Heliyon 9 (2023) e15504

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

20 years on: Confirmation of P. Anderson's (2002) paediatric model of executive functioning in a healthy adult sample[☆]

Jessica Scarfo [*], Emra Suleyman, Michelle Ball

Psychology, Victoria University; Ballarat Rd, Melbourne, Vic, 3011, Australia

ARTICLE INFO

Keywords:
Executive-function
Assessment
Cognition
Normative data
Neuropsychology
Psychometrics

ABSTRACT

Executive Functioning (EF) is a construct that encompasses multiple interrelated higher order skills, however, conceptualising this nebulous construct remains challenging. This study aimed to confirm the validity of Anderson's (2002) paediatric model of EF in a healthy adult sample using congeneric modelling. Measures of EF were selected based on utility with adult populations giving rise to minor methodological differences from the original paper. Separate congeneric models were constructed using each of Anderson's constructs in order to isolate the sub-skills represented by each (Attentional Control-AC, Cognitive Flexibility-CF, Information Processing-IP, Goal Setting-GS), with a minimum of three tests per subskill. One hundred and thirty-three adults (42 males and 91 females) aged between 18 and 50 (M = 29.68, SD = 7.46) completed a cognitive test battery comprising 20 EF tests. AC revealed a good fitting model $\chi^2(2) = 1.61$, $p = .447$, RMSEA = 0.000, CFI = 1.000, after removing the non-significant indicator Map Search ($p = .349$), and BS-Bk as BS-Bk was required to covary with both BS-Fwd (M.I = 7.160, Par Change = .706), and TMT-A (M.I = 5.759, Par Change = −2.417). CF revealed a good fitting model $\chi^2(8) = 2.90$, $p = .940$, RMSEA = 0.000, CFI = 1.000 after covarying TSC-E and Stroop (M.I = 9.696, Par Change = .085). IP revealed a good fitting model $\chi^2(4) = 1.15$, $p = .886$, RMSEA = 0.000, CFI = 1.000 after covarying Animals total and FAS total (M.I. = 4.619, Par Change = 9.068). Lastly, GS indicated a good fitting model $\chi^2(8) = 7.22$, $p = .513$, RMSEA = 0.000, CFI = 1.000 after covarying TOH total time and PA (M.I = 4.25, Par Change = −77.868). Therefore, all four constructs were reliable and valid, and the utility of a parsimonious EF battery is suggested. Investigation of the inter-relationships between the constructs using regression techniques, de-emphasises the role of Attentional Control and argue instead for capacity bound skills.

1. Introduction

Of all domains of cognitive function, Executive Functioning (EF) has arguably been the most vexing to researchers and clinicians alike. Whilst converging lines of scientific enquiry often narrow to a solution or bring clarity in research, it could be argued that EF

[☆] Author note: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. This manuscript is original, has not been published and is not under consideration for publication elsewhere. The datasets generated and analysed during the current study are not publicly available due to ethical constraints but are available from the corresponding author on reasonable request and permission from the appropriate human research ethics committee. Coding is not applicable. This study was not preregistered. Declarations of interest: none.

[*] Corresponding author.

E-mail addresses: jessica.burlak@vu.edu.au (J. Scarfo), emra.suleyman@vu.edu.au (Emra Suleyman), michelle.ball@vu.edu.au (M. Ball).

https://doi.org/10.1016/j.heliyon.2023.e15504
Received 6 August 2022; Received in revised form 27 March 2023; Accepted 11 April 2023

represents a varied and complex skill-set, and that it is this complexity that hinders its clear conceptualisation. Fundamental to the assessment of any cognitive skill is a sound theoretical underpinning of the classification and mechanisms associated with its function. Cognitive psychology literature has achieved clarity with respect to the definitions and underpinnings within intelligence, memory and learning, and language functioning [1–3] and these sound theoretical models underpin assessment, diagnosis and treatment processes. However, the cognitive construct of Executive Functioning still lacks many of the features necessary for a sound theoretical model. Executive Functioning is plagued by a lack of cohesion with respect to its definition, the skills that comprise it, and the operation and inter-relationship between these sub-skills as demarcated by models.

Numerous models of EF exist, including Behavioural Inhibition [4]; Four component conceptualisation [5,6]; Supervisory Attentional System (SAS [7–9]); and the Problem-solving framework [10]. However a model that more specifically emphasises the broad skills that encompass EF, and the consideration of the role of attention, is the paediatric model put forward by Anderson [11]. Whilst this theoretically driven model has been utilised as a conceptual framework in both healthy and clinical paediatric populations [12–14], it has never been empirically validated in either children or adults (Fig. 1). This paediatric model addresses most of the limitations circulating the literature regarding a) how EF skills contribute to the overall achievement of "independent purposive self-serving behaviour" [5] b) proposes an inherent hierarchy of cognitive processes in the overall performance of EF, and c) the degree to which the directionality of these skills interrelate, thereby considering the theoretical nature of EF in a comprehensive way. From the perspective outlined above, the strength of Anderson's model is that it adopts numerous tests from numerous paradigms to ensure a holistic assessment is achieved. On the face of it, this model presents an elegant solution that considers the integration and inter-operationalisation of an array of executive and non-executive skills within an overarching model of EF.

Attention is a construct consistently omitted from adult models of EF, despite this being included in most definitions. Evidence that attention must be considered when discussing EF comes from the fact that attentional difficulties manifest in many childhood disorders with an inherent executive dysfunction component (e.g., ADHD, autism, PDD), and shares functional attribution when an EF is engaged [15]. Although the explicit inclusion of an Attentional Control domain in Anderson's model is likely owing to the consideration of emerging skills in childhood, the consideration of attention in EF is largely absent in existing models. Compared to other more reductionist models of EF, the strength of Anderson's model is the consideration of selective attention, self-regulation, self-monitoring, and inhibition. These inclusions have high face validity for the definition of EF put forward by seminal neuropsychology author Muriel Lezak [5] who proposed that one needs self-regulation and self-monitoring to regulate EF performance. A particular strength of Anderson's model is that it also considers the influence of attention across different EF domains, and although not explicitly stated, the hypothesised model is drawn with unidirectional arrows pointing outwards from the Attentional Control sub-domain, suggesting that this might be a fundamental mechanism to EF.

A review of literature suggests that many tasks may assess cognitive flexibility, and whilst some studies have found support for its inclusion as a factor in isolation, explanations regarding which skills may contribute to effective performance on cognitive flexibility tasks are lacking, and support for this as a construct is limited. For example, previous research has proposed that a Cognitive Flexibility factor generally comprised the Wisconsin Card Sorting Test (WCST) variables [16–19] but does not necessarily clearly describe or define the construct with any clarity. Anderson's model instead distinguishes Cognitive Flexibility as a sub-domain with its own attendant skills, including divided attention, working memory, conceptual transfer, and feedback utilisation. This is a more precise conceptualisation that clearly defines the distinct components that are contended to be subsumed by Cognitive Flexibility.

Another fundamental skill to cognitive performance is processing speed. It is a key assertion of this paper that speed and accuracy need to be considered when conceptualising EF. The speed-accuracy trade-off has been well documented in the literature, whereby slowed processing speed is implicated when tasks increase in complexity, or vice versa, depending on goals, payoffs or motivation. Indeed, processing speed is a core component in the psychological assessment of intelligence, and has been examined and identified as

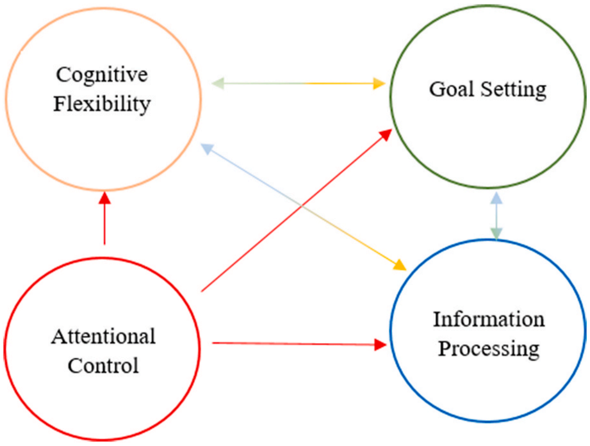


Fig. 1. Anderson's Proposed EF Model. Adapted from Anderson [11] with permission from personal communication October 12, 2021. The depicted constructs are suggested to involve integrated cognitive processes sharing frontal systems that receive and process information from posterior, motor, and subcortical regions.

an influential component to performance decrements in some clinical disorders. However, whilst many ‘information processing’ or ‘speed of processing’ factors have been proposed post hoc, it is rare to see this as a considered, distinct and measureable construct in an overall model of EF. Contention exists whether speed of processing is domain specific, or more globalised, which most likely contributes to the lack of consistency in the description of this construct where it has been variously labelled as information processing, complex attention, cognitive speed, reaction time, or even psychomotor speed [20]. Therefore, it is another strength of Anderson’s model that it includes speed as a construct in the operationalisation of EF task performance. A satisfying aspect of the Information Processing factor within Anderson’s model is that it proposes speed of processing as more than simply an ‘ingredient’ function. Instead, it is conceptualised as domain specific, bidirectional, and incorporating more than just the speed at which one completes a task. Within the Information Processing construct, efficiency, fluency and overall speed of processing are proposed, where it is thought that Cognitive Flexibility and Goal Setting are underpinned by Information Processing, thereby acknowledging the strength of association of IP, and also a degree of hierarchical recruitment of skills.

Lastly, EF comprises the ability to problem solve, maintain and shift attention, inhibit pre potent responses, plan, implement various strategies, and utilise feedback, all of which are necessary to adhere to goal-directed behaviour [21]. Goal-directed behaviour is therefore arguably the essence of EF, according to well-accepted definitions [5]. However, there remains a failure to offer a model of EF that captures Goal Setting as a construct, and a paucity of goal setting or planning factors represented in models of EF (not just in isolation to a test). This is certainly a criticism that can be applied to Miyake’s well subscribed model of EF [22]. Goal Setting and its subsumed cognitive skills (e.g., planning, organisation) are typically the latent ‘outcome’ in models of EF, but it has rarely been explored how to operationalise how a person would actually set and achieve goals. One model that has attempted to operationalise planning or problem solving (i.e., setting goals) is Zelazo and colleagues’ [10] problem solving framework. Their model considers EF as a macro construct, with hierarchical processes that are outlined to effectively plan to solve a problem. This model falls short however, because it lacks sufficient detail in relation to specific skills and their influence to effective problem solving. Lezak’s [5] conceptualisation of planning provides the most clarity as it includes consideration of a sequence of component skills necessary for its effective execution. Ultimately, the accumulation of skills at the lower level in the chain of processing skills are used to assess planning or Goal Setting. Furthermore, she considers that slowed processing speed at the beginning or end of a task may demonstrate reduced productivity, and considers compensatory performance, where if one has inflexible thought then issues may arise when attempting solve problems. This may in turn reflect an inability to shift, which may lead to perseverative behaviour. Anderson’s model attempts to explain how goal-directed behaviour is an executive, as opposed to a binary outcome (goal achieved, yes/no). Therefore, it is considered a strength that Anderson’s model conceptualises Goal Setting to include the subskills of initiation, conceptual reasoning, planning and strategic organisation.

A variety of EF tests are used to inform the models of EF and possible composite skills that underpin it. The obscurity of EF undermines the assessment, identification and treatment of executive dysfunction, particularly in adults, however the flaw of EF both as a theory and in its assessment is that clinicians tend to draw conclusions inefficiently, due to the lack of clarity regarding how these skills are delineated. It is therefore the contention of this study that the existing models of EF cannot be clarified until validity is first established regarding its specific latent constructs, in turn creating validity within a given set of measures. Therefore, as a platform to begin this inquiry, Anderson’s paediatric model was used as a template to be applied in the assessment of EF for healthy adults. It is recognised however, that in terms of cognition, children are not simply “mini adults”, and adults are not “overgrown children,” and that developmental trajectories of different EF skills differ. For example, most skills begin their protracted growth and come ‘online’ at different ages, and plateau around late adolescence to early adulthood [11,15,23–32].

