A MODEL OF EXECUTIVE FUNCTIONS IN VERY OLD COMMUNITY DWELLERS: EVIDENCE FROM THE SYDNEY OLDER PERSONS STUDY

Olivier Piguet[1,2], David A. Grayson[3], Robyn L. Tate[4], Hayley P. Bennett[1,2], Tanya C. Lye[1], Helen Creasey[1], William S. Brooks[1,2] and G. Anthony Broe[2,5]

([1]Centre for Education and Research on Ageing of the University of Sydney at Concord Repatriation General Hospital, Sydney, NSW; [2]Prince of Wales Medical Research Institute and the University of New South Wales, Sydney, NSW; [3]School of Psychology, the University of Sydney, Sydney, NSW; [4]Rehabilitation Studies Unit, Department of Medicine, the University of Sydney at Royal Rehabilitation Centre Sydney, Sydney, NSW; [5]Prince of Wales Hospital, Randwick, NSW, Australia)

Abstract

Executive functions (EF) are generally described as showing greater sensitivity to ageing compared to other cognitive domains. Numerous pitfalls exist in the measurement of EF due to loose definitions and lack of agreement on these concepts and uncertainty about the constructs being measured. To this date, the validity of EF constructs has not been examined in the old-old population. Performance of 122 randomly selected community dwellers aged between 81 and 97 years on nine EF tasks (seven of which commonly used in clinical practice) was examined. Factor analytic procedures using structural equation modelling (SEM) failed to satisfactorily explain the data according to four *a priori* models, the first two models reflecting two major constructs commonly found in current models of EF ("set" and "switch"), the last two reflecting task requirements. The best measure for each task was extracted using statistically driven analyses and further SEM revealed an orthogonal two-factor model which provided a good fit of the data, explaining between 8% and 25% of the total variance. This model can be interpreted in terms of reactive and spontaneous flexibility as proposed by Eslinger and Grattan (1993), with the first factor reflecting internally driven strategies and the second environment dependent strategies. Furthermore, these findings also suggest that: (a) unique tasks of EF may not be applicable to all age groups due to individual experience and changes in strategies; and (b) current clinical instruments may be inadequate to measure very specific aspects of the complex construct of EF.

Key words: executive function, ageing, measurement, structural equation modelling

Introduction

Studies on executive functions (EF) in ageing commonly report a poorer EF performance with advancing age compared to younger age groups. These changes include decline in concept formation and abstraction, decrease in mental flexibility, increased time required to shift mental sets, to adapt to novel situations or to solve new or unusual problems (Albert et al., 1990; Daigneault et al., 1992; Daigneault and Braun, 1993; Levine et al., 1995; Mittenberg et al., 1989; Parkin and Lawrence, 1994; Robbins et al., 1998; Shimamura and Jurica, 1994; Stuss et al., 1996; West, 1996). Deficits in EF seem to increase rapidly in the eighth decade of life and to show greater age-related decline compared to other cognitive functions such as language, memory and perception (Mittenberg et al., 1989; Tranel et al., 1994). These changes have been explained by a more rapid ageing in structures thought to predominantly mediate EF, namely the frontal lobes, compared to other brain structures (Raz et al., 1997). However, this "frontal aging hypothesis" has been challenged (Greenwood, 2000). Recent studies focusing on normal cognitive ageing showed evidence of decline related to age on some (e.g., Stroop test, Tower of London, Modified card sorting test) but not other (e.g., verbal fluency, cognitive estimates, use of common objects) measures of EF in neurologically intact individuals (Crawford et al., 2000; Rabbitt and Lowe, 2000; Wecker et al., 2000).

Knowledge about EF is hampered by an absence of clearly defined abilities or processes encompassed by this construct (e.g., "judgment", "will", "self-awareness", "creativity", "error utilization", "inhibition" or "planning") and by poorly defined measurement criteria. Despite this apparent lack of cohesion, EF are generally viewed as the overarching mechanisms responsible for the control of cognition (Phillips, 1997). They are very important in aspects of behaviour that require (1) setting up plans of actions (that may be composed of subtasks) for the achievement of a pre-determined goal; (2) the initiation and monitoring of the plans; and (3) the ability to shift and adapt plans in response to external or internal pressures (Luria, 1973). Executive functions are expressed in activities as varied as problem solving (Lezak, 1995), planning (Tranel et al., 1994), cognitive estimation (Shallice, 1988), prospective memory (Burgess, 1997; Shimamura, 1995), and when mental representation of tasks and goals are required (Pennington et al., 1996). They are also present in situations requiring concept formation, reasoning, cognitive flexibility ("spontaneous" and "reactive" flexibility) (Grattan and Eslinger, 1991) inhibition, working memory or resource allocation. Most clinical EF tasks generally aim to test one of the following aspects of cognition: plan setting, shifting and monitoring, or manipulation of abstract concepts. These tasks tend to be: (a) novel, so as to invoke higher levels of cognitive activation such as novel planning and goal setting; (b) of a certain duration, to tap into the monitoring aspect of behaviour and on-line manipulation of information; and (c) effortful, in order to engage cognitive power fully and assess the levels of monitoring and inhibition processes. Complexity is not necessary, being a relative construct which may be handled by other cognitive systems (Stuss and Alexander, 2000). Some degree of unpleasantness may be present as EF tasks are designed to be effortful and challenging (Denckla, 1996; Phillips, 1997).

Unlike other cognitive abilities, EF are independent of direct external or internal sensory inputs. Because of their primary role as coordinator of purposeful actions, the correspondence between the observed behaviours and the underlying cognitive processes is uncertain (Burgess, 1997). Because of this reliance on other cognitive systems for their expression, pure measures of EF probably do not exist. As such, descriptors of task demands (e.g., "concept shifting" or "inhibition") may be useful labels but do not necessarily imply corresponding different or discrete underlying cognitive processes. Furthermore, there is evidence that EF tasks that appear, on the surface, to require similar cognitive demands do not necessarily result in a consistent performance (Lowe and Rabbitt, 1997). Conversely, a disruption in an underlying cognitive process may be exhibited in various fashions (Kimberg and Farah, 1993). In addition, EF tasks tap processes incidental to their main purpose and the likelihood of measurement error is therefore greater than for tasks measuring other cognitive functions (Pennington et al., 1996). As Della Sala and colleagues (1998) indicated, EF are necessary but not sufficient for the adequate performance of EF tasks. Despite the existing confusion and lack of agreement in the literature about these terms, clinicians have embraced this nomenclature with enthusiasm, probably because of what appears to be their obvious face validity.

