PSIWORLD 2014

A new Block Design test: An exploratory study

Aurel Ion Clinciu*

Transilvania University of Brasov, Faculty of Psychology and Sciences of Education, Nicolae Balcescu 56, 500019 Brasov, Romania

Abstract

The present paper draws on the specific results of a vocational guidance battery of tests which are relevant to a newly created test of Block Design. This study involved 366 pupils at secondary school and high school levels (average mean of age was 16.1 years old, SD 1.80) and reveals information on the intrinsic psychometric qualities of the new Block Design Test. In addition, it offers insight into the specific aspects of the test that come as a result of its inter-relationship with other tests of the battery of intelligence used.

© 2015 The Authors. Published by Elsevier B.V.

Keywords: Block Design test; intelligence; vocational guidance.

1. Introduction

The Block Design test was originated by Kohs (1923) who created it as a comprehensive measure of nonverbal intelligence. Matarazzo appreciates that “The initial enthusiasm by its originator seems fully justified (…) and our own experience shows that it conforms to all criteria of a ‘good’ test.” (1980, p. 212). In 1930 Arthur incorporated Kohs designs in her Performance Scale. Later, adaptation of this test appear in a number of intelligence scales, like in Wechsler’s original or revised scales (W-B I, WAIS, WISC and WPPSI), in Stanford-Binet Fourth Edition (SB-IV) intelligence scales or in Bonnardel scales (B20, B101).

The adaptation of Kohs’s test resulted from the need to obtain a form that preserves the basic quality of the test, that of being an excellent tool of assessing visuo-spatial ability and perceptual organization, however taking into account the modification of the testing and scoring processes, thus creating a more manageable, faster, and overall better test. Its use was nonetheless limited by the fact that the test requires additional support (the blocks) and can only take place face to face, thus requiring more testing time. The test itself demands high levels of attention from both parties: the person being tested and the test administrator who runs the testing and the scoring sessions. To these potential difficulties, additional incidental factors may require consideration, such as any consistent previous experience with block or similar joining games of the testee, his/her level of attention focus, tiredness, any mental

* Aurel Ion Clinciu. Tel.: +40-268-318-149; fax: +40-268-470-190.
E-mail address: a_clinciu@yahoo.com.
illness condition. Furthermore, dexterity can influence scores positively, particularly for those using a trial-and-error approach. The administration time can be very lengthy and a learning set may be observed to influence scores.

The balance between the possible disadvantages that may be encountered and the advantages of this test positively leans towards the latter. Thus, through all the abbreviated forms of the W-B I, WAIS or WISC, it is a consistent recommendation that the block design test ought to be included in the duads, triads and tetrads, as it is considered to be the most representative of the non-verbal, performance, or perceptual organization sub-tests (Matarazzo, 1980, p. 253; Glasser & Zimmerman, 1967, pp. 134-144; Wechsler, 2002, p. 94).

By attempting to reduce the disadvantages mentioned above of the block design sub-test, the Wechsler scales of intelligence that have been developed since have used a lower number of patterns with an easier scoring process, which are in turn more rigorous in offering bonuses for speed of execution and improving the use of the answer sheet when scoring. However, the critical observations of Zimmerman and Woo-Sam (1973), for both the Block Design from the original WAIS and from the WISC have remained valid: high ceiling effect for young ages and low intelligence levels and low ceiling effect for older testees with higher intelligence levels.

2. Purpose of the study

The primary aim of the present paper was to develop a test of blocks (cubes) which would maximize its psychometric qualities and reduce some of the limitations noted previously. The impetus to build a new test of blocks had the specific aim of creating a substantial growth in the test’s diagnostic value through:

a. a careful selection of patterns which can be rigorously grouped by difficulty and complexity in three levels, consisting of three series of four models each;

b. the three series, ordered by difficulty, should be representative for early childhood, puberty and adolescence and should include starting items (first series) easy enough for young ages and exit items (series three) difficult enough to increase the ceiling effect of the test for older participants with higher levels of performance;

c. essential changes in the scoring algorithm in order to significantly increase the overall variability of the scores;

d. a uniform and highly differentiated method of granting bonuses for the speed of execution which would become the primary source of variability of the total scores.