1.1. *Methodological shortfalls of EF investigations*

In an attempt to bring clarity to the term, factor analytic studies have been implemented, albeit with limited success when considered collectively. Attempted fractionation of the skills that comprise EF further compounds the problem, because various factors structures have been identified in a range of populations. Thus, whilst significant gains have been made towards the fractionation of EF, the plethora of factors are likely a result of the methodological flaws of previous research. Given the limiting theory and constraints of testing, this is not an unexpected shortfall. For example, fractionation of the skills of EF was demonstrated by Miyake et al. [22] using Confirmatory Factor Analysis (CFA) comprising a three-factor model that included shifting, updating and inhibition, which they termed ‘unity and diversity’ of EF. This approach helped to expand traditional views of EF away from a simple homunculus construct. Miyake and colleagues’ [22] work has guided many researchers to date, lending support to a task impurity problem where it is difficult to tease out one skill only, as a single assessment task can in fact be measuring multiple skills at once (hence the unity and diversity of their factor structure). However, their work is criticised for restricting the number of tests and preconceived factors that these might tap into (e.g., no planning, organisation) hence limiting the scope of the analysis, and validity of their conclusions. The most comprehensive approach to factor analysis was a study conducted by Testa and colleagues [33], who attempted to avoid previous methodological limitations (e.g., small sample size, limited test selection). The researchers employed a holistic approach using an Exploratory Factor Analysis (EFA) on 19 EF tests with a large sample of 200 healthy adults and found six factors to comprise EF. Although this study took a comprehensive approach, the outcome was to suggest another alternative structure, and consensus remains elusive. Furthermore, inconsistencies extend beyond the mathematical limitations that results in an excess of factors generated, but the nomenclature varies considerably. Despite this, researchers continue to employ factor analyses without due consideration of the limitations imposed by the statistical method (e.g., sample size, power analysis, type and number of tests used).

Thus, whilst factor analysis is a commonly used technique to investigate construct validity, they dictate harsh demands in relation to sample size, power, and the number of tests included for analysis. To that end, the validity of findings is questionable because theory is applied post hoc to mathematical findings, relying on purely statistical support to fractionate any given EF construct that is well known to draw on an array of skills. The end result can be that this process does little more than simply contribute to the proliferation of a list of skills. Given EF tests can measure a number of separate, albeit often related EF skills, defining uniquely separable skills is problematic because tasks rarely tap into a single skill. If such tasks exist, they are arguably not executive in nature. Thus, task impurity obscures the clear conceptualisation of EF skills, and a comprehensive theoretical approach that gives justice to the multifaceted nature of EF is paramount. This means that researchers either need to be comfortable with low intercorrelations between tasks within their factor structures, or that models should be developed giving primacy to theory, and with mathematical concerns as secondary.

1.2. Rationale

A crucial first step must be to clarify what the purported function of each test is, and including more than one test for each function. Perhaps overlooked or not appropriately emphasised in previous research, is the value of assessing congeneric models in isolation, prior to addressing a model in its entirety. Congeneric models allow theoretically supported, complex latent constructs to be analysed to examine the extent to which indicators represent a true generic score [34]. That is, they assume that tests measure the same latent trait-a variable that is not directly observed but is rather inferred (through congeneric modelling) from other variables that are observed and directly measured using neuropsychological tests. However, unlike other measurement models, congeneric models are not bound by such strong restrictions and therefore serve numerous purposes [35]. For example, they allow different scales or measuring instruments to be used. Furthermore, a unique contribution of congeneric modelling is that it also addresses the importance of maximising individual reliabilities of composite scores that account for differences in weighting or contributions to the latent variable, as well as accounting for error, thus providing an indicator of how reliably these scales represent the same underlying construct. In essence, congeneric modelling uses theory as a guide to test placement, and can be said to provide evidence of the construct validity of the individual items used to measure a particular latent trait [34]. The parallel argument therefore is that statistically, the use of Confirmatory Factor Analysis (CFA) via path analysis (SEM) to assess a full structural model of EF and the interrelationships between latent variables without assessing the validity of the tests it is founded on is flawed. Typically, an SEM approach is used to identify a series of tests, enter them into a statistical program, from which latent variables are identified and analysed for their relationships amongst each other. This approach often constrains the number of measures that can be included because this number is influenced by the number of participants. This approach also restricts the validity of the latent construct that is best represented by a selection of tests. This is because the emphasis is not on the placement of tests that represent an overarching latent construct (such as principle components analysis or factor analysis), but rather, the emphasis is on the relationship amongst these constructs, often neglecting the fundamentals of what is being used to measure the construct in the first instance. Therefore, it is essential to use theory as a guide for test placement.

It is the purpose of the present study to confirm the validity of Anderson's paediatric model of Executive Function in a healthy adult population using a comprehensive approach including congeneric modelling. Anderson's model has been established to provide a template for assessment, yet has not been validated in the adult literature. This was achieved using a multi-step process, whereby several tests from numerous assessment paradigms were selected for inclusion in the overall study, with a triangulated approach to selecting good quality tests that have appeared in the literature for healthy adults. Thus, in order for researchers to successfully produce a coherent, effective, and psychometrically sound model and battery of EF respectively, researchers need to be unified in their approach to adopting numerous and varied EF measures.

1.3. Aims of present study

At the broadest level it was aimed to explore the overall nature of EF, with respect to what skills it comprises, how best to assess them, and the manner in which latent constructs relate to one another to explain goal-directed behaviour.

Specifically, the present study aimed to use congeneric modelling to confirm the validity of Anderson's paediatric model of EF in a healthy adult sample. Separate congeneric models were constructed using each of Anderson's constructs in order to isolate the subskills represented by each (Attentional Control-AC, Cognitive Flexibility-CF, Information Processing-IP, Goal Setting-GS). It was further aimed to assess the predictive validity of Anderson's four constructs, and the interrelationships between them, by performing a series of standard regression analyses.

1.4. Hypotheses

The EF constructs of Anderson's model are mostly bidirectional with the exception of AC which is unidirectional and therefore hypotheses will only be outlined as such. Specifically, it was hypothesised that;

1) All four constructs (AC, CF, IP, GS) purported by Anderson will meet the statistical requirements to demonstrate a variety of construct reliability and validity thresholds.
2) Attentional Control will be the strongest predictor, explaining the greatest variance in all other latent constructs. Thus, attention will be a significant domain that warrants its theoretical consideration within a model of EF, and not separate to it.
3) Information Processing will be the second strongest predictor of other latent constructs, therefore demonstrating that IP is an influential component in a model of EF.

2. Method

2.1. Participants

Overall, 133 adults (42 male and 91 females) aged between 18 and 50 ($M = 29.68$, $SD = 7.46$) were recruited for the current study. The percentage of participants in each age bracket was: 18–19 years (5.3%), 20–29 years (56.4%), 30–39 years (27%), and 40–50 years (11.3%). Inclusion/exclusion criteria included that participants be aged between 18 and 50 years with no previous diagnosis of a neurological or developmental disorder (e.g., autism or ADHD), or significant intellectual impairment, and no current diagnosis of an untreated psychological disorder. Finally, participants must not have undergone cognitive assessment in the previous two years.

Data was excluded for one participant who demonstrated an IQ < 70. Mean IQ as a function of group membership is presented in Table 1, in addition to other key demographic variables.

2.2. Materials

The selected measures included: Wechsler Abbreviated Scale of Intelligence: Vocabulary and Block Design Subtests (WASI: 2-subtest) [36], Trail Making Test; A, B [37], Test of d2; concentrate [38], Stroop test (Victoria version) [39], Block Span forwards and backwards; trials correct [40], Digit span forwards and backwards; trials correct [41], Picture Arrangement; raw score [41], Verbal Fluency; Animals and FAS [42], 5-point test; total unique designs [43], Austin Maze; total time [44], Tower of Hanoi; total time [45], Rey-Osterrieth Complex Figure Test; copy (ROCFT) [46], Wisconsin Card Sorting Test; 64-card (WCST) [47], Test of Everyday Attention measures (TEA) [48]; Elevator Counting with Reversal, Visual Elevator, Telephone Search while Counting (dual task), Behavioural Assessment of Dysexecutive Syndrome measures (BADS) [49]; Rule Shift time 2, Key Search, Zoo Map 1.

All tests were scored traditionally with the exception of the ROCFT, Key Search, and Zoo Map which were derived to create a speed and accuracy overall efficiency score, as part of a larger study. For example, ROCFT derived score (total time/copy score); Key Search derived (total time/raw score); Zoo Map derived score (total time/inversed raw score). The inversed raw score was to obtain the same directionality of what constitutes a better score as a negative score can be obtained in this task which would indicate poorer performance). For these measures, a lower derived score would indicate a better efficiency of performance.

Only the Austin Maze, blocks and digit span, and the TOH were administered on an iPad however for only the TOH, it was to demonstrate the starting and end configuration of discs and the participant completed the puzzles on the disc and prong board. Research has demonstrated the equivalency of cognitive constructs being measured by both conventional and computerised versions.

2.3. Procedure

Ethics approval was granted from the Victoria University Human Research Ethics Committee, and the authors have complied with APA ethical standards in the treatment of their sample, in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki). Participants were recruited through snowball sampling via social media and flyers around the campus at the university. Informed consent was obtained whereby participants were required to sign consent forms prior to testing, after reading the plain language statement, and had the opportunity to ask any questions. Participants had the right to withdraw at any time with no negative consequences to them. Testing duration was approximately 3 h, and participants were given the choice to spread this across two sessions if they preferred. Tests were administered in visual and verbal counterbalanced order between participants order to minimise order effects. The number of participants that were administered each of the four variations were thirty-three in versions B, C, and D, and thirty-four for version A.

A comprehensive test selection process was employed, that aimed to mirror Testa and colleagues' [33] approach. Specifically, psychometric properties of approximately 50 tests were meticulously scrutinised in order to assess them against specific inclusion criteria, and the list was then reduced according to these criteria (see Table S1 in Supplemental Material 1). The next step was to confirm Anderson's purported latent constructs using congeneric modelling. Clarification of test relationships achieved overall by the data reduction phase facilitated the assessment of each of the latent constructs identified in Anderson's model (Attentional Control; AC, Cognitive Flexibility; CF, Information Processing; IP, and Goal Setting; GS). Firstly, if Anderson placed a specific test under a certain

Table 1
Demographic characteristics of the sample (N = 133).

	Overall		Overall range	Total
	M	SD		
Age	29.68	7.46	18–50	
IQ	110.57	11.30	84–138	
Certificate/trade/diploma/TAFE				18
Secondary education				29
Tertiary education				56
Postgraduate education				28
Not disclosed				2
ANZSCO	66.17	17.75	8.9–100	

Note: Education was classified as commenced or completed. ANZSCO = Australian and New Zealand Standard Classification of Occupations.

construct, the current study matched this. Other tests were then placed as close to the theoretical placement recognised by Anderson. The value of the current approach is that placement of tests was underpinned by theory at the outset, instead of being applied post hoc to mathematical findings, and a parsimonious, psychometrically robust battery of tests to best represent the constructs at each step of the model building process can be offered. The last step was to then assess the predictive validity and interrelationships between verified EF constructs. This was achieved through a series of regression analyses on constructs that were computed to create composite scores based on their reliability and validity as found through congeneric modelling. Doing so allowed for the assessment of directionality outlined in Anderson's model, and provided clarity with respect to how different skills contribute to the overall achievement of more complex cognitive tasks, which skills in the hierarchical chain of cognitive processes contribute more towards the overall performance of EF, and the degree to which the directionality of these skills interrelate.