Few studies have investigated the empirical evidence and validity of the constructs underlying EF tasks (Boone et al., 1998; Della Sala et al., 1998; Kafer and Hunter, 1997). Kafer and Hunter (1997) examined four tests purported to measure planning and problem solving in a sample of 130 neurologically intact adults aged 17 to 55 years. Structural equation modelling failed to validate any of three hypothesised models and was further unable to extract an adequate model that would account for the data. The authors concluded that their findings questioned the existence of the "planning" and "problem solving" constructs underlying the tasks used. They postulated that the tests used were possibly not sensitive measures of these constructs in the general population because of skewed score distributions. In another study, Della Sala et al. (1998) assessed 48 adult frontal lobe patients on five "traditional" frontal tests and five non-frontal tests. Their results revealed greater positive correlations within the frontal tasks than between the frontal and non-frontal tasks. Factor analyses conducted on the five frontal tests, and on the complete set only yielded single factor solutions indicating that the two groups of tasks could not be differentiated on statistical grounds. The authors interpreted their results to indicate a necessary frontal involvement in all tests and that the difference was one of degree and not of kind.

To this date, factor analytic studies of EF tasks in elderly people are lacking. In this age group, in addition to brain structural changes that may (West, 1996) or may not (Greenwood, 2000) affect the frontal lobes more predominantly, affecting EF performance, other factors need to be taken into account. First, both inter- and intra-individual cognitive variability increase with age (Christensen et al., 1999), possibly reflecting such structural changes and expressed via the use of different approaches in dealing with cognitive demands. Second, the presence of diffuse neurodegenerative disorders associated with age, such as dementia further compounds this cognitive heterogeneity. These disorders are accompanied by cognitive changes in systems not necessarily directly implicated in the resolution of EF type problems but relied upon for their appropriate execution (Collette et al., 1999; Goldberg and Bilder, Jr., 1987; Lafleche and Albert, 1995).

Thus, the evidence presented above indicates that EF measurement varies, becoming increasingly complex with advancing age. In clinics involving elderly patients, who not uncommonly present with various simultaneously occurring pathological processes, such changes need be considered and their possible impact on EF task performance taken into account. Further, research of this complex and ever evolving construct and its possible fractionation is necessary so that the cognitive abilities supporting somewhat intangible constructs can be better identified. In an attempt to address these issues, the aim of this study was to identify the structure underlying the pattern of performance on a range of EF tasks in a group of community dwellers aged 81 years and over.

METHODS

Participants

Participants were recruited from the 6-year review of the Sydney Older Persons Study (SOPS) (Creasey et al., 2001; Waite et al., 2001). SOPS is a longitudinal study following a randomly selected representative sample of 630 community dwellers from the inner west of Sydney, Australia, aged 75

TABLE I
Demographic Characteristics for the Study Sample (N = 122)

Variable	N		%
Gender (M/F)	64/58		52/48
	Mean	SD	Range
Age	85.5	3.1	81-97
Education (Years)	10.4	2.1	6-19
Estimated full scale IQ	108.7	10.0	83-128
MMSE total score	26.7	2.6	18-30
Boston naming test; total score	46.7	7.4	28-60
WMS-R Logical memory; % forgetting	33.8	26.5	– 15-100
Judgment of line orientation	16.8	2.5	8-20

Note. MMSE = Mini Mental State Examination; WMS-R = Wechsler Memory Scale-Revised.

years and over at the time of recruitment in 1991-1993. All the SOPS participants capable of giving informed consent independently, or with the assistance of a relative if necessary, were included in the present study. Participants were excluded when they were deemed too cognitively or physically impaired, to prevent potential distress from high level cognitive testing or MRI procedures included in the study protocol (not presented here). In order to maintain the broadest possible panorama of cognition in this age segment, no other *a priori* medical exclusion criteria were applied. This study received ethics approval from the Central Sydney Area Health Service.

Of the 299 remaining SOPS participants potentially available, 167 agreed to participate in this study (56%). One participant died before testing could be completed and one participant with a Mini Mental State Examination (MMSE; Folstein et al., 1975) score of 13 points was excluded; 14 participants changed their mind and an additional 7 subsequently declined because of poor health. By the time of the study completion, another 6 participants were excluded because of continuing poor health. Three participants lived too far from the testing centre and could not be seen. Finally, because of the MRI strong magnetic fields, 8 participants with a cardiac pacemaker had to be excluded and another 5 declined to participate as they were claustrophobic. Thus, 122 individuals (64 men and 58 women) took part in the current study. The sample's demographic characteristics are presented in Table I as well as descriptive statistics for some measures of the integrity of the memory (Wechsler Memory Scale-Revised Logical Memory subtest; Wechsler, 1987), visuo-perceptual (subset of Judgment of Line Orientation; Benton et al., 1983) and language systems (Boston Naming Test; Kaplan et al., 1983). These findings indicate that the sample's cognitive test performance (in terms of posterior brain functions) is comparable to other samples of similar age and education (Ivnik et al., 1996; Neils et al., 1995). Examination of the range of MMSE scores reveals 3 participants with a score of below 21 points (18, 19 and 19 points). As the emphasis of the study was on a community sample, these participants were not excluded in order to maintain a more accurate reflection of the cognitive functioning of this age group.

Procedure and Materials

Testing occurred in two sessions one month apart on average, the first one at the participant's home and the second at the local university teaching hospital. Tests of EF were administered as part of a larger battery of cognitive tests, the order remaining constant across individuals. The aim was to select tasks that were commonly used in clinical practice and would tap into different theoretical constructs of EF. Further, with our elderly sample in mind, preference was given for EF tasks that were of short duration and that would minimise the possible impact of sensory deficits common in this age group. For these reasons, popular tests such as the Wisconsin Card Sorting Test (Heaton et al., 1993) or the Stroop test (Golden, 1978) were omitted, being deemed too complex, too long or susceptible to peripheral sensory deficits. The following tests were administered: (1) WAIS-R Similarities subtest (Wechsler, 1981); (2) New tower of London (New TOL) (Hanes, 1997); (3) Semantic fluency (animal category) (Spreen and Strauss, 1991); (4) Phonemic fluency (FAS) (Benton and Hamsher, 1983); (5) Ruff figural fluency task (RFFT) (Ruff, 1988); (6) Oral trail making test-part B (Oral Trails) (Kaye et al., 1990); and (7) the WAIS-R-Digit span backward subtest (DSB) (Wechsler, 1981).

