Journal of the International Neuropsychological Society (2007), **13**, 90–98.
Copyright © 2007 INS. Published by Cambridge University Press. Printed in the USA.
DOI: 10.1017/S1355617707070129

The relationship between general intellectual ability and performance on ecologically valid executive tests in a severe brain injury sample

RODGER Ll. WOOD[1] AND CHRISTINA LIOSSI[2]

[1]Brain Injury Research Group, Department of Psychology, School of Human Sciences, Swansea University, Swansea, United Kingdom
[2]School of Psychology, University of Southampton, Southampton, United Kingdom

(RECEIVED December 23, 2005; FINAL REVISION August 10, 2006; ACCEPTED August 10, 2006)

Abstract

Recent studies of brain injured and healthy individuals have provided empirical support for the theoretical proposition that executive function and general intelligence are closely associated by demonstrating that performance on tests of executive function is correlated with general intellectual ability (*g*). In the present investigation, the relationship between performance on the Wechsler Adult Intelligence Scale-III (WAIS-III), as a measure of *g*, and performance on recently developed ecological tests of executive function [i.e., Hayling and Brixton, Zoo Map and Key Search sub-tests from the Behavioral Assessment of the Dysexecutive Syndrome (BADS) battery], was examined in a sample of 118 severely brain injured individuals. The results indicated that (a) performance on tests of executive function share significant variance, and (b) a proportion of that shared variance is associated with performance on the WAIS-III. Correlations between conventional measures of executive function (i.e., Trails B and Controlled Oral Word Association) and WAIS-III were of comparable magnitude to the correlations between new, ecologically valid executive tests and WAIS-III. The results provide some support to the notion that tests of executive function measure non-specific intellectual functions, reminiscent of *g*. (*JINS*, 2007, *13*, 90–98.)

Keywords: Fluid intelligence, Crystallized intelligence, Executive function, Brain injury, Test validity, Ecological validity

INTRODUCTION

In 1927 Spearman argued that all mental tasks share a common factor, best described as general intelligence (*g*), based on some biological property of the nervous system. More recently, Jensen and Weng (1994) argued that *g* outweighs all other psychometric factors when predicting the outcomes of job training, job performance, and educational achievement. However, following traumatic brain injury there is a poor relationship between *g*, as measured by standardized tests of cognitive ability and the application of that ability in everyday problem solving activities (Burgess, 2000; Wood & Rutterford, 2004). Many patients with frontal brain damage may perform adequately on standardized tests of intelligence yet be markedly impaired on tasks that require quick thinking or a novel approach. To some extent, this dislocation in performance may be explained by the division of *g* into crystallized and fluid intelligence (Horn & Cattell, 1966). Crystallized intelligence refers to facts, procedures and skills that someone has acquired over their lifetime, a repertoire of information for dealing with familiar problems. In contrast, fluid intelligence deals with new and unusual problems that require mental speed and cognitive flexibility, reflecting the kind of abilities needed in everyday activities. Tests of intelligence, such as the Wechsler Adult Intelligence Scale-Revised (WAIS-R) (Wechsler, 1981) largely assess aspects of crystallized intelligence (Sternberg, 2000). However, some authorities argue that Spearman's *g* relies on *fluid intelligence*, a construct closely related to that of executive function, which is described by Temple (1997) as a "meta" cognitive activity that controls and integrates other cognitive activities. Executive func-

Correspondence and reprint requests to: Dr. Christina Liossi, School of Psychology, University of Southampton, Highfield, Southampton SO17 1BJ, United Kingdom. E-mail: cliossi@soton.ac.uk

https://doi.org/10.1017/S1355617707070129 Published online by Cambridge University Press

tions deal with novelty, planning and implementing strategies, monitoring performance, using feedback to adjust future responding, vigilance, and inhibiting task-irrelevant information (Rabbitt, 1997).

Tests that measure fluid intelligence seem to be more sensitive to frontal lobe dysfunction than conventional intelligence tests. Duncan et al. (1995) compared the performance of four patients with frontal brain damage, on a short version of the WAIS-R (as a measure of crystallized intelligence) and Cattell's Culture Fair Test (CFT) (Catell, 1973), as a fluid intelligence measure of *g*. They found a substantial discrepancy, with performance on the CFT being 22–38 points lower than on the WAIS-R. The same discrepancy was not found among a control group of five patients with posterior lesions. Duncan et al. (1997) conducted a further examination of the relationship between *g* and executive function, this time using 90 severely head injured patients. They administered conventional executive tests (CONVEX), such as the Wisconsin Card Sorting (Milner, 1963), Verbal Fluency (Benton & Hamsher, 1976), List Learning (Luria, 1966) and a spatial puzzle (Luria, 1966), plus a range of other cognitive tests not normally used to examine executive ability (e.g., recognition memory, digit span, object naming). Only modest correlations were found between: (a) the CONVEX tests and non-executive tests (median correlation .29); (b) CONVEX tests and ratings of everyday behavior (median correlation .12); and (c) between CONVEX tests themselves (median correlation .26). The authors concluded that the only common element shared by the tests was closely related to *g* and to a phenomenon they called "goal neglect" (disregard or neglect of a known task requirement. This can also be seen in normal controls under conditions of novelty and weak error feedback—see Duncan 1995).