2.3.1. Data handling, transparency and openness

We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. Data analyses utilised SPSS and AMOS version 26. All assumptions for congeneric modelling were assessed, and violation in relation to univariate and multivariate outliers was evident. Research utilising healthy samples and cognitive testing is likely to bias the spread of scores toward better performance than that observed in clinical samples. As some scores were limited in their range (i.e., floor and ceiling effects) and thus variability, one or two deviations from the top and bottom indices resulted in sensitivity to extreme outliers (e.g., ECR, TSC dual task decrement) and as such, only extreme outliers denoted by boxplots were treated. This left non-extreme univariate outliers in the data, which were numerous (and in fact too many to transform) and were left because they were legitimate values [50]. A specific issue related to the number of significant outliers and overall variability of performance associated with TSC dual task decrement was identified. Given variability in responses, TSC-E was used instead which refers to time per target weighted for accuracy of tone counting, but does not attempt to control for the decrement and individual variation in processing or psychomotor speed. Approximately 50 other values were truncated in the entire data set. Once raw scores were derived, extreme outliers decreased or did not occur. Further, Hair et al. [51] contend that SEM methods are robust to such violations.

Multivariate Normality (MVN) was checked as using Mardia's coefficient, and the critical ratio was greater than 1.96 on all models indicating potential threats to MVN (largest value was 5.402), therefore bootstrapped bias corrected confidence intervals were calculated to compare the p values with those generated without the bootstrapped standard error. There were consistently comparable p values and small biases evident, therefore the data was considered free from threats of multivariate non-normality.

To maintain integrity of sample size, and being cognizant of the importance of retaining variables for theoretical purposes, and to avoid specification error, Missing Value Analysis was conducted for all congeneric modelling. Analyses indicated that the permissible thresholds for MVA varied across the separate tests included. The TOH and the Austin Maze were found to violate these thresholds, and as a result the Goal Setting construct yielded a sample size of 113. The remaining constructs yielded a sample size of 133. For a detailed description of fit indices see (Table S2 in Supplemental Material 2). SEM is a large sample technique and much research has focused on appropriate sample sizes for SEM [51–53], and recommendations range between 100 and 200 [51]. However, there is no consensus on what constitutes an appropriate sample size [54] and considering some estimation models have been developed for as few as 60 participants despite some limitations [52], our sample seems reasonable. Hair et al. [51] and Raykov and Widaman [55] suggest there are a number of factors that impact upon, and therefore determine, adequate sample size. Based on their descriptions, *model misspecification* did not occur because a comprehensive approach was applied to include all theoretically relevant variables (as much as possible) and therefore this was not a problem that should be applied to the current data. Furthermore, rules surrounding *model size* suggests a minimum of 5 respondents per parameter, with 10 per parameter being adequate [51,56] and some have suggested as little as 3:1 or 2:1 [57]. All models apart from Goal Setting did not exceed the 10:1 guidelines, therefore the respondent to parameter ratio meets this assumption. However, Goal Setting required 13 parameters and with the reduction in sample size to $N = 113$, the minimum of 5 per indicator meets this assumption ($13 \times 5 = 65$). There were no extreme *departures from normality*, therefore an increase of 15 respondents per parameter was not needed, and therefore Maximum Likelihood Estimation could be used as the *estimation procedure,* which assumes multivariate normality, and therefore allows valid interpretations for smaller samples of 100–150 [51]. Nonetheless, for the reasons outlined above the current study's sample of 133 seems reasonable as it is supported by the parameters detailed above, however caution is advised when interpreting findings.

Cognitive data can yield results where higher scores can indicate either better (in the case of accumulating correct responses), or worse (in the case of a timed task) performance. Therefore, with respect to the interpretation of results, the hierarchical omega coefficient (also known as coefficient H) was used [58]. This statistic is considered to be a maximised optimum *construct reliability* measure for which, unlike standard methods, a negative sign does not impact the overall assessment of reliability [35]. This formula performs well on multidimensional measures by weighting each indicator based on factor loadings [59]. Coefficient H was used in conjunction with other traditional measures, (See Table S2 in Supplemental Material 2 for traditional reliability and validity formulae and fit indices) and its formula and guidelines is expressed below;

Equation 1. Construct/composite reliability according to coefficient H.

$$\mathrm{H} = \frac{1}{\left[1 + \left[\frac{1}{\frac{\lambda_1^2}{1-\lambda_1^2} + \frac{\lambda_2^2}{1-\lambda_2^2} + \ldots + \frac{\lambda_n^2}{1-\lambda_n^2}}\right]\right]}$$

where the λ's are the standardised factor loadings.

The datasets generated and analysed during the current study are not publicly available due to ethical constraints but are available from the corresponding author on reasonable request and permission from the appropriate human research ethics committee. Coding is not applicable. Materials have not been provided in text as they are either already made available in the public domain, or a private test battery, however see Supplemental Material 3 for further information.

3. Results

3.1. *Confirmation of Anderson's purported latent constructs in a healthy adult population using congeneric modelling*

Test selection processes yielded 23 variables for testing the confirmation of the four constructs in Anderson's model in a healthy adult population. Table 2 displays the means, standard deviation, z scores, and range of performance across the variables selected. Pearson's correlation between tests per construct can be found in Appendix A. Fig. 2 presents the results of the four congeneric models.

Four congeneric models were created to test the underlying structures of the latent constructs Attentional Control, Cognitive Flexibility, Information Processing, and Goal Setting (Fig. 2 & Appendix B for more detail). Results indicated that all four models upheld, and represented the latent constructs. However, individual factor loadings and reliabilities should be interpreted with caution, as some did not meet the recommended minimum cut off values (see Supplemental Material 2).

For all constructs, the standardised residuals were no greater than the critical value of 1.96 indicating a good fit, in addition to all indicator variables being significant. Furthermore all regression factor loadings and error variances were greater than the critical value, demonstrating significance, except for AM from the GS construct where the error variance was $p = .166$.

A parallel analysis was conducted on all four constructs to ensure the models were in fact congeneric, and results revealed the congeneric model was a significantly better model than the parallel.

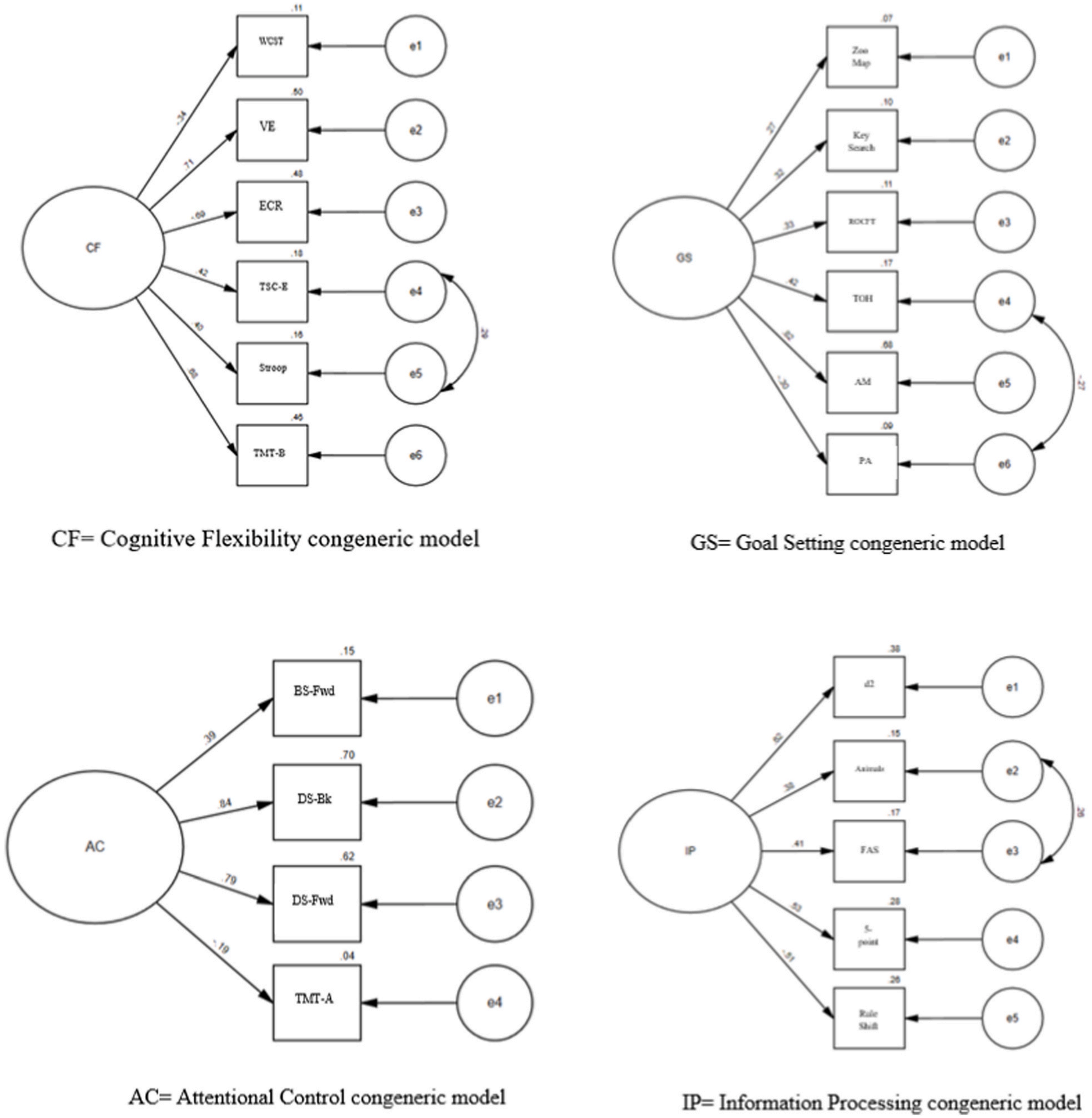
Table 2

Means, Standard Deviations, Raw scores, Z scores, and Possible Range of Performance of Variables Selected.

Variable	N	M *(SD)*	Minimum value *(Z score)*	Maximum value *(Z score)*	Range
Attentional Control					
BS-Fwd[a]	133	6.77 *(1.91)*	2.00 *(-2.49)*	11.00 *(2.21)*	0–16
BS-Bk[b]	133	6.55 *(1.84)*	3.00 *(-1.92)*	12.00 *(2.96)*	0–16
DS-Fwd[c]	133	6.74 *(2.31)*	2.00 *(-2.06)*	14.00 *(3.14)*	0–16
DS-Bk[d]	133	6.08 *(2.66)*	1.00 *(-1.91)*	14.00 *(2.98)*	0–16
TMT-A[e]	133	23.22 *(6.91)*	10.56 *(-1.83)*	49.81 *(3.85)*	0–180
Map Search[f]	133	68.72 *(8.65)*	43.00 *(-2.97)*	80.00 *(1.30)*	0–80
Cognitive Flexibility					
WCST total correct[g]	133	48.43 *(9.37)*	15.00 *(-3.57)*	59.00 *(1.13)*	0–64
VE[h]	133	3.77 *(0.86)*	2.00 *(-2.06)*	7.30 *(4.10)*	–
ECR[i]	133	7.90 *(2.23)*	2.00 *(-2.64)*	10.00 *(0.94)*	0–10
TSC-E[j]	133	2.95 *(0.88)*	1.70 *(-1.41)*	5.30 *(2.66)*	–
Stroop[k]	133	1.68 *(0.43)*	0.99 *(-1.63)*	2.83 *(2.67)*	0-∞
TMT-B[l]	133	57.18 *(17.80)*	20.06 *(-2.08)*	114.50 *(3.22)*	0–300
Information Processing					
d2CONC[m]	133	185.12 *(44.49)*	28.00 *(-3.53)*	294.00 *(2.45)*	−58–300
Animals total[n]	133	26.62 *(5.31)*	13.00 *(-2.56)*	41.00 *(2.71)*	0+
FAS total[o]	133	42.42 *(10.92)*	16.00 *(-2.42)*	71.00 *(2.62)*	0+
5-point total[p]	133	34.57 *(8.03)*	9.00 *(-3.18)*	53.00 *(2.30)*	0+
Rule Shift[q]	133	27.94 *(6.78)*	15.26 *(-1.87)*	56.50 *(4.21)*	0-∞
Goal Setting					
Zoo Map derived score[r]	113	7.37 *(6.60)*	0.85 *(-0.99)*	23.60 *(2.46)*	–
Key Search derived score[s]	113	5.50 *(4.55)*	0.54 *(-1.09)*	16.43 *(2.40)*	–
ROCFT derived score[t]	113	4.52 *(1.51)*	2.17 *(-1.55)*	11.17 *(4.40)*	–
TOH total time[u]	113	312.37 *(118.65)*	76.85 *(-1.98)*	727.80 *(3.50)*	0-∞
AM total time[v]	113	305.25 *(79.33)*	143.34 *(-2.04)*	483.07 *(2.24)*	0-∞
PA[w]	113	13.40 *(3.66)*	4.00 *(-2.57)*	19.00 *(1.53)*	0–22

[a]=Block Span Forward trials correct [b]=Block Span Backwards trials correct [c]=Digit Span Forward trials correct [d]=Digit Span Backwards trials correct [e]=Trail Making Test-A total time to complete [f]=Map Search [g]=Wisconsin Card Sorting Task total correct responses [h]=Visual Elevator timing score [i]=Elevator Counting with Reversal total correct [j]=Telephone Search while Counting weighted for accuracy of tone counting (measure of divided attention) [k]=Stroop test interference score [l]=Trail Making Test part B total time to complete [m]=d2 concentrate Hits-False Alarms [n]=Verbal Fluency semantic total admissible words in 3 min [o]=Verbal Fluency phonemic total admissible words in 3 min [p]=5-point total admissible designs in 3 min [q]=Rule Shift total time to complete [r]=Zoo Map 2 total time divided by total correct [s]=Key Search time divided by total correct [t]=Rey Osterrieth Complex Figure Task time divided by total correct [u]=Tower of Hanoi total time to complete all 13 trials [v]=Austin Maze total time to complete (trials to criterion; 2 error free trials) [w]=Picture Arrangement total raw score.

