DRAFT

Exploring the Internal Structure of the C-10

Xavier Jouve

Cogn-IQ.org

Abstract

The aim of this research by exploring its internal structure is to examine the meaning and the reliability of scores yielded by the C-10. Analyses were conducted on the data from 102 examinees, all of them having taken voluntarily the C-10 online. For this study, multidimensional scaling (MDS), internal consistency and factor analyses of the scale were performed. The three subtests of the C-10, namely the verbal analogies (ANA), the numerical sequences (SEQ) and the word similarities (SIM), were each observed to relate highly to a general ability factor (g(β)) with loadings ranging from .88 to .92. Reasoning (RG) and verbal (V) factors were also found but did not show high statistical significance. Furthermore, items of the test appeared to be scaled according to a Guttman's effect. The reliability coefficient as calculated with Spearman-Brown method showed a very satisfactory level at .97. Therefore, these findings suggest the C-10 can be regarded as reliable measure of a general cognitive functioning trait. Further research should be necessary in order to determine more appropriately what the nature of this trait is and to discuss the apparent lack of asymmetry of g.

Keywords: C-10, multidimensional scaling, factor analysis, test reliability, intelligence, asymmetry of g

Exploring the Internal Structure of the C-10

The C-10 is a liberally timed test that consists in three subtests, each one prepared with 20 items (for examples of items that make up the C-10, see Table 1). It aims at assessing cognitive ability in gifted individuals, most likely in adults and in older adolescents for whom traditional scales may be of too restrictive ceiling. Problems were written so that they were expected to be sufficiently discriminative in such individuals.

The first subtest (ANA) to be taken by the examinee is built with verbal analogies. This type of item is known to appropriately measure reasoning. According to Carroll (1993), analogies are deductive and inductive reasoning variables frequently classed as part of sequential reasoning (RG) in factor analyses. RG is one of the three main factors in the reasoning domain. Additionally, when using relatively infrequent words in the test problems, verbal analogies refer to verbal language comprehension (V; Carroll, 1980). For the aforementioned reasons, he also notes that analogies tests tend to correlate highly with tests of general intelligence (g). Both RG and V factors are respective major's components in the two other subtests of the C-10.

Inspired by Jouve's *Epreuve de Performance Cognitive* (2005), numerical sequences of the second subtest (SEQ) were prepared so that acquired knowledge in mathematics is reduced to bare minimum. Apart from primary-school level basic arithmetic, such as addition and subtraction, numbers are chiefly used as symbols and might have been replaced by drawings, or figures. This kind of items is governed by reasoning. Number series usually appeared as part of RG, induction (I) and even quantitative reasoning (RQ) in factor analytic studies (Carroll, 1993 p. 214). They give a good figure of a wide range of abilities in the domain of reasoning. In case of the C-10, it is believed that the quantitative reasoning load in problems is not significant thanks to the carefulness given to item preparation. C-10 number series completion subtest did

not aim at measuring a relatively advanced level of quantitative ability and attention was given not to introduce RQ as an artefact in the final measure. That is to say, C-10 sequences may produce a measure of RG and I. These are principal constituents of fluid intelligence (Gf), which according to Gustaffson (1984), is almost identical to factor g.

As in nearly all synonym search or more generally, vocabulary knowledge tasks, word similarities like those of the last C-10 subtest (SIM), are among the most frequently occurring variables measuring factor V (Carroll, 1993). In some studies evidence exists that tests of vocabulary knowledge can be regarded as a different trait than language development. This occurs more often in samples of adults as a consequence of the ability of some individuals to attain through experiences an advanced level of vocabulary. In such cases, although still very close to V, knowledge of vocabulary can be best described as a factor of lexical knowledge (VL; Huteau, Lautrey, 1999). As a key latent trait in the domain of language, extend of vocabulary is strongly related to both crystallized intelligence (Gc) and general intelligence.

According to what we know about the tasks included in the C-10, a reasonable expectation was to observe a significant trace of a link between each subtest. This must imply significant correlations between subtests, appropriate reliability of scores on the entire set of items, and a continuum from the easiest to the hardest item. Besides, the existence of a bi-factorial structure could be considered, with on one side a measure of reasoning, and on the other, a measure of verbal related area. The later would not be surprising because of the asymmetry of g in the high range of mental ability (Evans, 2002).