Obonsawin and colleagues (2002) explored the relationship between general intelligence (measured by WAIS-R) and performance on conventional tests of frontal lobe functioning in 123 *healthy* individuals. The executive tests they employed were measures of verbal fluency (Benton & Hamsher, 1976), Modified Card Sorting Test (Milner, 1963), the Stroop test (Stroop, 1935), the Tower of London test (Shallice 1982), Cognitive Estimation (Shoqeirat et al., 1990), and Paced Auditory Serial Addition task (Gronwall & Wrightson, 1974). They found that performance on all neuropsychological tests of frontal lobe function correlated significantly with performance on the WAIS-R. Many of the frontal tests correlated with one another but only weakly (0.21–0.46), supporting the findings of Duncan et al. (1997). It therefore seems that general intellectual ability contributes substantially to performance on neuropsychological tests sensitive to frontal lobe dysfunction.

Previous studies (Duncan et al., 1995; Duncan et al., 1997; Obonsawin et al., 2002) have used the WAIS-R to provide a measure of *g*. This has now been replaced by the WAIS-III, which contains new subtests that enhance the measurement of fluid intelligence (Matrix Reasoning, Symbol Search, and Letter-Number Sequencing). Also, tests that assess executive ability are changing because many traditional tests have failed to capture executive skills that feature in everyday functional tasks. One reason for this is that many traditional tests have a structure that minimizes their executive element. In some cases performance is monitored by the examiner who may inadvertently act as an individual's external executive control, further minimizing the demands on an individual's executive ability. Tests that have been developed more recently make demands on thinking that approximate those made on the individual in the everyday world and therefore, to successfully perform on such tasks, individuals need to draw extensively on their executive functions. Because of this, ecologically valid tests have the potential to be better measures of executive function, and overlap with measures of general intelligence to a different degree than traditional executive tests. If this hypothesis is true then the relationship between conventional tests of executive function and measures of general intelligence, and the relationship between ecologically valid measures of executive function and general intelligence, should differ, with the latter being weaker.

The aim of this study therefore, was to determine the degree to which a measure of *g*, based on the WAIS III, was associated with performance on recently developed ecological tests of executive ability (i.e., Hayling and Brixton tests; Burgess & Shallice, 1997), Zoo Map and Key Search subtests from the behavioral assessment of the Dysexecutive Syndrome (BADS, Wilson et al., 1996). In addition, two established tests, Trails B (Reitan & Wolfson, 1985) and the Controlled Oral Word Association Test (COWAT, Benton & Hamsher, 1976), were included as key reference measures to allow comparisons with the ecologically valid tests. These two tests are generally believed to be good measures of executive function and are frequently used in clinical and experimental studies investigating executive function and dysfunction. A number of predictions drawn from Duncan's hypothesis (Duncan, 1995) and previous empirical literature were tested in the present study: (1) there would be a significant positive correlation between measures of executive function and performance on WAIS-III; (2) there would be a significant positive correlation between different measures of executive function; (3) correlations between conventional measures of executive function (i.e., Trails B and COWAT with the WAIS-III) would significantly exceed the correlations between ecologically valid executive tests and WAIS-III performance measures; and (4) measures of general cognitive ability would account for the shared variance between tests of executive function.

METHODS

Ethical approval for this study was obtained from the Research Ethics Committee of the Department of Psychology, Swansea University. All participants gave their informed consent prior to their inclusion in the study, in compliance

https://doi.org/10.1017/S1355617707070129 Published online by Cambridge University Press

with regulations of the institution, and the guidelines of the Helsinki Declaration.

Participants

From 220 consecutive cases referred between 2001 to 2004 to the Brain Injury Research Group at Swansea University for neuropsychological assessment, 118 satisfied the present study's inclusion and exclusion criteria which were as follows: Inclusion criteria: (a) serious closed head injury based on Glasgow Come Score < 12 (Teasdale & Jennett, 1974) and Post Traumatic Amnesia (PTA) >24 hours (McMillan et al., 1996); (b) English as first language; (c) aged 18–60 years (d) willing to give consent for the assessment. Exclusion criteria included: (a) previous history of head injury, neurological, or psychiatric disorder, (b) alcohol or drug abuse, (c) speech, motor or perceptual deficit likely to interfere with neuropsychological assessment. Of the cases admitted to the study, the majority was male (84, 71.2%). Mean age at assessment was 35.47 years (*SD* = 13.54), mean Glasgow Coma Score 8.4 (*SD* = 3.97), mean PTA 17.31 days, (*SD* = 3.65) and mean time since injury 2.9 years (*SD* = 1.9). Ethnically, the sample was overwhelmingly White British (115 out of 118) with three being of Asian origin. 92 had abnormal CT scans, indicating frontal hemorrhagic or contusional injuries: left frontal injury = 33, right frontal = 41, bi-lateral injury = 18. All cases, even those with "normal" CT scans, were judged by referring clinicians to have suffered diffuse axonal or frontal contusional injuries, based on executive and behavior problems exhibited after injury.