Note: Maximum ranges for Fluency measures (Animals, FAS, 5-point) have not been provided because outcome scores depend upon a time limit rather than a set possible score. Similarly, the Stroop, TOH, and AM have no maximum score as they are also timed. Lastly, the derived scores (TSC-E, Zoo Map, Key Search, ROCFT) do not have a range because these were computed for the purpose of the present study and can also vary based on completion time.



CF= Cognitive Flexibility congeneric model

GS= Goal Setting congeneric model

AC= Attentional Control congeneric model

IP= Information Processing congeneric model

Fig. 2. Path diagram of (AC) Attentional Control N = 133, (CF) Cognitive flexibility N = 133, (IP) Information Processing N = 133, and (GS) Goal Setting N = 113.

3.1.1. *Attentional control*

A good fitting model of AC was evident after the removal of the non-significant indicator Map Search ($p = .349$), and the removal of BS-Bk because in order to improve the model, modification indices suggested the error variances for BS-Fwd and BS-Bk needed to be covaried (M.I = 7.160, Par Change = .706), in addition to BS-Bk and TMT-A (M.I = 5.759, Par Change = −2.417) $\chi^2(2) = 1.61$, $p = .447$, RMSEA = 0.000, SRMR = 0.039 AGFI = 0.970, GFI = 0.994, CFI = 1.000, IFI = 1.004, NFI = 0.984. As indicated in Fig. 2 the squared multiple correlations (SMC) indicate the reliability of the indicator variables and demonstrated that the latent construct Attentional Control accounts for between 4 and 70% of the indicators. The AVE coefficient was 0.377 which is under the cut off of 0.5, thus more than half the variance for these indicators is not accounted for by the construct of Attentional Control. Although some individual loadings are acceptable (except for BS-Fwd and TMT-A) the variance explained falls short. However, construct reliability was evident since coefficient H returned a value of 0.806 indicating stronger maximised reliability, and discriminant validity was demonstrated as the model was free of redundancy.

3.1.2. *Cognitive flexibility*

Analysis of CF revealed a good fitting model was evident after modification indices suggested the error variance of TSC-E and Stroop needed to be covaried to improve the model (M.I. = 9.696, Par Change .085) $\chi^2(8) = 2.90, p = .940$, RMSEA = 0.000, SRMR = 0.026, AGFI = 0.980, GFI = 0.993, CFI = 1.000, IFI = 1.036, NFI = 0.980. As indicated in Fig. 2 the squared multiple correlations (SMC) indicate the reliability of the indicator variables, and demonstrated that the latent construct Cognitive Flexibility accounts for between 11 and 50% of the indicators. The AVE, the coefficient was 0.315 which is under the cut off of 0.5, thus more than half the variance for these indicators is not accounted for by the construct of Cognitive Flexibility. Although some individual loadings are acceptable (except for WCST total correct, TSC-E, and Stroop) the variance explained falls short. However, construct reliability was evident since coefficient H returned a value of 0.767 indicating stronger maximised reliability. Discriminant validity however was only somewhat demonstrated, and it could be argued on theoretical grounds that TSC-E and the Stroop test were covaried. This could be based on the premise of a shifting skill required for both tasks. It is also argued that they both require high level task demand in that for the Stroop one must inhibit a strong pre potent response, and although individually the demands of the TSC-E are simple, when combined they require one to concentrate and complete two tasks simultaneously. The correlation between the two was $r = .29$ indicating a weak positive relationship.

3.1.3. *Information processing*

Analysis of IP revealed a good fitting model was evident after modification indices revealed there was need for the error variance of Animals total and FAS total to be covaried to improve the model (M.I. = 4.619, Par Change 9.068) $\chi^2(4) = 1.15, p = .886$, RMSEA = 0.000, SRMR = 0.022, AGFI = 0.987, GFI = 0.996, CFI = 1.000, IFI = 1.041, NFI = 0.984. As indicated in Fig. 2 the squared multiple correlations (SMC) indicate the reliability of the indicator variables and demonstrated that the latent construct Information Processing accounts for between 15 and 38% of the indicators. The AVE coefficient was 0.247 which is under the cut off of 0.5, thus more than half the variance for these indicators is not accounted for by the construct of Information Processing. Although some individual loadings are acceptable (except for Animals and FAS total) the variance explained falls short. However, construct reliability was evident since coefficient H returned a value of 0.633 indicating stronger maximised reliability. Discriminant validity is cautiously assumed, however theoretical reasons suggest that covarying Animals and FAS is not a concern and is expected as they require the same skills to be used in a different way with a correlation of $r = 0.26$ indicating a weak positive relationship.

3.1.4. *Goal setting*

Analysis of GS indicated a good fitting model was evident after modification indices revealed there was need for the error variances of TOH total time and PA to be covaried to improve the model (M.I = 4.25, Par Change = −77.868) $\chi^2(8) = 7.22, p = .513$, RMSEA = 0.000, SRMR = 0.056, AGFI = 0.944, GFI = 0.979, CFI = 1.000, IFI = 1.014, NFI = 0.884. As indicated in Fig. 2 the squared multiple correlations (SMC) indicate the reliability of the indicator variables and demonstrated that the latent construct Goal Setting accounts for between 7 and 68% of the indicators. Of these, AM was the only indicator meeting the recommended limit of 0.5, where Zoo Map, Key Search, ROCFT, TOH and PA were well below the absolute minimum of .4. where AM was the only indicator above the recommended limit of 0.5-.7. The AVE coefficient was 0.203 which is under the cut off of 0.5, thus more than half the variance for these indicators is not accounted for by the construct of Goal Setting. Although only the AM indicator was acceptable and the variance explained falls short, construct validity was still demonstrated indicated by the overall model fit and since coefficient H returned a value of 0.730 indicating stronger maximised reliability. Discriminant validity is cautiously assumed, with the covarying of TOH time and PA. This could perhaps indicate non-verbal anticipation, organisation and planning. The correlation between the two was $r = -0.27$ indicating a weak negative relationship.

3.2. *Assessing the interrelationships of verified constructs in a healthy adult population using regression analyses*

Composite scores were calculated according to Jöreskog and Sörbom [60] based on reliable and valid constructs that were identified through congeneric modelling. This takes into account factor score weights of each indicator variable, variance of the factor, estimated (or implied) covariance matrix and error variances. Factor score weights are proportional to the factor loadings and error variances so as to avoid incorrect estimation if simple unit weight addition of scores were to be used.

Table 3

Descriptive data and Pearson's correlation coefficients of the four validated constructs.

Construct	N	M *(SD)*	Min	Max	IP	AC	CF
IP[a]	113	65.80 *(17.12)*	21.52	105.89			
AC[b]	113	6.00 *(1.91)*	1.54	11.55	.350**		
CF[c]	113	3.74 *(1.51)*	0.45	9.39	−.634**	−.424**	
GS[d]	113	36.14 *(9.35)*	20.27	59.81	−.484**	−.149	.475**

Note. ** correlation is significant at the 0.01 level (two-tailed).

[a] = Information Processing.

[b] = Attentional Control.

[c] = Cognitive Flexibility.

[d] = Goal Setting.

Table 4
Results of the regression analysis for cognitive flexibility.

Construct	B	β	t	p	Partial r	sr^2	R	$AdjustedR^2$
AC[a]	−.186	−.235	−3.21	.002	−.294	4.8%		
GS[b]	.036	.226	2.88	.005	.266	3.8%		
IP[c]	−.039	−.442	−5.35	.000	−.456	13.4%		
Overall model							.698	.473

Note. Dependent variable = CF.
[a] = Attentional Control.
[b] = Goal Setting.
[c] = Information Processing.

All assumptions to satisfy standard multiple regression analyses were conducted and met. Missing data were handled using acceptable techniques of Missing Value Analysis Cases were excluded listwise, yielding a final sample of 113 (due to missing data within the Goal Setting construct).

Composite scores were then used in a series of standard multiple regression analyses to determine the predictive validity and strength of interrelationships of each construct. Given the number of repeat analyses a more conservative alpha level was set to 0.01. Descriptive analyses and Pearson's correlation coefficients are presented in Table 3.

As indicated in Table 3, all constructs demonstrated a moderate to strong significant relationship between each other. However, AC was found to be non-significant and weak to GS.

3.2.1. *Cognitive flexibility*

A standard multiple regression was run to determine the predictive validity of AC, GS, and IP collectively on CF. A significant model was found $F\,(3{,}109) = 34.52$, $p < .001$, $R = 0.698$, $AdjustedR^2 = 0.473$ indicating that 47.3% of the variance in CF, could be predicted by AC, GS, and IP. Table 4 presents this information.

As indicated in Table 4, all variables were significant predictors, and Information Processing holds the strongest relationship to Cognitive Flexibility.

3.2.2. *Information processing*

A standard multiple regression was run to determine the predictive validity of AC, CF, and GS collectively on IP. A significant model was found $F\,(3{,}109) = 30.43$, $p < .001$, $R = 0.675$, $AdjustedR^2 = 0.441$ indicating that 44.1% of the variance in IP, could be predicted by AC, CF, and GS. Table 5 presents this information.

As indicated in Table 5, CF and GS were both significant predictors to the model, however AC was non-significant. Cognitive Flexibility holds the strongest relationship to Information Processing.

Table 5
Results of the regression analysis for information processing.

Construct	B	β	t	p	Partial r	sr^2	R	$AdjustedR^2$
AC[a]	1.03	.115	1.47	.145	.139	1.0%		
CF[b]	−5.32	−.470	−5.35	.000	−.456	14.2%		
GS[c]	−.447	−.244	−3.03	.003	−.279	4.5%		
Overall model							.675	.441

Note. Dependent variable = IP.
[a] = Attentional Control.
[b] = Cognitive Flexibility.
[c] = Goal Setting.

Table 6
Results of the regression analysis for goal setting.

Construct	B	β	t	p	Partial r	sr^2	R	$AdjustedR^2$
AC[a]	.467	.095	1.06	.290	.101	0.7%		
CF[b]	1.94	.313	2.88	.005	.266	5.3%		
IP[c]	−.174	−.319	−3.03	.003	−.279	6.0%		
Overall model							.537	.269

Note. Dependent variable = GS.
[a] = Attentional Control
[b] = Cognitive Flexibility
[c] = Information Processing.