Portions of two computerised tests measuring working memory integrity were also included in the battery: the N-Back Colour 1- and 2-Back (Cohen et al., 1993) and the California Computerized Assessment Package (CalCAP) abbreviated version (Miller, 1996). For the N-Back, the aim is to respond as quickly as possible whenever a coloured circle briefly presented on a computer screen is identical to the stimulus presented immediately before (in the 1-Back condition) or to the one presented before the last one (in the 2-Back condition). Stimuli were

TABLE II

Description of the Measures of Executive Functions and Four A Priori Groupings

Task	Measure	Set	Switch	Mode
1) WAIS-R Similarities	Total words	Y	N	non-C
2) New tower of London	Highest level achieved	Y	Y	non-C
	Planning time item 3	Y	N	non-C
	Execution time item 3	N	Y	non-C
3) Semantic fluency	Switching score	N	Y	non-C
	Clustering score	Y	N	non-C
4) Phonemic fluency	Switching score	N	Y	non-C
	Clustering score	Y	N	non-C
5) Ruff figural fluency test	Total designs	Y	Y	non-C
	Total repetitions	N	Y	non-C
6) Oral trails-part B	Total time	Y	Y	non-C
	Efficiency score	N	Y	non-C
7) WAIS-R Digit span backward	Longest span	Y	Y	non-C
8) N-Back				
Colour 1-Back	Reaction time	Y	N	C
	Efficiency	N	Y	C
Colour 2-Back	Reaction time	Y	N	C
	Efficiency	N	Y	C
9) CalCAP				
Simple reaction time	Reaction time	Y	N	C
Complex reaction time	Reaction time	Y	N	C
	Efficiency	N	Y	C
Serial pattern matching 1	Reaction time	Y	N	C
	Efficiency	N	Y	C
Serial pattern matching 2	Reaction time	Y	N	C
	Efficiency	N	Y	C

Note. Y = belongs to model; N = does not belong to model (overlap is possible between set and switch models); C = computerised administration; non-C = non-computerised administration. Tasks: CalCAP = California computerized assessment package; WAIS-R = Wechsler Adult Intelligence Scale-Revised.

presented in a pseudo random order on the centre of a 12-inch colour screen. There were 38 different items comprising 37 possible pairs including 8 target pairs (9 pairs for the Colour 2-Back task). Stimulus presentation was 1000 msec with an inter stimulus interval (ISI) of 500 msec. Sitting at a distance of approximately 60-70 cm from the screen, participants were instructed to respond to the targets as quickly as possible by pressing with their dominant hand on a keypad connected to the external port. For each task, the reaction time was recorded as well as an efficiency rate which was a composite measure of accuracy and false positives.

The abbreviated CalCAP comprised four tasks varying in complexity. First was a simple reaction time (RT) task which required participants to respond to any stimulus as quickly as possible. Fifteen stimuli were presented with a random ISI ranging from 1000 to 5000 msec. The second task, a choice RT task, required a response only to the digit “7”. This task was composed of 100 trials with 15 target stimuli presented in a pseudo random order. Stimulus duration was 100 msec with an ISI of 800 msec. The third task was the “Serial Pattern Matching 1”, a task requiring a response to the presentation of two identical digits in sequence (i.e. one immediately after the other). The procedure was identical to the previous task except that it contained 20 target stimuli. The final task “Serial Pattern Matching 2” required participants to respond when they saw two digits in sequence and in increasing order (e.g. “2” followed by “3”). Number of trials, number of target stimulus were identical to the previous task. The CalCAP stimuli (single digits) were presented on the centre of a 12-inch colour screen. The instructions were very similar for the four tasks. Participants were asked to respond as quickly as possible to certain types of configurations by pressing the space bar on the keyboard and were asked to ignore all the other (irrelevant) stimuli. Each task was preceded by a maximum of 4 practice trials. The mean RT for the correct responses was recorded for all the tasks. An efficiency measure accounting for the number of false positive responses in the total number of responses was also recorded except for the single RT task where this measure was not applicable. The tests and measures collected are presented in Table II.

Planned Statistical Analyses

In order to examine the pattern of relations within the EF variables, four *a priori* models using different groupings of measures were tested with structural equation procedures. These groupings were: (1) “set” which included measures assumed to tap into cognitive abilities such as strategies, planning, problem solving and reasoning skills; (2) “switch” which included measures assumed to tap into aspects of EF such as flexibility, self-monitoring and set shifting; (3) “mode” which was taking into account test administration (computerised or not); and (4) “task” which examined grouping of measures by tasks. Allocation of measures to “set” and “switch” was based on task requirements, face validity and clinical judgment. For each model, maximum likelihood parameter estimates were obtained using AMOS 3.6.1

TABLE III
Descriptive Statistics for the EF Measures

Task	Measure	Mean	SD	Range	# of MVs
WAIS-R Similarities:	Total score	12.0	7.0	0-27	0
New tower of London:	Highest level achieved	4.2	1.6	1-9	7
	Planning time item 3	8.2	7.0	1-35	7
	Execution time item 3	22.5	24.0	3-174	7
Semantic fluency:	Switching score	7.7	3.1	1-18	1
	Clustering score	0.71	0.56	0-3.67	1
	Total score	13.2	4.3	3-31	1
Phonemic fluency:	Switching score	25.8	11.0	0-52	1
	Clustering score	0.12	0.13	0-0.55	1
	Total score	30.0	12.5	2-62	1
Ruff figural fluency test:	Total designs	37.6	13.5	12-77	11
	Total repetitions	8.1	9.0	0-66	11
Oral trails-part B:	Total time	56.7	26.6	15-108	4
	Efficiency score	87.0	20.0	0-100	4
WAIS-R Digit span backward:	Longest span	4.1	0.9	2-7	26
Colour 1-Back:	Reaction time	0.76	0.23	0.40-1.49	18
	Efficiency	63.40	37.70	– 87-100	18
Colour 2-Back:	Reaction time	1.05	0.36	0.43-3.03	18
	Efficiency	11.00	39.82	– 187-87	18
CalCAP simple RT:	Reaction time	525.49	264.54	289-2447	10
CalCAP complex RT:	Reaction time	554.40	81.37	413-831	12
	Efficiency	0.86	0.15	0.20-1	12
CalCAP serial pattern matching 1:	Reaction time	567.69	79.06	388-855	17
	Efficiency	0.56	0.23	– 0.05-1	17
CalCAP serial pattern matching 2:	Reaction time	584.99	95.01	236-866	27
	Efficiency	0.42	0.18	0.04-0.80	27

Note. CalCAP = California computerized assessment package; WAIS-R = Wechsler Adult Intelligence Scale-Revised; MV = missing value.