Measures

Patients were administered the following standardized tests by the first author, a clinical neuropsychologist:

The Wechsler Adult Intelligence Scale (3rd edition)

The WAIS III (Wechsler et al., 1998) is a standard clinical measure of intellectual functioning that provides scores for verbal and performance intelligence. Full Scale (FSIQ), Verbal Scale (VIQ), and Performance Scale (PIQ) scores were used in the analysis of data. WAIS-III has good reliability and validity in clinical populations (Tulsky et al., 1997). All participants received the full version of the WAIS III.

The Hayling & Brixton Tests

The *Hayling Test* (Burgess & Shallice, 1997) measures response completion and the ability to suppress inappropriate responses. Patients have to complete 30 sentences from which the last word is omitted. In the first half of the test (initiation condition) they complete sentences with a sensible word. In the second half of the test (inhibition condition) they have to supply a word that makes no sense in the context of the sentence (e.g., London is a very busy . . . *banana*). To achieve this it is necessary to inhibit the dominant (automatic) response before generating the new one. In the initiation section of the test the total response time for all 15 items is the performance measure (Hayling A). In the inhibition section, two measures are derived: the total time to respond to the items (Hayling B) and the total number of errors (Hayling C), where an error is a response that is related to the sentence. All three measures are expressed as scaled scores. The *Brixton Test* is a measure of rule attainment and rule detection. The patient is shown a 56-page stimulus book, one page at a time. All pages contain 10 circles in the same basic array, one circle is colored blue, all others are white. The patient has to predict where the blue circle will be on the next page, based on the sequence of previous pages. The total number of errors made on the task is summed and converted to scaled scores. The Hayling and Brixton Tests are mainly sensitive to left hemisphere damage. Using functional imaging techniques to identify cortical areas responsible for verbal initiation and suppression in normal subjects, left frontal activation was predominant on the Hayling Test (Nathaniel-James et al., 1997; Collette et al., 2001). The Brixton Test is associated with inductive reasoning and working memory, mediated by the left lateral frontal cortex (Reverberi et al., 2005). Both measures have good psychometric properties (*Hayling:* Split-half reliability for Hayling A = 0.93, Hayling B = 0.80, Hayling C = 0.72; Test-retest reliability for Hayling A = 0.62, Hayling B = 0.78, Hayling C = 0.52. *Brixton:* Split-half reliability 0.62; Test-retest reliability 0.71).

The Zoo Map and Key Search Tests

These two subtests from the Behavioral Assessment of Dysexecutive Syndrome (BADS) battery (Wilson et al., 1996) were designed to simulate real life situations. In the Zoo Map, participants are required to show how they would visit a series of designated locations, while following certain rules. The map and the rules are constructed so that there are only four variations on a route that do not infringe the rules. The first trial represents a high demand version of the task in which the planning abilities of the patient are rigorously tested. In the second, low demand, trial, the patient is simply required to follow written instructions. On both trials, the patient is expected to minimize errors by modifying performance on the basis of feedback. The Key Search test is more abstract and examines a person's ability to prepare an efficient plan of action in the context of a routine event. The patient is asked to imagine that a 100-mm square on an A4 size paper is a large field, in which they have lost their keys. They are asked to draw a line, beginning at a black dot, 50 mm below the square, to show the strategy they would use to search the field, to make absolutely certain that they would find their keys. The test evaluates a person's ability to monitor and evaluate their own performance, taking into account factors that are not explicitly stated in the instructions. Raw scores on each test were converted to profile scores for analysis, as per the BADS

https://doi.org/10.1017/S1355617707070129 Published online by Cambridge University Press

manual. Inter-rater reliability across the BADS tests ranges from .88–1.00.

Conventional Executive Tests

Eslinger and Grattan (1993) drew a distinction between two forms of frontal lobe function: reactive and spontaneous flexibility. Reactive flexibility was defined as a readiness to shift cognition and behavior freely, in response to changing demands of an external situation. Spontaneous flexibility describes the ready flow of ideas and answers in response to a question. In this study, both parts of the Trail Making Test (TMT) were administered but only Trails B was used as an exemplar of reactive flexibility; a fluency task (Controlled Oral Word Association) was used as a typical measure of spontaneous flexibility.

The Trail Making Test

The TMT (Reitan & Wolfson, 1985) is a measure of scanning, visuomotor tracking, divided attention, and cognitive flexibility involving connection of ascending order consecutive numbers (1-2-3 . . .) randomly arranged on a page (part A) and of numbers and letters in alternating order (1-A-2-B . . .) (part B). Trails B is a measure of executive function because the individual is required to keep the number and letter sequences in working memory so as not to lose place. Recent fMRI studies (Zakzanis et al., 2005) have confirmed the sensitivity of TMT to frontal regions in the left hemisphere. Time for completion was used as the primary score for each measure in this study. TMT has good reliability (0.67) and adequate construct validity (.36–.93).