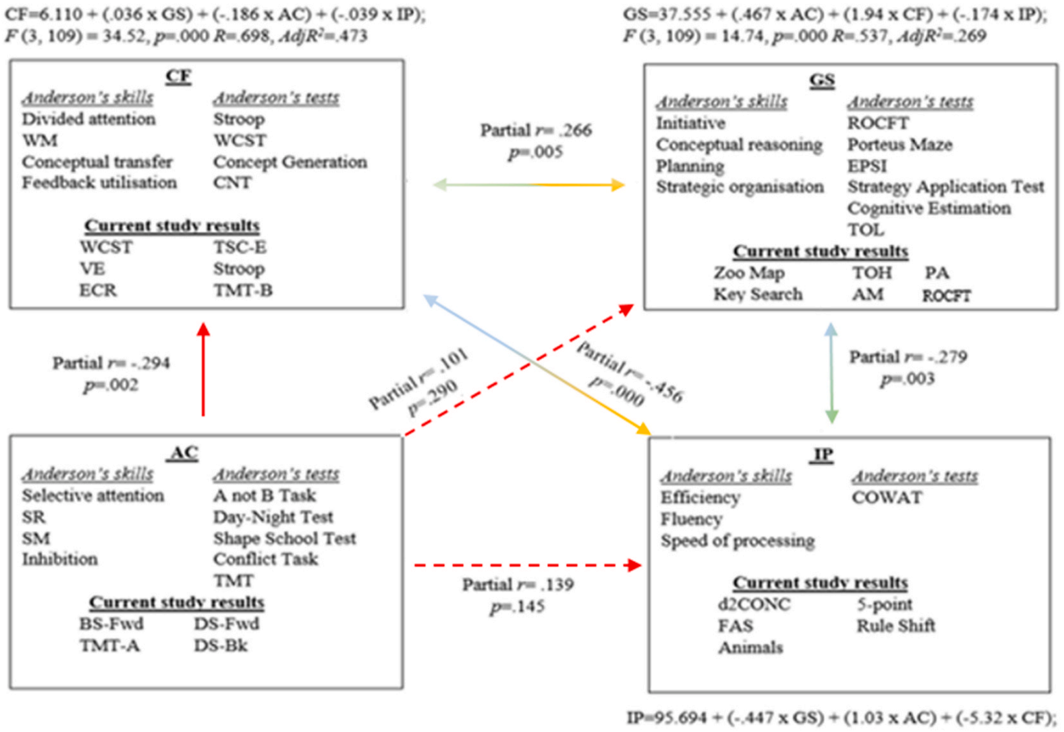


Fig. 3. Pictorial representation of regression analyses on constructs with subsumed tasks N = 113.

3.2.3. *Goal setting*

A standard multiple regression was run to determine the predictive validity of AC, CF, and IP collectively on GS. A significant model was found $F\,(3,109) = 14.74$, $p < .001$, $R = 0.537$, $AdjustedR^2 = 0.269$ indicating that 26.9% of the variance in GS, could be predicted by AC, CF, and IP. Table 6 presents this information.

As indicated in Table 6, both CF and IP were significant predictors, however AC was non-significant. Information Processing holds the strongest relationship to Goal Setting.

Overall results indicated that AC was not consistently statistically contributing to a model of EF. A pictorial representation of regression analyses on constructs with subsumed tasks and skills can be viewed in Fig. 3.

4. Discussion

4.1. *Confirmation of Anderson's purported latent constructs in a healthy adult sample using congeneric modelling*

This study is the first to demonstrate support for Anderson's paediatric model in a healthy adult sample. Results support the first hypothesis that all four constructs were statistically upheld, and therefore not only confirm the validity of Anderson's paediatric model, but also advances the pursuit of a psychometrically robust battery of EF that is underpinned by a theoretical framework.

4.1.1. *Attentional control*

According to Anderson [11], Attentional Control as a construct reflects self-regulation, self-monitoring, selective attention and inhibition, although with a healthy adult sample assessing these skills proved vulnerable to the limitations of traditional scoring. Although not all measures were retained in the final congeneric model, a significant proportion of variance of the overall construct was substantiated by the collective variables, and therefore the construct upheld.

Attentional Control (AC) was the only construct warranting removal of numerous tests due to the nature of some tasks not reaching statistical significance (Map Search) or being removed due to covariance (Block Span Backwards and TMT-A). Further, specific sub skills within this construct were unable to be assessed effectively (self-regulation and self-monitoring, inhibition). It is therefore argued that this construct best be reconceptualised as Simple Attentional Control (SAC) in healthy adults. This reconceptualisation is likely owing the fact that tasks used were clearly "too simple" to be considered executive in a healthy adult sample. Thus, Anderson's conceptualisation of AC in children may represent complex control at lower stages of development, however for adults, it does not (with this set of tasks). This is because maturation of the frontal lobes follows a developmental trajectory through to early adulthood, therefore influencing the way in which skills develop to underpin this construct. In essence, the emphasis or maturity is different, and therefore AC tasks load differently between the two populations. Thus, many of the tasks included in this construct were arguably facile for an unimpaired adult, due to cognitive resources and task load.

With respect to cognitive resources, selective attention theory models have proposed that stimuli are selected either early [61], intermediate [62], or late [63] in processing. High-load tasks use most, if not all of a person's resources thereby leaving no capacity to 'spill-over' to process and get distracted by irrelevant stimuli. In contrast, low-load tasks utilise few resources, thereby leaving 'spare' capacity to 'spill-over' and process irrelevant stimuli, meaning distractions can occur. Thus, it may be that the singular skills measured under the SAC construct might explain why this represents a simple task load in healthy adults. This is particularly evident with TMT-A, because beyond visual search, task requirements are to essentially count from 1 to 13, and Map Search is simply searching for symbols. These tasks may be too simple for a typical adult, to be considered "executive" in nature. Instead, both can be claimed to represent largely automatic skills, and not requiring of multiple, competing cognitive resources.

Secondly, simple task load may also explain current findings regarding span tasks. Traditionally, backwards span tasks should be more difficult than forwards as they require more effort [64]. However, complexity is a relative term, and the increase in difficulty between repeating stimuli forwards and backwards still may not be significant enough to warrant consideration as "executive" in terms of task load, as this is still reduced. Indeed, the current sample demonstrated this, as mean capacity difference between forwards and backwards of less than one unit.

Self-regulation and Self-monitoring sub skills were not effectively captured within the SAC construct. It was expected that any type of self-corrected errors would highlight some form of self-monitoring/regulation. However, participants were all healthy adults with IQ in the normal range who made few errors. In such a sample, it can be assumed that both self-regulation and self-monitoring are fundamental to all cognitive performance across a variety of domains, which may not be appropriately captured by error-based outcome scores. This is similar to Testa and colleagues' [33] study where they found a factor which they labelled Self-Monitoring and Set Maintenance which did not comprise errors, and instead comprised the Wisconsin Card Sorting Test (Factor B), Random Number Generation (repetition and cycling), and Contingency Naming Test (Factor C). This paper does not propose that self-regulation and self-monitoring skills are not inherent in Attentional Control, but does encourage a direction for future research to use tests that do not rely on errors to tap into these skills.

Finally, there was a lack of inhibition measures included. Although the Stroop is considered a prototypical measure of inhibition, this test was placed under CF in line with Anderson's description.

4.1.2. *Cognitive flexibility*

Cognitive Flexibility, as conceived in the Anderson model, is a construct that includes divided attention, working memory, conceptual transfer, and feedback utilisation. This construct was upheld in the current study, demonstrating the second highest amount of variance. It is a strength of current findings that highlight the notion that CF has been validated as an independent construct with its own set of attendant skills.

Consistent with previous definitions of cognitive flexibility that suggest a combination of multiple resources are essential for its execution [65], this study found that VE and TMT-B (i.e attentional switching, shifting, and cognitive flexibility) demonstrated the highest loadings and reliabilities under the CF construct. These findings strengthen the argument that any task that is dual in nature should be considered higher order.

4.1.2.1. Distinguishing attention and cognitive flexibility. The demarcation between AC and CF in Anderson's model, whilst not strictly upheld in this study, highlights the longstanding issue surrounding task purity of executive tasks. Effectively, a lower order test that represents a singular skill would be deemed a purer measure, as demonstrated by a higher amount of explained variance in the AC construct that was reconceptualised as SAC in the current study. By contrast, anything that is considered higher order (or higher in the chain of processing than simple attentional capacity) draws on multiple skills requiring parallel and/or switching processes, where one test is likely to recruit numerous skills. This would therefore explain the lower amount of shared variance in the CF construct. The demands of the additional skill introduce a level of task impurity that is reflected in the mathematical findings. This is particularly evident with the skills of inhibition and shifting, consistent with attention being fundamental to all cognitive performance [66], and also consistent with Kane et al. [67] who suggested that not all WMC tests are dual tasks, and instead, share processes with non-dual task measures. This might explain why some tasks not considered dual in nature were related to dual task measures and vice versa.

For example, as task demands increase and require one to carry out two tasks simultaneously or in close succession thereby becoming more complex, it is these fundamental skills that begin to work in an integrative manner that evolve to become more executive in nature. Therefore, inhibition would arguably belong under CF, consistent with the differences in developmental trajectories between skills. This is consistent with Garon et al. [68] who highlighted differences between simple and complex inhibition tasks are dependent on whether or not other skills are required, such as WM. This premise is further consistent with previous definitions regarding cognitive flexibility that suggest incorporating multiple sources of information simultaneously are essential for this construct [69]. This might explain current findings where the placement of inhibition is contentious, and why the Stroop test was an unreliable measure under CF, because there seems to be controversy if this is related to shifting or inhibition. It could be that inhibition may be functionally separable [22,70].

For example, if a healthy adult is required to engage in dual tasks, inhibition is a key component to Cognitive Flexibility because one must not only inhibit distractors (arguably a simple skill for healthy adults if used in isolation (e.g., SAC) but also draw on divided attention which requires one to consciously pay attention and distribute cognitive resources between two tasks at the same time [71]. Thus, an increase in task complexity implicates the cognitive resources and capacity available to divide one's attention. These skills are then used to incorporate WM to hold and manipulate information temporarily, where WMC uses controlled attention to maintain or supress WM [72,73], consistent with the placement of Elevator Counting with Reversal in the current study. Dajani and Uddin [70] refer to Hunter and Sparrow's work that suggests controlled attention is useful for any tasks that require goal maintenance, decision making, error monitoring, conflict resolution, effortful memory search or suppression of distracting information that is generally domain free [22], where greater WM demands are imposed when task shifting paradigms are used [70]. Thus, it is reasonable to suggest that for a child, the Stroop would be considered complex and higher order, which explains why Anderson placed this under the CF construct. By contrast, the Stroop test would clearly be facile for an unimpaired adult, as evidenced by the extreme lack of variability demonstrated in the current findings ($M = 1.68$, $SD = 0.43$), and therefore would be classified as a simple task measuring controlled, intended suppression of prepotent responses (given the automatic comprehension of the words). It could be argued that the healthy sample used (which inevitably dictates a low task load) did not make many self-corrections because they successfully performed the inhibition task, thus making the complexity of the tasks questionable.

Another reason to possibly explain why Anderson's descriptions between AC and CF overlapped besides upwards extrapolation of skills is the purported differences between shifting. Miyake and colleagues' [22] "Shifting" factor refers to shifting between tasks, mental sets, or operations, also referred to as 'attention shifting'. This involves "the ability to perform a new operation in the face of proactive interference or negative priming" [22] (p56). They discuss Posner and Raichle's idea that referred to this as executive-oriented shifts regulated by the frontal lobes, compared to spatial shifting that is regulated by posterior regions (e.g., parietal and mid-brain) [22]. This seems to be a term used to include a variety of shifting abilities, where as Dajani and Uddin [70] discuss differences between task and set shifting and their relative placement within a hierarchy. Set shifting has been described as the ability to shift attentional control *within* a task or between features of the same stimuli, typically considered a lower level form of cognitive flexibility [70]. This is consistent with the TMT-B in the current findings where one must switch between letters and numbers. Task switching, on the other hand, is more complex in nature because of the switching *between* tasks, with two different instructions [70]. This was demonstrated by the Telephone Search While Counting task of the TEA. One must listen to the tones, and also selectively attend to a visual stimulus simultaneously. Thus, whilst these two skills are posteriorly mediated and classified as simple in isolation, they become anteriorly mediated and more complex because of dual nature of the task (arguably). Together, these findings may explain why previous factor analytic studies have found a mixture of findings.