In order to investigate these hypotheses, analyses were performed using different methods. Studies applied multidimensional scaling (MDS), correlation approach with inter-subtests relationships and reliability coefficient, and principal component factor analysis.

Method

Participants

The sample of participants consisted in 102 individuals. Of these, 10 (10.2%) were females, 83 (84.7%) were males. The majority of these participants attended university (73.4%); among them, 13 (13.3%) have completed a doctorate, 13 others a masters, and 7 (7.1%) indicated to have studied until completion of a bachelor degree. The remaining 39 (39.8%) did not mentioned a precise college level. Even though the C-10 was only available in English, participants were not all English natives. The cohort gathered 27 (27.5%) English, 12 Germans (12.2%), and small groups of Frenchs, Greeks, Indians, Italians, Portuguese's and Spanish's - each of these groups of examinees accounted for not more than 6% of the total sample

People who spent time taking the C-10 were for the most part, members of clubs for intellectually gifted individuals, such as Cerebrals and Mensa, which the respective cut-off for applying is the percentile rank of 99.7 and 98 on a supervised intelligence test. These levels roughly correspond to 141 and 130 of a deviation-IQ using a mean of 100 and 15 points per standard deviation.

Procedure

The C-10 was proposed as a contest on the Cerebrals Society website during a six months period. Participation was free, and all visitors were welcomed to send their answers. Instructions given to the participants were straightforward to simply answer as best as they can with permission to use references for the verbal analogies and calculators for the number series if needed. The questionnaire was available in Portable Document Format (PDF) for being easily printed along with an electronic form to submit the answers.

Results and discussion

Multidimensional Scaling has been chosen to search for dimensionalities in items of the C-10. MDS models are spatial and measurement models (Tournois, Dickes, 1993). Classically, the scaling converts, by monotone transformation, data entries to Euclidian distances, in which only the order of those entries is taken into account by the transformation. The result is a representation into an Euclidian space. The classical MDS performs an ordinal transformation of a “proximity” matrix in order to represent it with distances into a Euclidian space.

For analysing the C-10 with MDS, ALSCAL (Takane, Young, De Leeuw, 1977) has been used. The scaling is based on a matrix of similarities gathering the Pearson’s correlations between items. According to Table 2, inadequacies by dimensionality show that above 4 dimensions the improvement of Kruskal’s Stress coefficient (1964) is insignificant. The Stress is the representation of proximities between test items. As can be seen in the reported values, it is not evident to determine between a three-dimensional and a four-dimensional solution which one fits the best. Moreover, both of them lead to a residual squared correlation (RSQ) that is close to 80% of common variance. Empirically, based on the content of the three different subtests, each one consisting in a single type of items, a solution for three dimensions could be reasonability retained.

However, a third solution has to be considered because general cognitive ability (g) is supposed to be involved in every task underlining the C-10 problems. If this initial hypothesis is confirmed, a two-dimensional solution should work, at least graphically.

In examining subtests inter-correlations shown in Table 3, one can see that all three are highly correlated. Interestingly, we can notice the slightly better correlation of verbal analogies with numerical sequences (r = .74) than with word similarities (r = .72). Moreover, the series and synonyms relationship has the lowest correlation at .66. Regarding the theoretical background and the factors that are supposed to be measured, these values are very consistent: the verbal analogies are a pivot between on one hand the domain of verbal ability, chiefly represented by the word similarities, and on the other hand, the numerical sequences mainly responsible for a measure of the domain of reasoning.

Additionally, the Spearman-Brown corrected split-half coefficient (Crocker & Algina, 1986) for the full scale raw score was .97 in this sample. This indicates a strong reliability of scores along with a low standard error of measurement (SEm = 2.62) when subtests are added all together. Reliability of scores is considered satisfactory at .80; and a value at or above .90 in magnitude is highly recommended and the most desirable for a use in cognitive assessment (Aiken, 2000; Nunnally & Bernstein, 1994; Salvia & Ysseldyke, 2001). Coefficients of this kind investigate the amount of error variance. Furthermore, they show the consistency of scores (Anastasi & Urbina, 1997), and if a set of variables reflects mostly true score. In other words, reliability coefficients indicate if the items of a test are able to yield all together with a measure that has its own meaning. Like the C-10, composite scores on general intelligence batteries such as for example the Wechsler Adult Intelligence Scale (Wechsler, 2008) usually have reliabilities of .95 or more. Scores reliability is a key aspect even for validity as the greater the reliability in scores, the greater the possibility for significant correlations with other internally consistent measures of what the test is supposed to assess (Lord, & Novick, 1968; Shields, & Caruso, 2004).