Controlled oral word association

COWAT (Benton & Hamsher, 1976) was administered as described by Benton and Hamsher (1976) and Goodglass and Kaplan (1972). Each participant was asked to produce orally as many words as possible beginning with the letters F, A, S, allowing 60 seconds for each letter. The score is the total number of words correctly produced over the 3-min period. COWAT is considered a "frontal" task because the organization of words by first letter is unfamiliar, and requires conscious, effortful, systematic organization, and the filtering of irrelevant information such as natural taxonomic categories. This test reflects abstract mental operation related to problem solving, sequencing, resisting distractions, intrusions, and perseverations. Impairment on the task is associated with the lateral, or the inferior prefrontal region, especially of the left hemisphere (Stuss et al., 1998). COWAT has good retest reliability (at one year 0.70). Concurrent validity with educational level is 0.32 and with WAIS VIQ, 0.14.

Statistical Analysis

Previous studies of this nature (e.g., Duncan et al., 1997) have reported correlations between *g* and executive function in the region of 0.30. *A priori* power analysis ascertained that 88 participants would be needed to detect a correlation of 0.30 with a significance level of 0.05 (two tailed) and power set at 0.80. Given that a factor analysis was also planned it was decided to recruit over 100 individuals with brain injury. Correlations using Pearson product moment correlation coefficient *r* were used to investigate the relationship between different measures of executive function as well as between measures of executive function and general intelligence. Trails A and B scores were reflected (multiplied by −1) before correlational analysis to ensure consistency of sign. A positive correlation always indicates that better performance on one test was associated with better performance on the other. Difference between correlations of ecologically valid tests and *g*, and conventional executive tests and *g* were examined with William's (1959) test for non-independent correlations (one-tailed).

RESULTS

Means (±SD) for all neuropsychological tests administered are presented in Table 1. The sample overall performed in the low to moderate average region on WAIS III, Hayling and Brixton, COWAT, and Trails B. Performance on the Key Search, Zoo Map, and Trails A tests was in the impaired or borderline-impaired range.

Performance on all neuropsychological tests of executive function correlated significantly with performance on the WAIS-III FSIQ, and PIQ. All executive tests, except Trails B, correlated with VIQ. Most of the correlations between executive tests were significant except the following: Trails B *versus* COWAT, Zoo Map, Brixton, Hayling B, and Hayling C; and COWAT *versus* Hayling C and Zoo Map (Table 2).

Table 1. Mean (±*SD*) for all neuropsychological tests administered ($N = 118$)

Test	Mean (*SD*)
WAIS-III (FSIQ)	88.76 (11.89)
WAIS-III (VIQ)	88.5 (13.17)
WAIS-III (PIQ)	88.95 (12.28)
Hayling A	4.81 (1.31)
Hayling B	5.33 (1.23)
Hayling C	5.37 (2.32)
Brixton	5.96 (1.91)
Zoo Map	2.71 (1.17)
Key Search	2.63 (1.21)
Trails A	−50.24 (23.97)
Trails B	−78.21 (42.18)
COWAT	27.5 (10.28)

Note. Hayling A, B, C and Brixton are reported as scaled scores; Zoo Map and Key Search are reported as profile scores; Trails A and B scores are reported as raw scores and are reflected (i.e., multiplied by −1); COWAT is reported as a raw score.

https://doi.org/10.1017/S1355617707070129 Published online by Cambridge University Press

Table 2. Pearson *r* correlations between performance on tests of executive function and WAIS-III FSIQ, VIQ and PIQ in severely brain injured individuals ($N = 118$)

	WAIS-III (FSIQ)	WAIS-III (VIQ)	WAIS-III (PIQ)	Hayling A	Hayling B	Hayling C	Brixton	Zoo Map	Key Search	COWAT	+Trails A
WAIS-III (VIQ)	0.94**										
WAIS-III (PSIQ)	0.93**	0.77**									
Hayling A	0.38**	0.31**	0.39**								
Hayling B	0.27**	0.27**	0.22*	0.59**							
Hayling C	0.28**	0.31**	0.21*	0.24**	0.55**						
Brixton	0.25**	0.27**	0.21*	0.37**	0.56**	0.35**					
Zoo Map	0.35**	0.31**	0.34**	0.29**	0.26**	0.34**	0.11				
Key Search	0.41**	0.39**	0.39**	0.36**	0.53**	0.39**	0.33**	0.59**			
COWAT	0.39**	0.35**	0.41**	0.24**	0.19*	0.03	0.23*	0.16	0.27**		
+Trails A	0.21*	0.17	0.25**	0.27**	0.08	−0.016	−0.08	0.02	0.12	0.18	
Trails B	0.21*	0.17	−0.23*	0.26**	0.17	0.05	0.05	0.15	0.20*	0.13	0.76**

Note. +Although not an executive test Trails A is included for completeness purposes.
**Correlation is significant at the 0.01 level (2-tailed).
*Correlation is significant at the 0.05 level (2-tailed).