(Arbuckle, 1997). Currently, in the domain of structural equation modelling, a large number of indices are available to test the goodness of fit of a model. (See Byrne, 2001 for a good introduction text on these increasingly popular methods). In this study, goodness of fit was established using the goodness of fit index (GFI), the adjusted GFI (Arbuckle, 1997) and the root mean square error approximation (RMSEA) (Browne and Cudeck, 1992). A GFI in excess of .90 indicates a good fit of the model onto the data (Marsh et al., 1988); an adjusted GFI will always be lower than the GFI, imposing a penalty related to the number of parameters being estimated, and values in excess of .90 indicate very good fits; whilst a RMSEA value of less than 0.04 is indicative of a close fit of the model (Browne and Cudeck, 1992). The models with their respective groupings are presented in Table II. The data, correlations and SDs, used for these analyses are presented in the Appendix, with missing values (MVs) imputed. For all the EF measures, existing missing values were inferred from the scores from the remaining variables within the set using a multiple regression procedure. In other words, the obtained score, in replacement of the MV, was a composite score reflecting the pattern of associations among predictors (i.e., the other variables) with the predictor having the MV.

Results

The descriptive statistics for the EF measures are presented in Table III. Within the structural equation modelling procedures, a single factor model fitting "set" was examined first. This was followed by a second model adding "switch" as a second factor before the third model, adding "mode", was examined. Finally, a "task" model was examined independently. These models did not include the total number of words for the two fluency tasks as they were deemed to reflect a composite measure of the number and size of word clusters already represented in the data set (switching and clustering scores respectively).

None of the *a priori* models was able to provide an acceptable fit of the interrelations among the set of variables (see Table IV). The initial single factor ("set") model gave rise to a GFI of .552. The GFI improved only marginally after the introduction of the "switch" and "mode" factor (.686 and .722 respectively). The "task" model, positing 9 factors underlying the nine tests was unable to converge and gave rise to an inadmissible solution because of the presence of negative variance or correlations larger than 1.00 among some variables.

Given the unsuccessful modelling of measures according to theoretical neuropsychological constructs, the components from each test were examined individually. It appeared that certain measures exhibited partial logical constraints. For example, the switching and mean cluster size scores from the Semantic fluency task had a negative correlation (r = – .50) and were "competing" in that high scores on one component could only be achieved at the expense of the other. The same phenomenon was observed to a lesser degree for the RFFT between the number of repetitions and the number of unique designs (r = – .14).

TABLE IV
Results of the Confirmatory Factor Analyses for Four A Priori Models

Model	Admissible	χ^2	df	GFI	AGFI	RMSEA
"Set"	Yes	1093.494	262	.552	.487	0.162
"Set" & "Switch"	Yes	711.228	247	.686	.619	0.125
"Set", "Switch" & "Mode"	Yes	595.997	234	.722	.643	0.113
"Task"	No	610.154	222	.725	.628	0.120

Note. GFI = goodness of fit; AGFI = adjusted goodness of fit; RMSEA = root mean square error of approximation.

Given these constraints, a statistically driven analysis of the EF measures was conducted to clarify the underlying structure within each test with the goal of extracting one meaningful measure per task. For each task, this was achieved by first identifying and removing existing constraints. Unnecessary measures, as determined by the results of the analyses performed on each task were also removed. As such, to resolve the competition between the switching and clustering measures from the two verbal fluency tasks, these were replaced with their respective global measure (i.e., total number of words). Single indicators for the WAIS-R DSB and Similarities subtests were retained. For the seven CalCAP measures, no logical constraints, such as the ones described above, were present. Maximum likelihood factor analytic investigations within this set of measures revealed two factors with an eigenvalue greater than 1. Confirmatory factor analysis of this two-factor solution using AMOS 3.6.1 (Arbuckle, 1997) produced a GFI of .969 (χ^2 = 14.837; df = 8) but yielded a correlation of .968 between the two factors, these being in essence indistinguishable. The CalCAP measures were therefore summarised into one composite score composed of their z-score weighted total. A similar procedure was applied to the N-Back measures. As two factors could not be fitted to four indicators, these measures were summarised with their average z-score weighted total into one composite score. Following the same procedure, a composite score was constructed from the time and performance measures from the Oral Trails task as they seemed to tap a common construct conceptually. For the New TOL test, the highest level of complexity successfully completed was retained. Finally for the RFFT, a composite measure was constructed taking into account the number of errors in the total performance.

The factorial structure of this set of nine variables was investigated with AMOS 3.6.1 (Arbuckle, 1997). An initial exploratory model fitted with one general factor gave rise to a GFI of .905, an adjusted GFI of .842 and a RMSEA value of 0.094 (χ^2 = 55.89; df = 27; p = .001). The introduction of a second factor saw the exploratory model improve significantly with a GFI of .943, an adjusted GFI of .864 and a RMSEA of 0.078 (χ^2 = 33.107; df = 19; p = .023); increment χ^2 = 22.783 on 8 df; p = .004. The highest modification index between residuals was ascertained (12.447 between the two verbal fluency measures) and a correlated uniqueness parameter between these residuals (yielding r = .42) was added to the model. This improved the fit of the two-factor model substantially, giving rise to a GFI value of .970, an adjusted GFI of .925 and a RMSEA value of 0.00 (χ^2 = 16.553; df = 18; p = .554)[1]. This model was then "reified" (i.e., an oblique factor rotation chosen) by defining a 2-factor model with the first factor loading exclusively on N-Back and not on DSB and, conversely, the second factor loading exclusively on DSB but not on N-Back. Constraining the latent Factor1-2 correlation from this model (r = – .221; Critical Ratio = – 1.126) to 0 resulted in no loss of information: GFI = .968, adjusted GFI = .923 and RMSEA = 0.00 (χ^2 = 17.854; df = 19; p = .532). This indicated that the correlation between the two factors indicated by N-Back and DSB tasks was well within chance bounds (i.e., not significantly different from 0) and strongly supported a two-orthogonal-factor model as acceptable for the nine tasks. Finally, the loadings from FAS and Similarities to Factor-1 (the factor indicated by N-Back) were not significant, so a final model was run with these set to zero: GFI = .966, adjusted GFI = .926 and RMSEA = 0.00 (χ^2 = 19.260; df = 21; p = .568). The estimates from this final model are shown in Table V.