An additional aim of our research was to create a test of blocks which could be administered independently as a visuo-spatial intelligence test and, at the same time, could be included as a sub-test in a more complex battery of intelligence tests. These requirements have had an impact on the construction, administration, scoring and calibration processes of the test as a whole. The 12 patterns of the Block Design test that we have created are completely new. The first series of two-by-two blocks includes lines which separate and individualize each of the blocks, which in turn results in the easy identification of a suitable working mode. The second series is still a two-by-two blocks design however this series eliminates the separating lines and proposes models with more perceptually complex shape-fond relationships. The third series has a three-by-three blocks design and brings the most complex, demanding and productive patterns of the test. In order to reduce the search time, all models use blocks with white and red sides only (two with white sides, two with red ones and finally two bicolor ones). The achievement of the second objective (three series ordered by difficulty, with low output and high output ceiling) is to be analyzed next. To significantly increase the variability of the scores, special attention was given to the scoring method. Experience up to date in administering this test indicates that there is a low likelihood of participants not being able to complete all the patterns in the time given. The general rule of ending the testing episode after three consecutive failed attempts was rarely used. It can therefore be inferred that should the testing time be sufficiently large, almost most participants can solve the 12 models. Thus, this would suggest that the main differentiator between participants lies in the method of awarding bonus points for the speed of execution. The bonus for the speed of execution for the first series of patterns is of 45 seconds, for the second series 75 seconds and for the third series 150 seconds. For each interval of 15 seconds one bonus point is awarded, therefore the first series can award two bonus points, the second series can award four and the final series can award nine bonus points for the speed of execution. This results in a maximum of 60 bonus points for a quick, correct and full completion of a pattern. However, since not all patterns can be fully completed by all participants, another source of increased variability in scores comes from scoring the patterns which are only partially completed. Thus, for each uncolored side that is picked and fitted in the correct place the participant receives a point. Another point is awarded for each bicolor side...
placed in the correct place, but in an incorrect position, while a bicolor block placed in the correct place and position receives two points, one for place and one for position. There is a total of 120 points for all patterns which, together with the total of 60 points awarded as bonuses, lead to a maximum total score of 180 points. These scores seem to offer a wide range of variability, high enough not to discriminate between ages or different levels of performance.

Secondly, I aimed to develop a battery of intelligence tests destined for the diagnosis of aptitudes which would support vocational guidance in line with Holland’s Self-Directed Search (SDS) questionnaire. As this issue was also considered from a personality perspective (the Big Five model) a final objective was that of determining the relationships between the test of Block Design and the other intelligence tests used on one hand, and the interest and personality variables on the other hand.

3. Method

3.1. Participants and procedure

There were 366 participants tested, of which 30 were in the 6th grade, 30 in the 8th grade, 134 in the 9th grade and 172 in the 11th grade. The average age was 16.1, SD1.80. The participants were pupils from four important schools in Brasov, each with a different program: theoretical, technical, IT and humanistic. Participation was voluntary, based on parental consent and also approvals of the school inspectorate and the school manager were obtained. The students and their parents were informed about the nature and purpose of the study and were assured about confidentiality of their results. In order to stimulate students’ participation, the psychologist informed them that each participant would receive tailored feedback and this happened once all results were analyzed. Participants were advised they could withdraw at any point during the study and no penalty would be given, however none withdrew. The administration of the test was made collectively for seven tests and/or questionnaires of intelligence, personality and interests. The Block Design test was performed individually by each pupil, at the end of the testing session, under the observation of two test-administrators who worked simultaneously in different corners of the classroom. Special care was given in order to completely eliminate the possibility of contamination. The allocated time for the whole procedure was equivalent to a full school day, namely six hours, of which three hours was fully dedicated to the individual administration of the Test of Blocks.

4. Results and discussions

The factor analysis of the battery of tests we used led to the identification of a solution with three factors, covering 73.62% of the total variance. The three factors we identified were figural reasoning, spatial-graphic and verbal. Figural reasoning covered 47.22% of the variance. As factor of fluid intelligence, it included tests (Number Series, Matrices, Quantitative Reasoning and Block Design) that proposed inductive reasoning and analysis in the visual bi-dimensional field. The second factor covered 15.56% of the total variance and was formed of two newly developed drawing tests, Draw a Person and Bender-Gestalt Standard, which measured the level of conceptualization of a few important graphic constructs. The third factor covered 11.61% of the total variance and constituted the verbal component of this battery. It was formed of the Definition and Verbal Recombination tests, of vocabulary and verbal fluency respectively.

| Table 1. Factorial structure of intelligence battery |
|-----------------------------------------------------|
| **Comunalities** | **Fluid intelligence** | **Crystallized intelligence** |
| Number Series    | .767                   | .857                        |
| Matrices         | .792                   | .801                        |
| Quantitative Reasoning | .695            | .794                        |
| Block Design     | .611                   | .557                        |
| Draw a Pearson Test | .742             | .851                        |

Fluid intelligence Crystallized intelligence
As it can be noticed, the three factors that were extracted produced high communality for all the sub-tests of the battery. The Block Design test splits its variance between the first and the second factors, and can thus be classified as both a test of inductive figural reasoning and an analysis of visual bi-dimensional space. Additionally, the Verbal Recombination test of the third factor correlated highly with the first factor, which resulted in a decrease in the purity and specificity of each of the factors extracted. To determine the most appropriate factor structure for this battery of tests a confirmatory factor analysis was additionally required. In short, Block Design is a fair measure of those aspects of general intelligence which emphasizes the ability to visualize and construct various geometric designs from different component parts, involving the ability to perceive, analyze and synthesize spatial relationships.