After a Bonferroni correction for multiple comparison error ($\alpha = 0.004$) there was no significant difference between correlations of ecologically valid tests and *g*, and conventional executive tests and *g* (see Table 3)

When the executive tests correlation matrix (Table 2) was covaried with WAIS-III FSIQ (FSIQ has been used as a measure of *g* in previous investigations), all correlations decreased (except that between Brixton *vs*. Hayling B, which increased, albeit, modestly) and few remained significant, most of them between subcomponents of the same test (Table 4). Therefore, the shared variance on the scores for these neuropsychological tests is mostly accounted for by performance on the WAIS-III FSIQ. The fact that some correlations remained significant could probably be attributed to method variance (i.e., variance produced by the methodological approach chosen to assess the ability).[a] Another possibility relates to the fact that WAIS-III scores are age corrected, unlike the scores for Hayling and Brixton and Key Search and Zoo map. This could weaken the effect compared to what it would be if a raw WAIS score was available and partialled out.

A principal components analysis was carried out with the scores from all tests of executive function (Table 5). Varimax rotation yielded 2 factors (eigenvalue >1.0) that accounted for 52.99% of the total variance. Scores from Hayling A, Hayling B, Hayling C, and Brixton loaded highly on the first factor. Scores from Zoo Map, Key Search, and Trails B loaded highly on the second factor. COWAT factor scores loaded almost equally on the two factors. The score for factor 1 correlated significantly with performance on the WAIS-III FSIQ, VIQ, and PIQ. The score for factor 2 again correlated significantly, but less than did factor 1, with performance on WAIS-III FSIQ, VIQ, and PIQ (FSIQ *vs*. Factor 1: 0.33**, VIQ *vs*. Factor 1: 0.33**, PIQ *vs*. Factor 1: 0.28, FSIQ *vs*. Factor 2: 0.28**, VIQ *vs*. Factor 2: 0.25**, PIQ *vs*. Factor 2: 0.27).[b]

Table 3. Differences between correlations of ecologically valid tests with FSIQ, and conventional executive tests with FSIQ ($N = 118$, $df = 115$)

	t	*p*
Hayling A *versus* Trails B	1.6	0.05
Hayling A *versus* COWAT	0.09	0.46
Hayling B *versus* Trails B	0.52	0.3
Hayling B *versus* COWAT	1.1	0.14
Hayling C *versus* Trails B	0.78	0.22
Hayling C *versus* COWAT	0.9	0.18
Brixton *versus* Trails B	0.32	0.37
Brixton *versus* COWAT	1.3	0.09
Zoo Map *versus* Trails B	1.21	0.11
Zoo Map *versus* COWAT	0.36	0.36
Key Search *versus* Trails B	1.82	0.03
Key Search *versus* COWAT	0.19	0.42

DISCUSSION

This study aimed to establish relationships between newly developed, ecologically valid tests of executive function

[a]Method variance refers to variance that is attributable to the measurement method rather than to the constructs the measures represent. Thus, a measure's variance is composed of true variance, systematic variance and random error variance. For a comprehensive discussion of common method biases in behavioral research see Podsakoff et al. (2003).

[b]**. Correlation is significant at the 0.01 level (2-tailed).

https://doi.org/10.1017/S1355617707070129 Published online by Cambridge University Press

Table 4. Partial correlations between performance on tests of executive function, controlling for WAIS-III FSIQ, in severely brain injured individuals ($N = 118$)

	Hayling A	Hayling B	Hayling C	Brixton	Zoo Map	Key Search	COWAT
Hayling B	0.55**						
Hayling C	0.15	0.52**					
Brixton	0.40**	0.53**	0.30**				
Zoo Map	0.18	0.18*	0.27**	0.02			
Key Search	0.24**	0.47**	0.31**	0.26**	0.52**		
COWAT	0.10	0.090	−0.09	0.15	0.03	0.13	
Trails B	−0.05	0.11	0.03	0.12	0.23*	0.22**	0.002

Note. **Correlation is significant at the 0.01 level (2-tailed)
*Correlation is significant at the 0.05 level (2-tailed)

(Hayling and Brixton, and Zoo Map and Key Search) and a measure of *g*, based on the WAIS-III FSIQ. Trails B and COWAT were used as reference measures, to allow comparisons with ecological tests of executive ability. In line with previous studies using brain injured patients (Duncan et al., 1995; Duncan et al., 1997) and healthy individuals (Obonsawin et al., 2002), performance on all neuropsychological tests of executive function correlated significantly with performance on the WAIS-III FSIQ, and PIQ. Similarly, most executive tests correlated among themselves. The matrix of positive but predominantly low correlations, with little substantial clustering of executive tasks, suggests that many ecologically valid as well as conventional executive tasks have as little in common between them as they have in common with non-executive tasks. Thus, if a distinct executive control mechanism exists, variation in its efficiency is not easily discerned as a systematic, replicable pattern of associations between mean performance scores on current neuropsychological tasks. It is important to point out, however, that the tasks considered to tap similar executive functions (e.g., Key Search and Zoo Map) tended to show moderate correlations with one another, whereas not correlating as strongly with the tasks considered to tap other executive functions, therefore showing signs of convergent and discriminant validity. Based on evidence presented earlier, it seems reasonable to conclude that valid assessment of executive functions requires the use of batteries of neuropsychological tests, each evaluating different cognitive processes. However, a note of caution; although the pattern of correlations between executive tasks can be taken to infer the existence of a fractionated executive system, each implementing a distinct executive function, Della Sala and colleagues (1998) have pointed out that such a conclusion is precluded by the theoretical (i.e., task impurity problem) and methodological limitations of the correlational and associated factor-analytic approach (i.e., difficulty in interpreting what constructs different factors obtained in many factor analytic studies really represent).