Inspection of the relative contribution of each factor to the explained variance for each EF measure allowed interpretation in terms of three distinct groups: (a) a group of measures showing loadings on Factor 2 in excess of 85% (DSB, Similarities, FAS and RFFT); (b) a group of measures with approximately equal relative contributions from Factor 1 and 2 to the explained variance (Semantic fluency, Oral Trails, CalCAP and New TOL); and finally (3) N-Back, loading exclusively on Factor 1. The intercorrelations among the nine EF composite measures are presented in Table VI with the tasks grouped according to their respective loadings. The overall average correlation was .35, however, within and between group average correlations were more varied with a positive correlation coefficient of .446 among the measures loading primarily on

[1]This over association between these two measures cannot be ignored. The introduction of the correlated uniqueness parameter can be seen as a third minor common factor for these two tasks only, whilst remaining uncorrelated with the other two factors.

TABLE V
Final Two-factor Model for the Nine EF Measures

	Factor 1	Factor 2	Variance explained (%)	Ratio (F1/F2)	Relative contribution to explained variance	
					Factor 1	Factor 2
Colour N-Back	0.499	0	24	Infinite	1.00	0.00
DSB	0	0.614	23	**0**	0.00	1.00
Similarities	0	4.519	18	**0**	0.00	1.00
FAS	0	7.892	16	**0**	0.00	1.00
RFFT	3.342	10.048	25	**0.33**	0.10	0.90
Semantic fluency	1.098	2.226	12	0.49	0.20	0.80
Oral Trails	0.384	0.488	23	0.79	0.38	0.62
CalCAP	0.275	0.359	13	0.77	0.37	0.63
New TOL	0.654	0.484	8	1.35	0.65	0.35

Note. CalCAP = California computerized assessment package; DSB = Digit span backward; FAS = Phonemic fluency; New TOL = New tower of London; RFFT = Ruff figural fluency test.

TABLE VI
Correlations (× 100) and SDs for the Nine EF Measures Grouped according to their Respective Factor Loadings

Loading	Measure	SDs	1	2	3	4	5	6	7	8	9
			Factor 1	Factor 2				Mixed			
Factor 1	1. Colour N-Back	0.71	100								
Factor 2	2. DSB	0.88	10	100							
	3. Similarities	6.94	9	43	100						
	4. FAS	12.41	12	42	43	100					
	5. RFFT	15.13	26	55	43	43	100				
Mixed	6. Semantic fluency	4.29	24	33	35	61	48	100			
	7. Oral Trails	0.92	36	40	41	41	40	42	100		
	8. CalCAP	0.77	31	29	38	36	40	35	47	100	
	9. New TOL	1.54	31	29	15	24	35	35	39	29	100

Note. All correlations larger than .15 significant at the .005 level.
CalCAP = California computerized assessment package; DSB = Digit span backward; FAS = Phonemic fluency; New TOL = New tower of London; RFFT = Ruff figural fluency test.

Factor 2 (DSB, Similarities, FAS and RFFT) and .379 among those loading equally on both factors (i.e. "Mixed"; Semantic fluency, Oral Trails, CalCAP and New TOL). Among the groups, the correlation between N-Back (i.e., Factor 1 measure) and the Factor 2 measures was much lower (average r = .141) whilst the correlations between the "Mixed" group and Factor 1 and 2 groups were .307 and .366 respectively.

Discussion

Bearing in mind the smallness of our sample and the large number of EF tests examined, in this sample composed of elderly community dwellers, none of the postulated EF models was found to yield even approximate fit of the covariance data. Somewhat surprisingly, the EF measures did not fit the groupings according to existing models of EF cognitive constructs thought to reflect the more basic (core) cognitive requirements underlying tasks of EF. Thus, the "set" grouping supposedly requiring means-goal analysis, plan of actions, strategy application, activation of past experience and appropriate activation of schemas, plans or scripts; or the "switch" grouping supposedly requiring on-line maintenance of activation, discrepancy analysis between "now and end goal", and shifting of plans of actions failed to explain the data adequately. The presence of logical constraints between measures of the same task (yielding negative correlations as "method" effects) such as switching and clustering scores in Semantic fluency made it impossible to examine the exact role of some components and may have compounded these findings. Our statistically driven approach, devoid of any *a priori* theory, was an attempt to clarify the existing underlying constructs within each test, to reduce the existing dependency among some measures and to shed light on the relations among these tasks. From these investigations, the associations among these measures can be explained by two orthogonal factors with three distinct groups emerging from these analyses. Factor 1 was defined such that one measure, N-Back, loads exclusively on it. Factor 2 was defined such that DSB loaded exclusively on it, and we found that three other measures essentially load only on this factor as well. Surprisingly, these two factors are in essence uncorrelated (this feature of the data can directly be seen from the near zero correlations between N-Back and the four Factor-2 indicator tests in Table 6). Finally, the variance on the last four measures is explained by contributions from both Factors 1 and 2. The variance explained in these measures by the two common factors was low (8%-25%), indicating that the tests probably are somewhat idiosyncratic to a major extent.

One interpretation emerging from the results of the factor analytic investigations fit Eslinger and Grattan's (1993) theoretical framework of EF. The measures that predominantly load on Factor 2 (FAS, RFFT, Similarities and DSB) are all characterised by commonly encountered stimuli in the context of unusual demands and manipulations. An optimal performance is reliant on the generation of a diversity of responses and creative approaches that are all internally driven. Such tasks require divergent thinking or, in the words of Eslinger and Grattan (1993): spontaneous flexibility. In contrast, N-Back, which is loading exclusively on Factor 1, is a task that is unusual from the outset and requires unusual responses. In this instance, self-generated strategies cannot be applied for appropriate performance and correct responses are only achieved in response to stimuli triggered by the environment. This task requires participants to shift response set in relation to external cues and relies on convergent thinking or reactive flexibility (i.e., the "readiness to freely shift cognition and behaviour according to the particular demands and context of a situation"; Eslinger and Grattan, 1993, p. 18). Whilst a verbal, internally driven, strategy may be used, the task requires a constant update of information within working memory which can only be achieved successfully by responding to the environment. The tasks that belong to the "Mixed" group (i.e., loading on both factors) necessitate concurrent activation of on-line maintenance of information (spontaneous flexibility) within a given framework of activation (reactive flexibility). For example, Semantic fluency requires the production of words within a single semantic category and is therefore subject to clear external boundaries. Spontaneous flexibility remains however a necessary component reflected in the choice of strategies (e.g., clusters and switches among subcategories of animals). In this group, the difference in factor loadings would reflect the strength of the existing framework (e.g., less for Semantic fluency and greater for New TOL).