In essence, distinguishing between the two constructs SAC and CF may therefore come down to increased complexity stemming from the dual nature of a task (parallel or switching). It could be argued that SAC reflects bottom up, and CF reflects top down processes. Therefore, anything that is dual in nature should be considered higher order or complex attention, and this is likely to be subsumed within in the Cognitive Flexibility construct. Cognitive Flexibility should therefore be considered a complexity bound construct.

4.1.3. *Information processing*

Within the Information Processing (IP) construct Anderson suggests that one must be efficient and fluent. Together these processes describe how quickly and accurately one can produce an appropriate response. IP was a well-supported construct because all tests demonstrated consistency with respect to their loadings, however reliabilities were lower than desired.

These findings are consistent with previous research that found processing speed is a quantifiable domain capable of measurement in its own right, with various statistical extractions methods and populations [16,20,74–76]. Interestingly, in their PCA of mixed participants ($N = 92$) Chiaravalloti and colleagues [20] found a separable Information Processing domain. However, they also found that different measures of information processing loaded across various factors, suggesting IP may not be a unitary construct, and identified different factors relating to simple and complex processing speed. Furthermore, Pires et al. [76] found in their CFA that EF, Verbal Abilities (VA) and Processing Speed (PS) loaded separately, concluding that EF and Processing Speed are related, yet separable constructs.

The IP construct has been relatively underrated within cognitive models of EF, not only because of the lack of internal consistency of the construct itself, but mainly because it has received little consideration as an EF at all. A review of the literature suggests that IP is rarely considered "executive", and is often thought of as a separate cognitive skill. Therefore, a possible explanation for why many of the tests in the current study demonstrated low reliability in relation to the overall IP construct could be because speed is fundamental to *all* cognitive performance. This is consistent with multifactorial notions of EF. All of the tests within this construct have been purported at some stage as measures of speed within the literature, most likely because "processing speed has been used to refer to a variety of measures that get used in different ways that may tap underlying components to varying degrees" [77] (p270). Whilst it is not a novel finding that speed is fundamental to cognitive performance, it is likely the reason why studies have proposed processing speed as a globalised function [78–80], that is representative of a systemic mechanism that is not specific to a particular task [75]. It is further likely the reason why some have suggested it is important to remove the influence of processing speed to understand the contributions of other higher order cognitive processes [77]. Indeed, most processing speed measures are simple tasks which minimise the contribution of other higher cognitive functions [81]. Thus, the ambiguity in the construct leads researchers to question what the role of PS is to EF, and if it is best conceptualised as domain specific, or a more globalised function. Given the lower than desired reliabilities found in this construct, it may be sensible to suggest that whilst IP was found to be a measurable construct, it may also in fact be a globalised function that both recruits and underpins a variety of different skills.

It is the contention of this paper to suggest that IP, whilst not traditionally conceptualised as an EF, is a critical foundation skill to overall EF, much in the same way as attentional control. In this study and more broadly, IP is a construct that is consistently measured using EF type paradigms [20,77]. Therefore, this measurable construct extends beyond simple reaction time measures as suggested in some previous studies [74,82]. It could be argued that the IP construct shares notional similarities with SAC, and is therefore best conceptualised as a fundamental construct that underpins higher order performance, however these constructs should not be considered similar beyond the notional level of complexity.

The study findings argue strongly that whilst attention can be demarcated based on complexity, where more simple attention should not be considered 'executive' and more complex attention is arguably better defined as 'Cognitive Flexibility', IP when measured beyond simple RT, is unequivocally an Executive Function. Furthermore, it is likely an EF that permeates all aspects of cognition, hence the generalised nature of its definition in previous studies. In any model of EF, especially Anderson's Executive Function model, it is certainly evident that IP provides additional explanatory value, and strong validity, as all tasks under IP were strong indicators and none were removed.

The IP construct focusses on the strategic production of output under time constraints, and the level of complexity of a task contributes to the speed with which it can be completed. There are clear distinctions between individuals in basic speed of processing capacity, but within a model of EF it is the ability to maintain speed when complexity is added that is key. Essentially, efficiency within this construct refers to faster performance with more correct responses, where a lower score indicates better efficiency, which relies on intact frontal systems [11].

4.1.4. *Goal setting*

GS as a construct reflects initiative, conceptual reasoning, planning, and strategic organisation [11]. Not all purported skills (e.g., initiation which is more qualitatively measured) were explicitly captured by the included tests, and some tests demonstrated very weak loadings and reliabilities. This may explain why this construct explained the least amount variance overall, further supporting the notion of task impurity. Reconciling the mathematical limitations of including the set of tests in this construct meant that poor indicators were retained, and only one test met the minimum requirement for SEM. However if all were removed, the construct itself would not be validated. It is therefore a strength of congeneric modelling that allows theory as a guide to test placement, and the model achieved a significant fit overall, confirming the latent construct of Goal Setting.

For a researcher and clinician alike, the inability to tease out individual skills is a potentially vexatious issue. Understanding and distinguishing between separate cognitive skills is desirable in bringing precision and clarity to both theoretical and clinical outcomes. Whilst this may be more simply achieved for lower level skills, task impurity makes this considerably murkier when considering higher order functions.

A review of the literature highlights that goal setting/formation is usually treated as an outcome rather than a measurable construct in its own right, and subsequently a paucity of validation of this construct exists in the literature. Goal setting is most certainly not a singular cognitive skill. The formulation of goals requires the recruitment of an accumulation of skills that might be considered lower order in the chain of processing. To set a goal, one must plan, and in order to plan one must use an array of skills that might include for example, working memory to temporarily hold and manipulate information, inhibition of competing or distracting stimuli, and dividing and/or shifting of attention. Thus, skills preceding the GS construct such as those in CF, AC, and IP could be thought to play a role in effective execution of a goal. The only model that could be found that includes Goal Setting as a meaningful construct is the problem solving framework [10]. This framework highlights a hierarchy of steps that are necessary for one to formulate and execute a plan. However, there is a failure to consider which skills at the bottom of the hierarchy are necessary for advancement, and how they work within a set of skills to contribute to EF as a 'macro construct'. For example, a person may be capable of completing complex tasks, however only when asked to do so, and simply lack initiative where it does not occur for them to do anything [5]. Therefore, previous approaches to validate this construct have failed because mathematical procedures fail in the support of higher order constructs due to task impurity. It is therefore argued that it is more meaningful to use theory as a guide to group skills to represent this construct.

Study findings indicated the Austin Maze was the strongest indicator and demonstrated the strongest reliability, suggesting the AM to be a very good fit within Goal Setting. Similarly, the TOH was an arguably strong indicator, however the reliability index was poor, possibly due to the error variance being correlated with PA, suggesting there may be an element of shifting of sequences or mental visualisation and planning common to both. It could be argued that the Austin Maze, given its strength of relationship to the construct, is measuring something else other than 'goal setting' or 'planning'. There is a lack of consistency in the literature regarding what each task measures. For example, the Maze has variously been purported to measure learning efficiency [83], prospective working memory [33], and planning [84,85] and Tower tasks as measuring inhibition [22], working memory capacity [86], and problem solving [87]. This further strengthens the contention that any task considered higher order is likely to encompass a variety of skills, contributing further to the notion of task impurity.

The results of the current study advise caution when using ROCFT, Key Search and Zoo Map as measures of strategic organisation, planning, and reasoning tasks respectively in healthy adults due to these tests resulting in the lowest reliabilities for the GS construct (and indeed across all constructs). It could be that the ROCFT is hampered by inefficient scoring procedures. Traditionally, the ROCFT has continued to be used in both research and clinical practice in paediatric and clinical populations, however a search of the literature identified that it is rarely used in studies of EF, in particular with healthy adult samples. It was selected for inclusion in this study following Anderson. Similarly, Key Search and Zoo Map (selected based on inclusion in Testa's study) have scarce support for their reliability and validity in relation to what they are measuring. Traditionally, these tests form part of the Behavioural Assessment of Dysexecutive Syndrome (BADS) test battery which has been used extensively in clinical practice, however individual subtests are rarely used in either research or clinical practice. Bennett et al. [17] explored the BADS using a clinical population and found the Zoo Map test loaded onto a factor with other measures, and likely reflected sequencing and self-monitoring. They also found that Key Search loaded onto a factor with other measures labelled as initiation and cessation, and control of action. However, interestingly no planning or goal setting factors were found, when these are stated to be tests of planning. The one study similar to this project that used these measures in healthy adults was Testa et al. [33] however, their study loaded Zoo Map onto a factor which they labelled as Task Analysis, and Key Search did not make it to the final factor solution.

In summary, the complexity of the GS construct stymies the ability to clearly define and assess it, as is demonstrated by the paucity of validation evidence within the literature. This is because task impurity impedes the ability to break down the multifactorial skills that are required to be integrated to reach an end state goal. Whilst other skills that are defined by dual tasks also have this problem, such as those falling within the CF construct, the issue escalates as tasks increase in complexity, or require the integration of more than two skills simultaneously. Thus, Goal Setting is arguably the most complex construct, because it is proposed to subsume such skills as planning, strategic organisation, initiative and conceptual reasoning, which are all complex and multifactorial concepts in themselves. The results of the current study represent a unique advancement because a formidable construct that has remained challenging to define and validate has been offered and supported statistically by giving primacy to its theoretical basis. Goal Setting as it is operationalised in this study, reflects the highest level of EF able to be assessed by existing measures and should be considered at the apex of complexity, with other latent constructs beneath it conceptually.

4.2. Predictive ability of the model

This paper has outlined support for Anderson's concepts of Attentional Control (re-conceptualised as Simple Attentional Capacity) Information Processing, Cognitive Flexibility and Goal Setting as separable constructs. Regression analyses were then conducted to assist with model confirmation by defining the relationships between them, and to assist in understanding whether a hierarchical organisation was supported. Study findings indicated all constructs were unique in their contribution and played a significant role within a model of EF, as can be seen in Fig. 3. Results indicated that AC only contributed with statistical significance to CF, and not to IP and GS. Therefore, the second hypothesis that stated Attentional Control would be the strongest predictor, explaining the greatest variance in all other latent constructs was not supported.

It was also interesting that the unique contribution as indicated by partial coefficients demonstrated that IP and CF demonstrated the strongest relationship to each other. Therefore, it could be argued that these two are considered important constructs in the overall conceptualisation of EF. Furthermore, given the significant positive relationship between CF to GS, it could also be implied that a hierarchy of skills is evident in the overall conceptualisation of EF. For example, this positive relationship suggests that as one's ability to shift and divide attention, and to inhibit irrelevant stimuli in order to use WM resources to maintain and manipulate information, so too does their ability to set goals and plan out how to achieve them effectively in an organised fashion. Essentially, those tasks that are dual in nature, should facilitate effective goal formation.

The third hypothesis stated that IP would be the strongest predictor after AC in the model. However overall, Information Processing was the strongest, and therefore most influential predictor thereby providing only partial support for the hypothesis. This result demonstrates that IP is an influential component in a model of EF. Furthermore, a negative relationship was found between both IP and CF, and IP and GS which is consistent with Anderson's premise that bidirectional relationships may account for the enhancement or decrement in tasks that mediate performance at a higher level. For example, given that IP in the current study has been conceptualised as a capacity bound construct, it could be that the speed-accuracy trade-off is evident. The allocation of resources for one to be fluent and efficient (IP) decreases as the complexity of task (CF, GS) increases. Similarly, where the complexity of task (CF, GS) decreases, the capacity for resources to be efficient and fluent (IP) in task performance increases. In sum IP was shown to be an influential construct as its capacity bound nature is likely to be embedded in all performance outcomes. When conceptualising CF and GS as complexity bound, it follows that the higher an individual's capacity, the more resources they may allocate to tasks increasing in complexity.