Table 5. Rotated factor matrix (Varimax rotation) from principal components analysis of tests of executive function scores ($N = 118$)

	Factor 1	Factor 2
Hayling A	0.69	0.04
Hayling B	0.88	0.17
Hayling C	0.70	0.26
Brixton	0.72	0.02
Zoo Map	0.23	0.79
Key Search	0.48	0.66
Trails B	−0.09	0.71
COWAT	0.15	0.12

When we tested the possibility that associations between tests of executive function are mediated by intellectual ability, by covarying the correlations with performance on the WAIS-III, all correlations decreased and few of them remained significant. This suggests that shared variance on scores for executive tests can be accounted for, at least to some extent, by performance on the WAIS-III FSIQ. Two issues are raised by empirical evidence suggesting that executive functioning is related to intelligence. First, the relationship between a measure of general intellectual ability and performance on a test of executive function provides support for Duncan's hypothesis (Duncan, 1995) that components of general intellectual ability are measured by tests of executive function. Secondly, the study clearly demonstrates that although performance on executive tests is sensitive to intellectual functioning, this does not in any way account for the full variance of performance on these tests.

The results of the factor analysis indicate that executive tests load on two factors; the most prominent factor, which demonstrated the higher correlation with FSIQ, can be considered to reflect *g*, whilst the less prominent factor, which demonstrated the lower correlation with FSIQ, probably reflects the specific executive component of the tests (i.e., *e*). It is therefore interesting that some of the recently developed, ecologically valid tests of executive function (i.e., Zoo Map and Key Search) that simulate real life situations, loaded highly on the second factor (i.e., they showed high "*e*"), reinforcing the impression that they have strong

https://doi.org/10.1017/S1355617707070129 Published online by Cambridge University Press

executive requirements. Similarly Trails B, a well-recognized test of executive function, also loaded highly on factor two (i.e., showed high "*e*"). However, the situation was different for the remaining tests (i.e., Hayling A, Hayling B, Hayling C, and Brixton). These loaded highly on the first factor (i.e., they showed high "*g*"). In contrast to the findings of Obonsawin and colleagues (2002), COWAT loaded almost equally on the two factors. However, differences between the samples prohibit a direct comparison.

Despite the fact that some ecological tests had a high *e* loading in the factor analysis, correlations between conventional measures of executive function (i.e., Trails B and COWAT and WAIS-III, did not significantly exceed the correlations between ecologically valid executive tests and WAIS-III performance), therefore, not supporting the relevant hypothesis. This finding probably reflects (a) the general problems associated with measurement of executive function (Rabbitt et al., 1997), (b) the documented limitations of the ecological validity of specific tests (Odhuba et al., 2005; Wood & Liossi, 2006), and (c) the use of a coarse-grained scoring method (i.e., stern scores; the difference between stern scores correspond to 0.5 of an *SD*) (Crawford & Henry, 2005).

The results of this investigation need to be treated cautiously. Individuals with severe brain injury provide a good population to study the relationship between *g* and executive function because head trauma usually results in executive problems and injuries are mostly diffuse, which provides variability in the abilities under investigation, necessary for a correlational study. However, at the same time, the heterogeneity of the sample in terms of age and severity of injury (with its associated intellectual impairment) is potentially problematic. Theories of factor differentiation have existed since 1927, when Spearman postulated that the factor structure among cognitive abilities varies with age (Spearman, 1927). An elaboration of differentiation theory, described as the *ability differentiation hypothesis* by Reinert and colleagues (1965), further proposes that the factor structure among intellectual abilities differs according to level of intellectual ability. Therefore, the results of the present investigation may not stand with older populations, or populations that have much lower or higher levels of functioning. However, it needs to be noted that recent reviews (Carroll, 1993) and empirical investigations (Anstey et al., 2003) have failed to find clear evidence for the differentiation of abilities according to age or intellectual ability. The lack of a control group also made it impossible to identify an overlap between mild executive dysfunction on test performance attributable to brain injury from executive difficulties that seem to exist within the general population (Chan & Manley, 2002).

It is also acknowledged that, as yet, there has not been a clear specification of the cognitive processes that underlie higher level cognition, to explain how higher-level cognitive processes are related to attention and working memory. Various authors have proposed that *g* is related to properties of working memory, such as simple memory capacity or ability to exclude distracting or task irrelevant information (Ackerman et al., 2005). Because research on executive performance in relation to general intelligence progresses, the taxonomies of intelligence and executive function should become clearer, as will the interactions and dissociations among them. Further research is needed to elucidate the nature and extent of individual differences in these cognitive functions. This should increase our understanding of test components, as well as the variety of factors that contribute to impaired performance.