Alternative explanations to the data set were unsatisfactory. These findings were not explained by task requirements, our results showing no specific groupings for the measures by modality (such as language versus non-language based processes, motor or visuospatial abilities). Similarly, decline in speed of processing, which has been put forward as the primary underlying process accounting for changes in EF performance with ageing (Salthouse, 1994; Salthouse and Babcock, 1991), did not contribute to the various groupings of measures. Our findings are in contrast with other studies. Kensinger et al. (2000) and Shimamura (2000) describe DSB and N-Back as measuring the same cognitive construct, namely updating and manipulation of information. Our results suggest that, in a very old sample, these two tasks require quite different processes for appropriate performance. Within a working memory and age difference framework, it is possible that N-Back and DSB load on a common factor but not when other aspects of EF are considered.

Several methodological aspects of this study deserve comment. First, the total variance explained by the two factors remains small on all EF measures (between 8 and 25%). This indicates the presence of uncorrelated specific factors, as well as noise, associated with the task and suggests that this group of tasks is not strongly related. Nevertheless, within the shared portion of the variance, the processes can be explained along the lines of two underlying factors reflecting internal/external or divergent/convergent thinking processes. Second, it may be argued that these findings may not be applicable to other tests of EF such as the WCST or the Stroop but this needs to be determined. Nevertheless, the tasks were selected on the basis of sound criteria: to be commonly used in clinical practice; to be relevant in the context of the assessment of very old participants; and to map onto previously described underlying cognitive constructs within the EF domain. Thirdly, the contribution of certain components, such as switching and clustering, could not be examined fully. Although both may be associated with EF, because of task requirements, these cannot be measured simultaneously. When they arise from the same task, there is a risk of obtaining findings that are influenced not by underlying EF constructs but by task constraints thus distorting the model. Had the "switching" and "clustering" scores been collected on separate occasions, it is predicted that the negative correlation between the two measures would have been attenuated or would have even become positive. Finally, the discrepancy between these findings and other studies may be related to the nature of our sample. This community sample comprises randomly selected "old-old" participants, some of whom may present with some level of cognitive impairment. Whilst not homogeneous in that respect (i.e., not composed of completely healthy subjects only), this sample more accurately reflects the cognitive functioning of this age group, with its increasingly common within and between-subject variability in test performance (Christensen et al., 1999). It is also more similar to the type of patients clinicians are likely to encounter in their practice.

Beyond the uncertainty regarding the face validity of some of the existing tasks of EF (Burgess, 1997; Levine et al., 1998; Phillips, 1997; Rabbitt, 1997), our findings underscore the difficulty in investigating the measurement of EF, particularly in the oldest segment of the population. Test performance data on this age group remain scarce. Perhaps the specific cognitive requirements for the successful completion of tasks of EF vary depending on the cognitive development of the participants. Denckla (1996) has suggested that some EF tasks tap into different cognitive constructs in children and in adults. It may be fallacious to entertain the thought of a unitary EF task, applicable to all age groups, given that a novel task at a certain age may not be so at a later stage of development due to continuing exposure to similar situations. Consequently, it is possible that how very old participants approach tasks of EF differ from younger adult participants. At this point in time, as suggested by Kafer and Hunter (1997), available instruments may not allow clinicians to measure the many hypothesised facets of EF. Alternatively, it is also possible that measures of EF are age specific and that different tasks may be required at different ages.

The size and representativeness of the sample is certainly a strength for a study involving participants aged 81 years and over although the number of participants may arguably be small by statistical standards. These results are presented as a conceptual analysis in an attempt to raise awareness and discuss an important methodological issue in the assessment of EF rather than to provide a definitive answer to a very complex question. Nevertheless, the results of these analyses conducted on the EF measures gave rise to a model of EF which is intuitive and fits previous findings (Eslinger and Grattan, 1993). Obviously, this model will need to be examined and tested further. In any case, continuing research on EF is needed not only to better understand the subtleties of human cognition but also to develop appropriate instruments that will improve the measurement of specific cognitive constructs.

Acknowledgement. This research was supported in part by a grant from the National Health and Medical Research Council of Australia, the Ageing and Alzheimer's Research Foundation and an Infrastructure Stream C grant from the Department of Health of New South Wales, Australia.

References

Albert MS, Wolfe J and Lafleche G. Differences in abstraction ability with age. *Psychology and Aging, 5:* 94-100, 1990.

Arbuckle JL. *Amos User's Guide Version 3.6*. Chicago: Smallwaters Corporation, 1997.

Benton AL and Hamsher K. *Multilingual Aphasia Examination*. Iowa City: AJA, 1983.

Benton AL, Hamsher KD, Varney NR and Spreen O. *Contributions to Neuropsychological Assessment*. New York: Oxford University Press, 1983.

Boone KB, Pontón MO, Gorsuch RL, González JJ and Miller BL. Factor analysis of four measures of prefrontal lobe functioning. *Archives of Clinical Neuropsychology, 13:* 585-595, 1998.

Browne MW and Cudeck R. Alternative ways of assessing model fit. *Sociological Methods and Research, 21:* 230-258, 1992.

Burgess PW. Theory and methodology in executive function research. In P Rabbitt (Ed), *Methodology of Frontal and Executive Function.* Hove, Sussex: Psychology Press, 1997, pp. 81-116.

Byrne BM. *Structural Equation Modeling. Basic Concepts, Applications and Programming.* London: Lawrence Erlbaum, 2001.

Christensen H, MacKinnon AJ, Korten AE, Jorm AF, Henderson AS and Jacomb P. Dispersion in cognitive ability as a function of age: A longitudinal study of an elderly community sample. *Aging, Neuropsychology and Cognition, 6:* 214-228, 1999.

Cohen JD, Forman SD, Braver TS, Casey BJ, Servan-Schreiber D and Noll DC. Activation of the prefrontal cortex in a nonspatial working memory task with functional MRI. *Human Brain Mapping, 1:* 293-304, 1993.

Collette F, Van Der Linden M and Salmon E. Executive dysfunction in Alzheimer's disease. *Cortex, 35:* 57-72, 1999.