Study findings demonstrated that GS as represented by the current study was the only construct that was the most 'latent' and 'intangible', consistent with previous research. However, this study has found that skills below in the chain of processing, specifically those under CF and IP constructs, are influential skills that contribute towards effective GS. The strength of the relationship between Cognitive Flexibility and Goal Setting demonstrates a degree of overlap in support of the notion that CF skills contribute to GS outcomes. At the same time, the strength of the relationship also supports the notion that they are separable constructs, and lends weight to the premise of a hierarchy of EF skills. The same logic could be applied to the relationship between IP and GS, which demonstrated a similar pattern. Not only do these results confirm the premise of Anderson's paediatric model of EF for use with healthy adults, but they also support Miyake and collegues' [22] premise that EFs are both unified and diverse, and move us away from the once applied premise of viewing EF as a homunculus.

4.3. Limitations and future directions

First, sample size was a limitation. Although sample size was deemed appropriate for the current study, a larger sample size would be able to demonstrate greater strength of the individual loadings and therefore overall model. Sample size was a contentious issue given the large test battery taking approximately 3 h to complete. However, one of the many strengths of the current study was the ability to take a large test battery and reduce this data to manageable form, thereby aiding test selection. Therefore, these findings will be beneficial for future researchers to be selective in their measures and narrow down test selection based on the current findings. Therefore, this study can be replicated in not only a sample of healthy adults, but also older populations using the selected tests, which would mean that the time to complete the test battery would be significantly reduced, in turn aiding the ability to obtain a larger sample size.

Demographic constitution of the study sample reflected a fairly homogenous group with respect to SES and intelligence, and therefore results of the current study are able to be generalised to the wider population. However, the large number of female participants compared to males is noteworthy. Although there have been mixed findings with respect to sex differences in cognitive performance on EF measures, a more even distribution would have been desired. Future research may wish to evaluate this model as a function of sex. In all variables measured for research, variability of scores is essential to uphold assumptions for statistical testing. The issue of outcome scores based on errors or time exemplifies this. Errors by nature will elicit a binary outcome (correct or incorrect, yes/no) which drastically limits variability. In contrast, time encapsulates an almost infinite spectrum of performance, and this variability is necessary for statistical testing. However, it seems that regardless of the type of outcome measure used for a higher order task, partitioning variance will always remain difficult because of the multifactorial skills subsumed for effective performance on a complex task. Path analysis using SEM would have provided more robust conclusions regarding causality amongst the set of variables, however limitations imposed by sample size made regression analyses most appropriate.

This paper has demonstrated that there are multiple ways in which to measure EF, supported by a variety of tests underpinned by a variety of skills, consistent with EF theory. It is acknowledged that interpretation of current findings is constrained, particularly when the labelling of skills, tests, and constructs alike also contribute to the proliferation of taxonomies, much the same way as previous research. Thus, it is not to say that congeneric modelling is the best way to examine EF, but is rather an alternative view offered towards obtaining a clearer conceptualisation of the skills that comprise it.

The current findings have significant implications that can be applied to the implementation of reliable and valid assessment methods to aid the diagnosis and intervention procedures of a range of adult disorders with inherent executive dysfunction. A reliance on a variety of skills is not only necessary, but essential, to “enable a person to engage successfully in independent purposive self-serving behaviour” [5] (p35).

4.4. Conclusion

EF theory is mired to using task impurity as an explanation for all the shortcoming of EF assessments. However, the lack of shared variance between tests within a construct should be valued because one test is likely to reflect numerous skills, which is why statistically partitioning variance will always remain challenging. Rather than viewing the limited amount of variance explained by a construct as a shortfall of EF theory, perhaps a paradigm shift is necessary where the application of the pejorative term ‘task impurity’ can shed some light to what this actually means for researchers. We would expect the shared variance explained, and reliabilities of individual variables to be lower than usual for higher order EF constructs because of the multifactorial nature of EF that require many different, yet related, EF skills. Therefore, when a construct fails to explain a large proportion of variance it should be clear that these tasks are more executively driven, and perhaps not poor fitting measures because of the array of skills required to complete a higher order task. Thus, the premise of shifting the traditional view of understanding these as poor measures is necessary, and instead, the lack of variance should be valued because it signifies the recruitment of higher order skills.

Finally, the work of Peter Anderson [11] almost two decades ago has proven an exceptional platform from which to explore definitions, tests and constructs of EF. This paper has reconciled various issues highlighted within the literature. In summary, these results confirm that the simplicity of Attentional Control as represented by more posterior attention tasks, does not significantly explain EF performance at the higher levels. SAC is redundant in a model of EF in healthy adults. Complex attention is required instead, where this may be represented as one of the constituent skills of CF. Furthermore, considering GS has been conceptualised as the epitome of Executive Functioning, and at the top of the hierarchy of EF skills, it seems reasonable to apply this construct as the outcome or end state goal. This is consistent with most definitions of EF, where complexity bound skills lower in the chain of processing predict performance at a higher level. EF assessment has long been limited by a lack of sound theoretical underpinnings of the mechanisms and processes that explain performance. New models of EF must consider the role of Information Processing speed not only as a mere component, but as a central element that likely mediates the relationship between lower order and higher order skills.

With cautious consideration of the caveats outlined in this study, the included measures in all constructs could be considered a parsimonious battery of Executive Function that will allow both clinician and researcher alike to delineate foundational skills from those at the apex of higher order functioning. Thus, these findings provide further understanding of the complex cognitive processes associated with EF. There is no ‘best’ way to measure EF because multiple skills are used in different ways. However, this study has helped identify which skills comprise EF, and highlights the importance of the relationships between constructs that work in an integrative manner, in turn helping us to understand the way in which the mind coordinates complex cognitive processes. Ultimately, this paper has brought clarity towards the clearer conceptualisation of the nebulous construct that is Executive Functioning.

Acknowledgements

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix A. Descriptives and Pearson's Correlations of Test Indicators Per Construct

Table A.1. Means, Standard Deviations, and Pearson's Correlations of Indicators to Test the Construct Attentional Control

Attentional Control							
	N	M*(SD)*	1.	2.	3.	4.	5.
1. BS-Fwd	133	6.77 *(1.91)*	–				
2. BS-Bk	133	6.55 *(1.84)*	.352**	–			
3. DS-Fwd	133	6.74 *(2.31)*	.297**	.211*	–		
4. DS-Bk	133	6.08 *(2.66)*	.328**	.355**	.660**	–	
5. TMT-A	133	23.22 *(6.91)*	−.168	−.266**	−.156	−.143	–
6. Map Search	133	68.72 *(8.65)*	.249**	.127	.075	.004	−.339**

$^{*}p < .05$. $^{**}p < .01$.

Table A.2. Means, Standard Deviations, and Pearson's Correlations of Indicators to Test the Construct Cognitive Flexibility

Cognitive Flexibility							
	N	M*(SD)*	1.	2.	3.	4.	5.
1. WCST total correct	133	48.43 *(9.37)*	–				
2. VE	133	3.77 *(0.86)*	−.211*	–			
3. ECR	133	7.90 *(2.23)*	.238**	−.514**	–		
4. TSC-E	133	2.95 *(0.88)*	−.161	.287**	−.255**	–	
5. Stroop	133	1.68 *(0.43)*	−.093	.278**	−.280**	.414**	–
6. TMT-B	133	57.18 *(17.80)*	−.264**	.465**	−.449**	.338**	.290**

$^{*}p < .05$. $^{**}p < .01$.

Table A.3. Means, Standard Deviations, and Pearson's Correlations of Indicators to Test the Construct Informational Processing

Information Processing						
	N	M*(SD)*	1.	2.	3.	4.
1. d2CONC	133	185.12 *(44.49)*	–			
2. Animals total	133	26.62 *(5.31)*	.237**	–		
3. FAS total	133	42.42 *(10.92)*	.250**	.377**	–	
4. 5-point total	133	34.57 *(8.03)*	.307**	.202*	.266**	–
5. Rule Shft	133	27.94 *(6.78)*	−.335**	−.192*	−.169	−.264**

$^{*}p < .05$. $^{**}p < .01$.

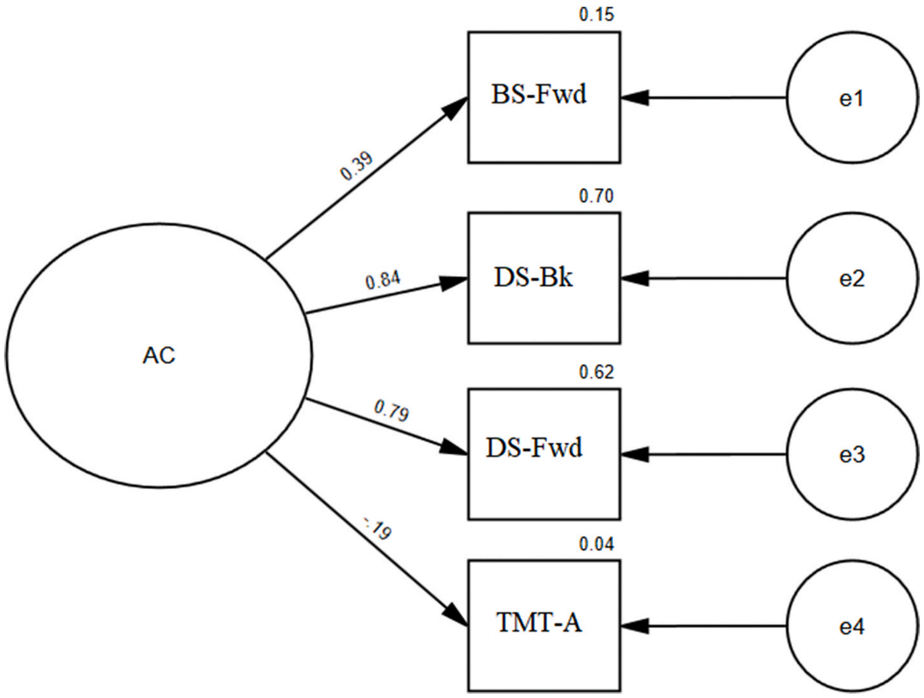
Table A.4. Means, Standard Deviations, and Pearson's Correlations of Indicators to Test the Construct Goal Setting

Goal Setting							
	N	M*(SD)*	1.	2.	3.	4.	5.
1. Zoo Map derived score	113	7.37 *(6.60)*	–				
2. Key Search derived score	113	5.50 *(4.55)*	.156	–			
3. ROCFT derived score	113	4.52 *(1.51)*	.222*	.111	–		
4. TOH total time	113	312.37 *(118.65)*	.052	.001	.158	–	
5. AM total time	113	305.25 *(79.33)*	.201*	.277**	.245**	.363**	–
6. PA	113	13.40 *(3.66)*	−.095	−.079	−.128	−.363**	−.242**