The relationship between intelligence and executive function is obscure. The field still lacks a compelling theory and operational definitions of executive functions and intelligence. This state of affairs, makes Deary's (2005) observation that "these associations are often correlations between two unknowns" (pp. 226) all more relevant. The covariance among test scores however, is "real." Sternberg (2005) postulated that the best way to understand intelligence is through converging operations. Towards this direction, this study, along with previous similar investigations (Duncan et al., 1995; Duncan et al., 1997; Obonsawin et al., 2002) and recent neuroimaging studies that demonstrate the role of the dorsolateral prefrontal cortex in performing fluid intelligence (Gray et al., 2003) and executive tests (Duncan & Owen, 2000) converge on a unified set of results and begin to show what Sternberg (2005) wished for, that the "results transcend the methods used to elicit them" (p. 244).

ACKNOWLEDGMENTS

The authors thank Prof John Crawford and Dr John Duncan for their helpful comments on an earlier version of this paper. The information in this manuscript and the manuscript itself is new and original and is not currently under review by any other publication, and has never been published either electronically or in print. The authors have no financial relationships or conflict of interest to disclose.

REFERENCES

Ackerman, P.L., Beier, M.E., & Boyle, M.O. (2005). Working Memory and Intelligence: The Same or Different Constructs? *Psychological Bulletin*, *131*, 30–60.

Anstey, K.J., Hofer S.M., & Luszcz, M.A. (2003). Cross-sectional and longitudinal patterns of dedifferentiation in late-life cognitive and sensory function: The effects of age, ability, attrition, and occasion of measurement. *Journal of Experimental Psychology: General*, *132*, 470–487.

Benton, A.L. & Hamsher, K. (1976). *Manual for Multilingual Aphasia Examination*. Iowa City: University of Iowa.

Burgess, P.W. (2000). Strategy application disorder: The role of the frontal lobes in human multitasking. *Psychological Research*, *63*, 279–288.

Burgess, P.W. & Shallice, T. (1997). *The Hayling and Brixton Tests*. Bury St. Edmunds, UK: Thames Valley Test Company.

Carroll, J.B. (1993). *Human cognitive abilities: A survey of factor*

https://doi.org/10.1017/S1355617707070129 Published online by Cambridge University Press

analytic studies. Cambridge, England: Cambridge University Press.

Catell, R.B. (1973). *Measuring intelligence with a culture fair test*. Champagne, IL: The Institute For Personality And Ability Testing.

Chan, C.K. & Manley, T. (2002). The application of dysexecutive measures across cultures: Performance and checklist assessment in neurologically healthy and traumatically brain-injured Hong Kong Chinese volunteers. *Journal of the International Neuropsychological Society*, *8*, 771–780.

Collette, F., Van der Linden, M., Delfiore, G., Degueldre, C., Luxen, A., & Salmon, E. (2001). The functional anatomy of inhibition processes investigated with the Hayling task. *Neuroimage*, *14*, 258–267.

Crawford, J.R. & Henry, J.D. (2005). Assessment of executive deficits. In P.W. Halligan & N. Wade (Eds.), *The Effectiveness of Rehabilitation for Cognitive Deficits*, (pp. 233–245). London: Oxford University Press.

Deary, I.J. (2005). The principles of cognition and the abilities of man: A natural collaboration. *Cortex*, *41*, 225–227.

Della Sala, S., Gray, C., Spinnler, H., & Trivelli, C. (1998). Frontal lobe functioning in man: The riddle revisited. *Archives of Clinical Neuropsychology*, *13*, 663–682

Duncan, J. (1995). Attention, intelligence, and the frontal lobes. In M.S. Gazzaniga (Ed.), *Cognitive Neurosciences* (pp. 721–733). Cambridge, MA: MIT Press.

Duncan, J. & Owen, A.M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences*, *23*, 475–483.

Duncan, J., Burgess, P., & Emslie, H. (1995). Fluid intelligence after frontal lobe lesions. *Neuropsychologia*, *33*, 261–268.

Duncan, J., Johnson, R., Swales, M., & Freer, C. (1997). Frontal lobe deficits after head injury: Unity and diversity of function. *Cognitive Neuropsychology*, *14*, 713–741.

Eslinger, P.J. & Grattan, L.M. (1993). Frontal lobe and frontal striatal substrates for different forms of human cognitive flexibility. *Neuropsychologia*, *31*, 17–28.

Goodglass, H. & Kaplan, E. (1972). *An assessment of aphasia and related disorders*. Philadelphia: Lea and Fibiger.

Gray, J.R., Charbris, C.F., & Braver, T.S. (2003). Neural mechanisms of general fluid intelligence. *Nature Neuroscience*, *6*, 316–322.

Gronwall, D.M. & Wrightson, P. (1974). Delayed recovery of intellectual function after minor head injury. *Lancet*, *9*, 605–609.

Horn, J.L. & Cattell, R.B. (1966). Refinement and test of the theory of fluid and crystallized intelligence. *Journal of Educational Psychology*, *57*, 253–270.

Jensen, A.R. & Weng, L.J. (1994). What is a good *g*? *Intelligence*, *18*, 231–258.