Crawford JR, Bryan J, Luszcz MA, Obonsawin MC and Stewart L. The executive decline hypothesis of cognitive aging: Do executive deficits qualify as differential deficits and do they mediate age-related memory decline? *Aging, Neuropsychology and Cognition, 7:* 9-31, 2000.

Creasey H, Waite LM, Grayson DA, Bennett HP, Dent O and Broe GA. The impact of the neurodegenerative disorders on ageing: An overview of the Sydney Older Persons Study. *Australasian Journal on Ageing, 20:* 11-16, 2001.

Daigneault S and Braun CMJ. Working memory and the Self-Ordered Pointing Task: Further evidence of early prefrontal decline in normal aging. *Journal of Clinical and Experimental Neuropsychology, 15:* 881-895, 1993.

Daigneault S, Braun CMJ and Whitaker HA. Early effects of normal aging on perseverative and non-perseverative prefrontal measures. *Developmental Neuropsychology, 8:* 99-114, 1992.

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

Denckla MB. A theory and model of executive function. A neuropsychological perspective. In GR Lyon and NA Krasnegor (Eds), *Attention, Memory, and Executive Function.* Baltimore: Paul H. Brooks Publishing, 1996, pp. 263-278.

Eslinger PJ and Grattan LM. Frontal lobe and frontal-striatal substrates for different forms of human cognitive flexibility. *Neuropsychologia, 31:* 17-28, 1993.

Folstein MF, Folstein SE and McHugh PR. "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12:* 189-198, 1975.

Goldberg E and Bilder RM, Jr. The frontal lobes and hierarchical organization of cognitive control. In E Perecman (Ed), *The Frontal Lobes Revisited*. New York: Lawrence Erlbaum, 1987, pp. 159-187.

Golden CJ. *Stroop Color and Word Test*. Chicago: Stoelting, 1978.

Grattan LM and Eslinger PJ. Frontal lobe damage in children and adults: A comparative review. *Developmental Neuropsychology, 7:* 283-326, 1991.

Greenwood PM. The frontal aging hypothesis evaluated. *Journal of the International Neuropsychological Society, 6:* 705-726, 2000.

Hanes KR. *New Tower of London*. Carlton, Australia: Vanguard, 1997.

Heaton RK, Chelune GJ, Talley JL, Kay GG and Curtis G. *Wisconsin Card Sorting Test (WCST) Manual Revised and Expanded*. Odessa,Fl: Psychological Assessment Resources, 1993.

Ivnik RJ, Malec JF, Smith GE and Tangalos EG. Neuropsychological tests' norms above age 55: COWAT, BNT, MAE Token, WRAT-R Reading, AMNART, STROOP, TMT, and JLO. *The Clinical Neuropsychologist, 10:* 262-278, 1996.

Kafer KL and Hunter M. On testing the face validity of planning/problem-solving tasks in a normal population. *Journal of the International Neuropsychological Society, 3:* 108-119, 1997.

Kaplan EF, Goodglass J and Weintraib S. *The Boston Naming Test*. Philadelphia: Lea and Febiger, 1983.

Kaye K, Grigsby J, Robbins LJ and Korzun B. Prediction of independent functioning and behaviour problems in geriatric patients. *Journal of the American Geriatrics Society, 38:* 1304-1310, 1990 .

Kensinger EA, Milligan DK, Locascio JJ and Corkin S. The effect of Alzheimer's disease and Parkinson's disease on the components of working memory. In JH Growdon, RJ Wurtman, S Corkin, and RM Nitsch (Eds), *Molecular Genetics of Dementia*. Cambridge, MA: Center for Brain Sciences and Metabolism Charitable Trust, 2000, pp. 291-297.

Kimberg DY and Farah MJ. A unified account of cognitive impairment following frontal lobe damage: The role of working memory in complex, organized behavior. *Journal of Experimental Psychology: General, 122:* 411-428, 1993.

LAFLECHE G and ALBERT MS. Executive function deficits in mild Alzheimer's Disease. *Neuropsychology, 9:* 313-320, 1995.

LEVINE B, STUSS DT and MILBERG WP. Concept generation: Validation of a test of executive functioning in a normal aging population. *Journal of Clinical and Experimental Neuropsychology, 17:* 740-758, 1995.

LEVINE B, STUSS DT, MILBERG WP, ALEXANDER MP, SCHWARTZ M and MACDONALD R. The effects of focal and diffuse brain damage on strategy application: Evidence from focal lesions, traumatic brain injury and normal aging. *Journal of the International Neuropsychological Society, 4:* 247-264, 1998.

LEZAK M. *Neuropsychological Assessment*. New York: Oxford University Press, 1995.

LOWE C and RABBITT P. Cognitive models of ageing and frontal lobe deficits. In Rabbitt P (Ed), *Methodology of Frontal and Executive Function.* Hove, Sussex: Psychology Press, 1997, pp. 39-59.

LURIA A. *The Working Brain*. New York: Basic Books, 1973.

MARSH HW, BALLA JR and MCDONALD RP. Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin, 103:* 391-410, 1988.

MILLER EN. *CalCAP California Computerized Assessment Package Manual*. Los Angeles: Norland Software, 1996.

MITTENBERG W, SEIDENBERG M, O'LEARY DS and DIGIULIO DV. Changes in cerebral functioning associated with normal aging. *Journal of Clinical and Experimental Neuropsychology, 11:* 918-932, 1989.

NEILS J, BARIS JM, CARTER C, DELL'AIRA AL, NORDLOH SJ, WEILER E and WEISINGER B. Effects of age, education, and living environment on Boston Naming Test performance. *Journal of Speech and Hearing Research, 38:* 1143-1149, 1995.

PARKIN AJ and LAWRENCE A. A dissociation in the relation between memory tasks and frontal lobe tests in the normal elderly. *Neuropsychologia, 32:* 1523-1532, 1994.

PENNINGTON BF, BENNETTO L, MCALEER O and ROBERTS RJJ. Executive functions and working memory: Theoretical and measurement issues. In LG Reid and NA Krasnegor (Eds), *Attention, Memory and Executive Function.* Baltimore: Paul H. Brookes Publishing Co., 1996, pp. 327-348.

PHILLIPS LH. Do "Frontal Tests" measure executive function? Issues of assessment and evidence from fluency tests. In P Rabbitt (Ed), *Methodology of Frontal and Executive Function.* Hove, Sussex: Psychology Press, 1997, pp. 191-213.

RABBITT P. Introduction: Methodologies and models in the study of executive function. In P Rabbitt (Ed), *Methodology of Frontal and Executive Function.* Hove, Sussex: Psychology Press, 1997, pp. 1-38.