Results indicate that the blocks sub-test highlighted a clear hierarchy, on three successive levels, of the degree of difficulty of the 12 patterns, both by speed (the time required to complete the task) and the quality of the test completion (the blocks placed in the correct slots plus the ones oriented correctly and placed in the appropriate slots). The average mean time for the first two series was under two minutes and reached almost six minutes for the difficult final series. By combining the two scoring criteria (speed and quality), the scores indicated an average progression of 28, 45 and 78 points for series I, II and III respectively. The average time required for the administration of the test was under 8 minutes, thus indicating an extremely manageable test. The average score of the whole sample was 151 points, and strongly skewed to the left (asymmetry). This average value could leave significant space to differentiate between scores at the medium and high levels of performance. The additional 151 points in the left side of the distribution would allow the creation of fine score groups for low and medium ages but also for lower levels of intelligence. Therefore the limitation this sub-test may have had in the Wechsler test, of the difficult start to the test for younger participants and the uncomplicated finish of the test for the older participants, appears to have been eliminated in this newly developed test. However, a more extensive and detailed answer to this issue will only be possible once the test is extensively administered and used for participants of all ages, and the standards for normative samples are developed.

More specific issues raised by the use of any psychometric test relate to its fidelity and validity. In this case, when looking at the fidelity of the test, the internal consistency was determined after eliminating the first three items, with close to zero variance. Cronbach’s alpha indicated a good coefficient of 0.80. The correlation between the nine items and the scale ranged from 0.40 (second series items) to 0.70 (third series items), which suggested a homogenous construct, defined most towards the end of the test. The evaluation of fidelity through the test-retest method will be assumed in the follow-up stages of this research. The value of the Block Design test was studied through its correlation with the intelligence tests of the battery (concurrent validity) as well as through its correlation with an external criterion (predictive validity). The results of this analysis are summarized in Table 2.

|                           | Bender-Gestalt Standard | Definitions | Verbal Recombination |
|---------------------------|-------------------------|-------------|----------------------|
|                           | .727                    | .846        | .875                 | .919                  | .679 | .492 | .636 |
The correlation between the battery of intelligence tests and the general point average was high (r = 0.63), which would indicate a strong argument for its predictive validity. As expected, this is higher for the crystallized component of the battery (r = 0.62) than for the fluid one (r = 0.53), and higher for the Romanian language (r = 0.61) than for mathematics (r = 0.41), where the learning continues and sequencing plays a decisive role in obtaining high levels of performance. The correlations of the Block Design with the other tests of the battery were medium to high, ranging from 0.26 (Definitions) to 0.59 (Matrices). For a non-verbal test, the correlation of 0.35 with the criterion variable (Grade Point Average) for mainly adolescents was an enough strong one. It could also be noted that the new Block Design formula did not offer significant score differences by gender.

The Block Design is not only an excellent test of general intelligence, but it also provides a high level of information when used in qualitative analyses. One can gain valuable insights into the participant’s performance by watching “how” he/she completes the task set. Additionally, another interesting issue is that of the method employed by the participant in assembling the designs by either following the figure as a whole versus breaking it into component parts (synthetic or analytic approach). Furthermore, factors such as personal style, motivation, test anxiety, emotional reaction on the part of the participant, impulsivity, and temperamental clues can evolve and be observed during the administration of the test. In conclusion, the newly created Block Design test appears to be a powerful psychometric instrument, robust both in terms of fidelity and validity, however further approaches and research are necessary to reach its final version.

References

Glaser, A. J., & Lee Zimmerman, I. L. (1967). Clinical interpretation of the Wechsler Intelligence Scale for Children (WISC). New York: Grune & Stratton.

Jackson, D. N., Iliescu, D., & Glință, F., (2008). MAB-II. Multidimensional Aptitude Battery. Manual tehnic. Cluj-Napoca: Editura Sinapsis.

Kohs, S. C. (1923). Intelligence measurement: a psychological and statistical study based upon the Block-Design tests. New York: Macmillan.

Matarazzo, J. D. (1980). Wechsler’s Measurement and Appraisal of Adult Intelligence. (Fifth ed.). New York: Oxford University Press.

Wechsler, D. (1976). WISC-R manual. Wechsler Intelligence Scale for Children – Revised. Anglicized version. Oxford: NFER Publishing Company.

Wechsler, D. (2002). WAIS – III. Wechsler Adult Intelligence Scale – Third edition. WMS – III. Wechsler Memory Scale – Third edition. Technical manual. San Antonio: The Psychological Corporation.

Zimmerman, I. L., Woo-Sam, J. M., & Glaser, A. J. (1973). Clinical Interpretation of the Wechsler Adult Intelligence Scale. New York: Grune & Stratton.