Luria, A.R. (1966). *Higher Cortical Functions in Man*. London: Tavistock.

McMillan, T.M., Jongen, E.L.M., & Greenwood, R.J. (1996). Assessment of post-traumatic amnesia after severe closed head injury: Retrospective or prospective? *Journal of Neurology, Neurosurgery, and Psychiatry*, *60*, 422–427.

Milner, B. (1963). Effects of different brain lesions on card sorting: The role of the frontal lobe. *Archives in Neurology*, *9*, 100–110.

Nathaniel-James, D.A., Fletcher, P., & Frith, C.D. (1997). The functional anatomy of verbal initiation and suppression using the Hayling Test. *Neuropsychologia*, *35*, 559–566.

Obonsawin, M.C., Crawford, J.R., Page, J., Chalmers, P., Cochrane, R., & Low, G. (2002). Performance on tests of frontal lobe function reflect general intellectual ability. *Neuropsychologia*, *40*, 970–977.

Odhuba, R.A., van den Broek, M.D., & Johns, L.C. (2005). Ecological validity of measures of executive functioning. *British Journal of Clinical Psychology*, *44*, 269–278.

Podsakoff, P.M., MacKenzie, S.B., Lee, J., & Podsakoff, N.P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*, 879–903.

Rabbitt, P. (1997). Methodologies and Models and the Study of Executive Function. In P. Rabbitt (Ed.), *Methodology of Frontal and Executive Function*. Hove, UK: Psychology Press.

Reinert, G., Baltes, P., & Schmidt, L.R. (1965). Faktorenanalytishce Untersuchungen zur Differenzuierungshypothese der Intelligencz: Die Liestungsdifferenzierungshypothese (Factor analytic investigation of the differentiation hypothesis of intelligence: The differential performance of intelligence). *Psychologische Forschung*, *28*, 246–300.

Reitan, R.M. & Wolfson, D. (1985). *The Halstead-Reitan Neuropsychological Test Battery*. Tucson, AZ: Neuropsychology Press.

Reverberi, C., Lavaroni, A., Gigli, G.L., Skrap, M., & Shallice, T. (2005). Specific impairments of rule induction in different frontal lobe subgroups. *Neuropsychologia*, *43*, 460–472.

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London B*, *298*, 199–209.

Shoqeirat, M.A, Maize, A., MacDonald, C., Meudal, P., & Pickering, A. (1990). Performance on tests sensitive to frontal lobe lesions by patients with organic amnesia. *Journal of Clinical Psychology*, *29*, 401–408.

Spearman, C. (1927). *The Abilities of Man*. London: Macmillan.

Sternberg, R.J. (2000). *Handbook of Intelligence*. New York: Cambridge University Press.

Sternberg, R.J. (2005). The importance of converging operations in the study of human intelligence. *Cortex*, *41*, 243–244.

Stroop, J.R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643–662.

Stuss, D.T., Alexander, M.P., Hamer, L., Palumbo, C., Dempster, R., Binns, M., Levine, B., & Izukawa, D. (1998). The effects of focal anterior and posterior brain lesions on verbal fluency. *Journal of the International Neuropsychological Society*, *4*, 265–278.

Teasdale, G. & Jennett, B. (1974). Assessment of coma and impaired consciousness: A practical scale. *Lancet*, *8*, 1–4.

Temple, C. (1997). *Developmental cognitive neuropsychology*. East Sussex, UK: Psychology Press.

Tulsky, D., Zhu, J., & Ledbetter, M.F. (1997). *The WAIS III, WMS III Technical Manual*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-Revised*. New York: The Psychological Corporation.

Wechsler, D. (1998). *Wechsler Adult Intelligence Scale-Third Edition*. San Antonio, TX: The Psychological Corporation.

Wechsler, D., Wycherley, R.J., Benjamin, L., Crawford, J.R., & Mockler, D. (1998). *Manual for the Wechsler Memory Scale–Third Edition* (U.K.). San Antonio, TX: The Psychological Corporation.

Williams, E.J. (1959). The comparison of regression variables. *Journal of the Royal Statistical Society*, *21*, 396–399.

https://doi.org/10.1017/S1355617707070129 Published online by Cambridge University Press

Wilson, B.A., Alderman, N., Burgess, P.W., Emslie, H., & Evans, J.J. (1996). *Behavioural Assessment of the Dysexecutive Syndrome*. Bury St. Edmunds, UK: Thames Valley Test Company.

Wood, R.Ll. & Liossi, C. (2006). The ecological validity of executive tests in a severely brain injured sample. *Archives of Clinical Neuropsychology*, *21*, 429–437.

Wood, R.Ll. & Rutterford, N. (2004). Relationships between measured cognitive ability and reported psychosocial activity after bilateral frontal lobe injury: An 18-year follow-Up. *Neuropsychological Rehabilitation*, *14*, 329–350.

Zakzanis, K.K., Mraz, R., & Graham, S.J. (2005). An fMRI study of the Trail Making Test. *Neuropsychologia*, *43*, 1878–1886.

https://doi.org/10.1017/S1355617707070129 Published online by Cambridge University Press