RABBITT P and LOWE C. Patterns of cognitive ageing. *Psychological Research, 63:* 308-316, 2000.

RAZ N, GUNNING FM, HEAD D, DUPUIS JH, MCQUAIN J, BRIGGS SD, LOKEN WJ, THORNTON AE and ACKER JD. Selective aging of the human cerebral cortex observed in vivo: Differential vulnerability of the prefrontal gray matter. *Cerebral Cortex, 7:* 268-282, 1997.

ROBBINS TW, JAMES M, OWEN AM, SAHAKIAN BJ, LAWRENCE AD, MCINNES L and RABBITT PMA. A study of performance on tests from the CANTAB battery sensitive to frontal lobe dysfunction in a large sample of normal volunteers: Implications for theories of executive functioning and cognitive aging. *Journal of the International Neuropsychological Society, 4:* 474-490, 1998.

RUFF RM. *Ruff Figural Fluency Test Administration Manual*. San Diego: Neuropsychological Resources, 1988.

SALTHOUSE TA. The aging of working memory. *Neuropsychology, 8:* 535-543, 1994.

SALTHOUSE TA and BABCOCK RL. Decomposing adult age differences in working memory. *Developmental Psychology, 27:* 763-776, 1991.

SHALLICE T. *From neuropsychology to mental structure*. New York: Cambridge University.Press, 1988.

SHIMAMURA AP. Memory and frontal lobe function. In MS Gazzaniga (Ed), *The Cognitive Neurosciences*. Cambridge, MA: MIT Press, 1995, pp. 803-813.

SHIMAMURA AP. The role of the prefrontal cortex in dynamic filtering. *Psychobiology, 28:* 207-218, 2000.

SHIMAMURA AP and JURICA PJ. Memory interference effects and aging: Findings from a test of frontal lobe function. *Neuropsychology, 8:* 408-412, 1994.

SPREEN O and STRAUSS E. *A Compendium of Neuropsychological Tests*. New York: Oxford University Press, 1991.

STUSS DT and ALEXANDER MP. Executive functions and the frontal lobes: a conceptual view. *Psychological Research, 63:* 289-298, 2000.

STUSS DT, CRAIK FIM, SAYER L and FRANCHI D. Comparison of older people and patients with frontal lesions: Evidence from word list learning. *Psychology and Aging, 11:* 387-395, 1996.

TRANEL D ANDERSON SW and BENTON A. Development of the concept of "executive function" and its relationship to the frontal lobes. In F Boller and J Grafman (Eds), *Handbook of Neuropsychology (vol 9).* Amsterdam: Elsevier, 1994, pp. 125-148.

WAITE LM, BROE GA, GRAYSON DA and CREASEY H. The incidence of dementia in an Australian community population: the Sydney Older Persons Study. *International Journal of Geriatric Psychiatry, 16:* 680-689, 2001.

WECHSLER D. *WAIS-R Manual*. New York: The Psychological Corporation, 1981.

WECHSLER D. *WMS-R Manual*. New York: The Psychological Corporation, 1987.

WECKER NS, KRAMER JH, WISNIEWSKI A, DELIS DC and KAPLAN E. Age effects on executive ability. *Neuropsychology, 14:* 409-414, 2000.

WEST RL. An application of prefrontal cortex function theory to cognitive aging. *Psychological Bulletin, 120:* 272-292, 1996.

Olivier Piguet, Prince of Wales Medical Research Institute, Barker St, Randwick NSW 2031, Australia. e-mail: o.piguet@unsw.edu.au

(*Received 14 June 2002; reviewed 30 September 2002; revised 2 December 2002; accepted 4 March 2003; Action Editor Carlo Umiltà*)

APPENDIX

Correlations (× 100) and SDs for the Initial 24 Component Scores (see Table II for test-domain labels)

Measure	SD	1	2	2	2	3	3	4	4	5	5	6	6	7	8	8	8	8	9	9	9	9	9	9	9
1) Total words	6.97	100																							
2) Highest level achieved	1.55	15	100																						
2) Planning time item 3	7.01	13	6	100																					
2) Execution time item 3	23.98	22	44	12	100																				
3) Switching score	3.12	26	22	3	11	100																			
3) Clustering score	0.56	6	11	−3	8	−50	100																		
4) Switching score	10.97	38	20	−4	13	39	7	100																	
4) Clustering score	0.13	21	20	5	8	12	19	9	100																
5) Total designs	13.53	40	29	11	19	24	11	34	1	100															
5) Total repetitions	9.00	12	16	−6	23	4	2	20	−1	−14	100														
6) Total time	26.61	44	29	22	27	22	11	38	15	53	−1	100													
6) Efficiency score	20.04	32	43	13	32	20	9	26	17	33	−5	69	100												
7) Longest span	0.89	43	29	20	27	28	−5	43	−3	46	23	46	28	100											
8) 1-Back Reaction time	0.23	−2	2	8	13	12	7	6	−12	5	1	12	30	−14	100										
8) 1-Back Efficiency	37.70	0	28	29	37	17	−3	5	−7	24	−5	20	30	12	48	100									
8) 2-Back Reaction time	0.36	7	27	15	27	6	10	21	0	14	−1	20	21	3	38	38	100								
8) 2-Back Efficiency	39.82	22	32	23	24	11	−1	5	−3	48	−7	34	24	28	8	50	24	100							
9) Simple Reaction time	264.54	25	10	−3	15	11	2	21	18	34	−9	33	15	20	16	13	10	19	100						
9) Complex Reaction time	81.37	26	19	−3	20	21	1	22	5	34	−10	31	24	1	29	29	35	14	36	100					
9) Complex Efficiency	0.15	39	33	11	41	20	5	21	8	33	−2	35	22	18	15	26	34	18	46	71	100				
9) Serial 1 Reaction time	79.06	−1	0	−1	−12	17	2	12	−5	13	−7	11	−2	−2	14	4	6	−4	18	39	14	100			
9) Serial 1 Efficiency	0.23	32	23	18	21	16	12	29	19	49	−12	57	40	26	16	29	24	22	39	63	53	38	100		
9) Serial 2 Reaction time	95.01	11	−2	−3	−10	15	−3	15	13	18	−4	17	4	−2	12	−4	10	−17	9	35	18	63	37	100	
9) Serial 2 Efficiency	0.18	34	34	13	9	28	8	29	12	47	−8	48	36	25	17	17	20	25	17	48	39	33	51	40	100