Figure 1 shows the two-dimensional scaling of C-10 items. Even though, as seen previously, the positioning inadequacies are not perfectly suitable, we can note the singularity of the representation. Items configuration appeared as a horse-shoe figure. This particular case is known as Guttman's effect (1955). Numerous illustrations have been published of similar cases (Davies, & Coxon, 1982; Kruskal, & Wish, 1978). A horse-shoe figure typically represents one dimension which can only be seen in a two-dimensional solution. Obviously, if a one-dimensional analysis is used, all the variables will be set accordingly as no perspective will be available. However, if the results of a two-dimensional classical MDS provide with a "U" structure, this happens as a consequence of the existence of a single dimension. The plots have been arranged two-dimensionally, and the half-circle represents a continuum in which all constraints of one-dimensional dissimilarities have been respected.

References

Aiken, L. R. (2000). *Psychological testing and assessment* (10th ed.). Needham Heights, MA: Allyn & Bacon.

Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice Hall.

Carroll, J. B. (1980). Measurement of abilities constructs. In *Construct validity in Psychological Assessment; Proceedings from a Colloquium on theory and application in education and employment* (pp. 23-41). Princeton, NJ: Educational Testing Service.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor analytic studies*. New York: Cambridge University Press.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Fort Worth, TX: Harcourt Brace Jovanovich College.

Davies, P. M., & Coxon, A. P. M. (1982). *Key texts in multidimensional scaling*. London: Heinemann.

Evans, M. G. (2002). The implications of the asymmetry of g for predictive validity. In Werner Auer-Rizzi, Erna Szabo, & Caecilia Innreiter-Moser (Eds.), *Management in einer Welt der Globalisierung und Diversitaet: Europaeische und nordamerikanische Sichtweisen* (pp. 433-441). Stuttgart: Schaeffer-Poeschel Verlag.

Gustaffson, J. E. (1984). A unifying model for the structure of intellectual abilities. *Intelligence, 8*, pp. 179-203.

Guttman, L. (1955). A generalized simplex for factor analysis. *Psychometrika, 20*, pp. 173-192.

Huteau, M., & Lautrey, J. (1999). *Evaluer l'Intelligence: Psychométrie cognitive*. Paris: Presses Universitaires de France.

Jouve, X. (2005). *Epreuve de Performance Cognitive - Forme Courte, EPC-C: Manuel*. Paris: Les Editions du Centre de Psychologie Appliquée.

Kruskal, J. B. (1964). Non metric multidimensional scaling: a numerical method. *Psychometrika, 29*, pp. 115-129.

Krusal, J. B., & Wish, M. (1978). *Multidimensional scaling*. Sage Publications: London.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

Shields, A. L., & Caruso, J. C. (2004). A reliability induction and reliability generalization study of the CAGE questionnaire. *Educational and Psychological Measurement, 64*, pp. 254-277.

Takane, Y., Young, F.W., & De Leeuw, J. (1977). Nonmetric individual differences multidimensional scaling: an alternative least squares method with optimal scaling features. *Psychometrika, 42*, pp. 7-67.

Tournois, J., & Dickes, P. (1993). *Pratique de l'échelonnement multidimensionnel: De l'observation à l'interprétation*. Bruxelles: De Broeck Université.

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale-Fourth Edition: Technical and interpretive manual*. San Antonio, TX: Pearson Assessment.

Table 1

Examples of C-10 items

Subtest	Item	Answer
ANA	If HEAD relates to HAT, then FOOT relates to (?)	SHOE
SEQ	1, 10, 100, (?), (?)	1000, 10000
SIM	(A _ I K L N T) means nearly the same as SPEAKING	TALKING

Note. ANA = Verbal Analogies; SEQ = Numerical Sequences; SIM = Word Similarities.

Table 2

Items Positioning Inadequacies According to the Number of Dimensions

Dimensions	Stress	RSQ
1	.48	.42
2	.31	.64
3	.21	.76
4	.15	.84
5	.12	.88
6	.10	.91

Note. Stress = Kruskal's Stress Formula 1; RSQ = Residual Squared Correlation.

