moderate correlation.

Lastly, the authors examined MVPT-3 performance differences between three exceptionality groups against the overall population mean. For example, those examinees with academic difficulties or head injuries should be more likely than their undiagnosed peers to possess visual perception deficits. Scores for examinees classified as developmentally delayed, head injured, or learning disabled were compared against the mean of 100. Each group earned significantly lower scores as compared to this mean. Examinees classified as developmentally delayed earned a mean MVPT-3 score of 69.46, falling more than two standard deviations below the mean. Examinees classified as head injured earned a mean score of 80.16, falling approximately 1.33 standard deviations below the mean. Lastly, examinees classified as possessing a learning disability earned an average score of 88.24. The lower MVPT-3 scores for each of the exceptional groups support the hypotheses.

The analyses conducted support the construct validity of this instrument. However, no comparisons were made between MVPT-3 scores and real-life tasks requiring visual perceptual skills. Such analyses would strengthen support for construct validity. The authors did not conduct a factor analysis as they hypothesized that only one construct is measured by this instrument.
Concurrent. The authors fail to address the concurrent validity of the MVPT-3. Unfortunately, the only data analyses regarding correlations with other visual perception tests were conducted on the original 1972 edition (MVPT). This test contained only 36 items and was not normed for adults. At that time, the MVPT was compared to four other tests, including the Frostig Test (DVPT, $r = .73, p < .01$), the DVPT-2 ($r_s$ ranged from .27 to .78 for individual subtests/composites, $p < .01$ for all subtests/composites), the Metropolitan Readiness Tests ($r = .40$ for Matching and .31 for Copying, $p < .05$ for Matching and not significant for Copying), and the Durell Analysis of Reading Difficulties ($r = .46, p < .01$). The MVPT and the Matching subtest of the Metropolitan Reading Test correlated .40, significant at the .05 level. The MVPT and the Copying subtest correlation equaled .31 and was not significant. Based on the moderate to high relationship found with the other visual perception tests, the authors concluded that the MVPT and these other tests measure similar constructs. Despite this similarity, considerable unique variance existed and likely resulted from the elimination of the motor component contained in the other tests. These results suggest strong concurrent validity for the original MVPT. However, these data cannot be used to infer the validity of the MVPT-3. Further research must be conducted to determine whether the newest edition correlates with other current measures of visual perception.

Other validity. The authors fail to provide predictive validity data in the manual. The MVPT has been used to determine driver recertification for adults who have suffered strokes or head injury. The authors claim that the MVPT is more predictive than other visual perception tests as to who would fail on-road driving tests. The authors did not note how predictive the MVPT test scores were, only that they were better than other tests. This omission of specific data does not allow one to assume that MVPT scores accurately predict on-road driving ability. Furthermore, numerous items have been added to the instrument since its original publication.

Conclusion

Overall, the MVPT-3 appears to be examiner and examinee friendly. Administration is brief and easy, requiring no specialized assessment training. The authors designed this test based on previous research suggesting the importance of five interconnecting perceptual processes to overall visual perceptual ability.

Research on the test format supports the claim that it is acceptable for motorically impaired individuals. Furthermore, the authors contend that the test is acceptable for research, screening, and diagnostic purposes, a claim that is suspect for some ages; for example, for ages 4, 5, and 7, the MVPT-3 is best used for research purposes only. Examiners should examine reliabilities before administering this test to ensure that the test meets the minimum criteria for their purposes.

Although the authors present data that support the content, construct, and concurrent validity of this instrument, these analyses are not always definitive or in some cases are based on data from earlier editions. Furthermore, item gradient and item difficulty problems exist at some ages. Examiners must consider these limitations when interpreting test scores.

Examiners are advised to use the MVPT-3 but with caution for some examinees. Future studies will continue to determine the validity of the revised instrument.

Sara Jean McCane
University of Tennessee, Knoxville
General Description

The Young Children’s Achievement Test (YCAT; Hresko, Peak, Herron, & Bridges, 2000) is an individually administered achievement test designed to evaluate preschool, kindergarten, and first-grade children for risk of school failure. The test is comprised of five subtests specifically intended to assess general information, reading, mathematics, writing, and spoken language. The YCAT was developed to be used with children ranging from 4 years 0 months to 7 years 11 months old. The authors stress the importance of early identification for preventing academic problems, noting that early academic ability is the best predictor of later academic ability. Furthermore, they believe that results from the YCAT can be used in developing and monitoring educational interventions.

Test authors recommend that examiners be formally trained in test administration and have an understanding of test construction and test statistics; knowledge of test administration, scoring, and interpretation; and specific knowledge in testing children age 4 to 7.

Materials for the test consist of an examiner’s manual that provides an overview of the YCAT; information about administration, scoring, and interpreting results; and support for the technical adequacy of the test. An easel picture book is used to present items for the five subtests. In addition, a student response form is provided for examinees to record written responses. Similarly, a profile/examiner record booklet is used by examiners to record student responses.

Test scores are derived from age-based norms. Raw scores can provide numerous derived scores, including age equivalents, percentile ranks, standard scores ($M = 100$, $SD = 15$), normal curve equivalents (NCEs), $z$ scores, $T$-scores, and stanines. Appendices for converting raw scores are provided in the examiner’s manual. Administration of the YCAT generally takes 25 to 45 minutes. Because each subtest is untimed, precise administration time is unavailable.