Table 3

Correlations Between C-10 Subtest Raw Scores

Subtest	ANA	SEQ	SIM
ANA	1.00		
SEQ	.74	1.00	
SIM	.72	.66	1.00

Note. ANA = Verbal Analogies; SEQ = Numerical Sequences; SIM = Word Similarities.

Table 3

Principal Components Factor Loadings of C-10 Subtests

Subtest	Factor g(β)	Factor 1 (RG)	Factor 2 (V)
ANA	.92	.71	.57
SEQ	.89	.91	.32
SIM	.88	.36	.92

Note. ANA = Verbal Analogies; SEQ = Numerical Sequences; SIM = Word Similarities.

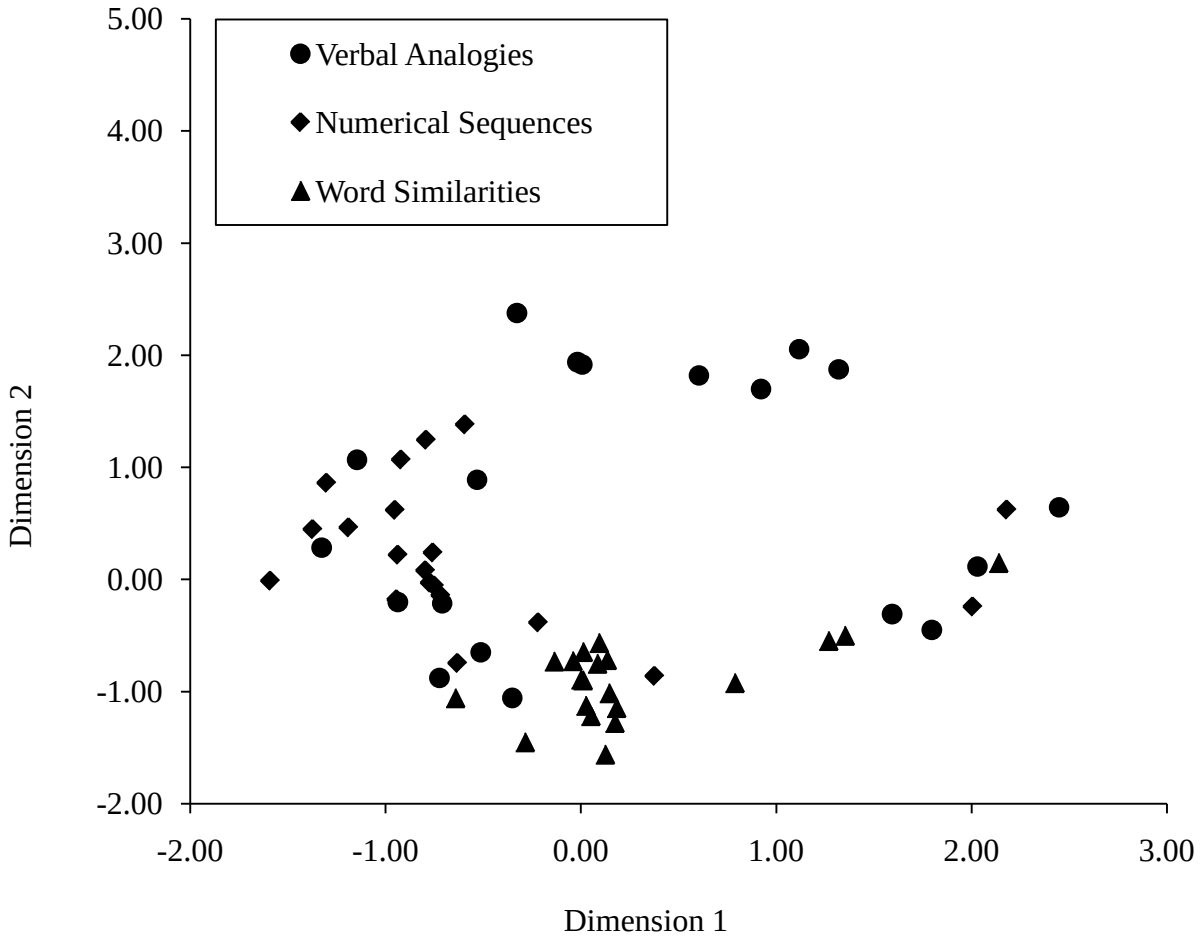

Figure 1. Two-dimensional scaling of the C-10 items.