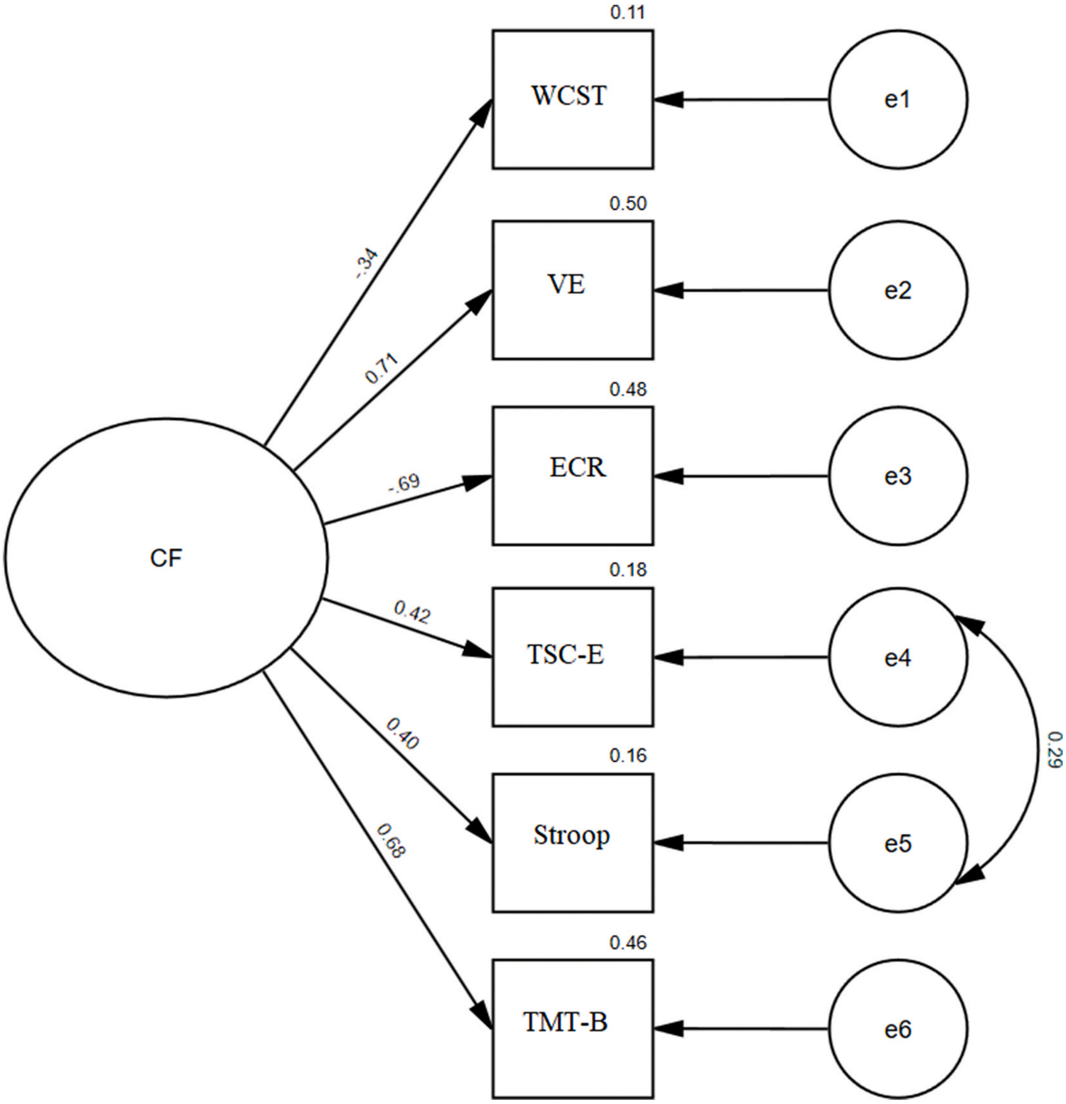
$^{*}p < .05$. $^{**}p < .01$.

Appendix B

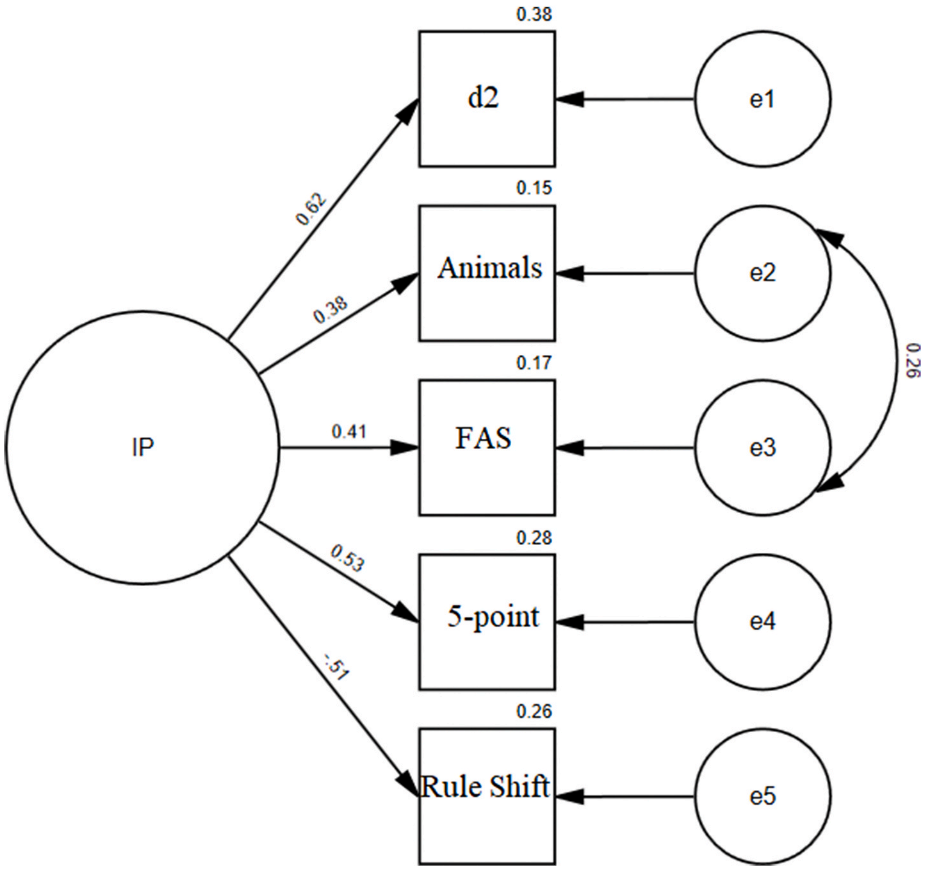
Congeneric Models



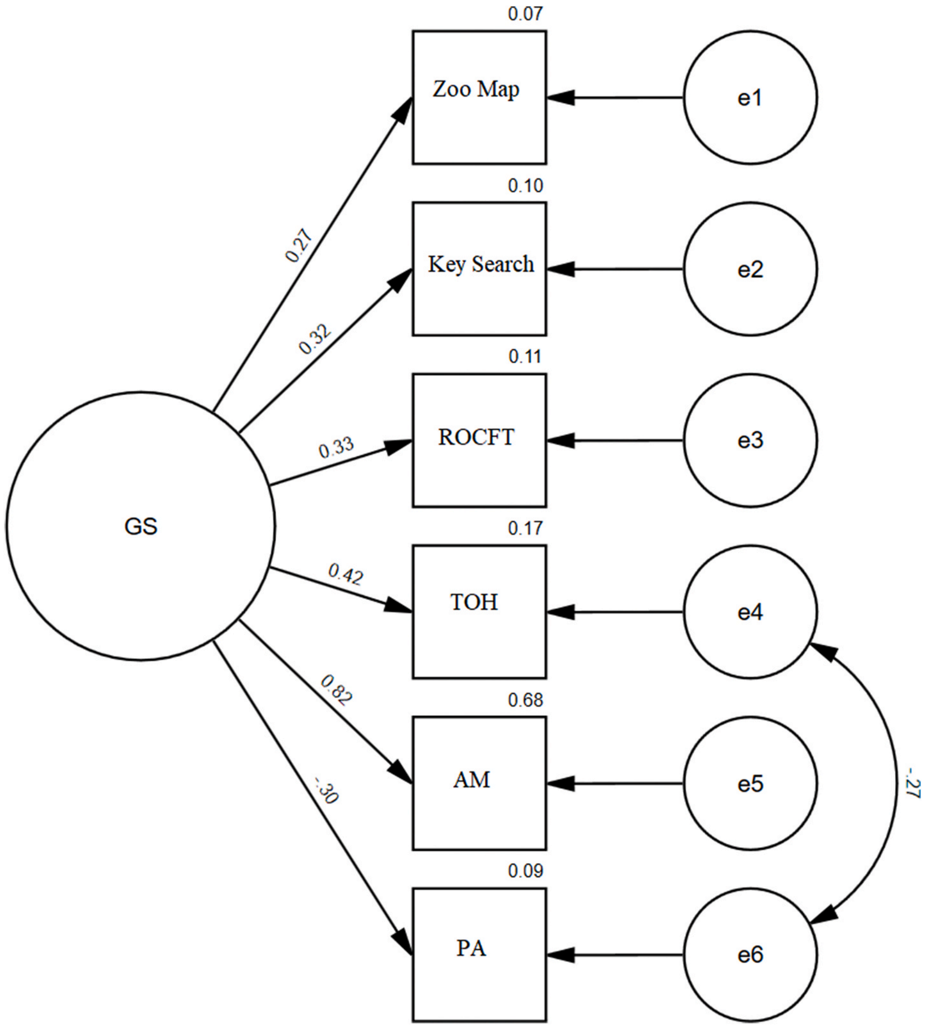
Note: N = 133. Block Span and Digit Span; trails correct, TMT-A; total time to complete



Note: N = 133. WCST; total correct. VE; timing score. ECR; correctly counted strings. TSC-E; dual task decrement without motor component. Stroop; C/D. TMT-B; total time to complete



Note: N = 133. d2; Concentrate score. Verbal Fluency; Animals and FAS total words. 5-point; total unique designs. Rule Shift; total time to complete task 2



Note: N = 113. Zoo Map; total time/inversed raw score. The inversed raw score was to obtain the same directionality of what constitutes a better score as a negative score can be obtained in this task which would indicate poorer performance. Key Search; time/raw score. ROCFT; time/copy score. TOH; total time to complete. AM; total time to complete. PA; raw score.

Appendix C. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2023.e15504.

References

[1] A.M. Adams, S.E. Gathercole, Phonological working memory and speech production in preschool children, J. Speech Hear Res. [Internet] 38 (2) (1995) 403–414, https://doi.org/10.1044/jshr.3802.403.

[2] T.P. Alloway, S.E. Gathercole, S.J. Pickering, Verbal and visuospatial short-term and working memory in children: are they separable? Child. Dev. [Internet] 77 (6) (2006) 1698–1716, https://doi.org/10.1111/j.1467-8624.2006.00968.x.

[3] A. Baddeley, G.J. Hitch, Working memory, in: G.H. Bower (Ed.), The Psychology of Learning and Motivation : Advances in Research and Theory, Academic Press, New York, 1974, pp. 47–89.

[4] R.A. Barkley, Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD, Psychol Bull [Internet] 121 (1) (1997) 65–94, https://doi.org/10.1037/0033-2909.121.1.65.

[5] M.D. Lezak, The problem of assessing executive functions, Int. J. Psychol. [Internet] 17 (1–4) (1982) 281–297, https://doi.org/10.1080/00207598208247445.

[6] M.D. Lezak, D.B. Howieson, D.W. Loring, H.J. Hannay, J.S. Fischer, Neuropsychological Assessment, fourth ed., Oxford University Press, New York, NY, 2004.

[7] D. Norman, T. Shallice, Attention to Action: Willed and Automatic Control of Behaviour, in: R. Davidson, R.G. Schwartz, &. D. Shapiro (Eds.), Consciousness and Self-Regulation: Advances in Research and Theory, Plenum Press, New York, NY, 1986, pp. 1–18.

[8] T. Shallice, P.W. Burgess, The domain of supervisory processes and temporal organization of behavior, Phil. Trans. Roy. Soc. Lond. B: Biological Sciences [Internet] 351 (1996) 1405–1412, https://doi.org/10.1098/rstb.1996.0124.
[9] D.T. Stuss, T. Shallice, M.P. Alexander, T.W. Picton, A multidisciplinary approach to anterior attentional functions, Ann. N. Y. Acad. Sci. [Internet] 769 (1) (1995) 191–211, https://doi.org/10.1111/j.1749-6632.1995.tb38140.x.
[10] P.D. Zelazo, A. Carter, J.S. Reznick, D. Frye, Early development of executive function: a problem-solving framework, Rev. Gen. Psychol. [Internet] 1 (2) (1997) 198–226, https://doi.org/10.1037/1089-2680.1.2.198.
[11] P. Anderson, Assessment and development of executive function (EF) during childhood, Child. Neuropsychol. [Internet] 8 (2) (2002) 71–82, https://doi.org/10.1076/chin.8.2.71.8724.
[12] H.L. Bodimeade, K. Whittingham, O. Lloyd, R.N. Boyd, Executive function in children and adolescents with unilateral cerebral palsy, Dev. Med. Child. Neurol. [Internet] 55 (10) (2013) 926–933, https://doi.org/10.1111/dmcn.12195.
[13] B. Høie, A. Mykletun, P.E. Waaler, H. Skeidsvoll, K. Sommerfelt, Executive functions and seizure-related factors in children with epilepsy in Western Norway, Dev Med Child Neurol [Internet] 48 (6) (2006) 519–525, https://doi.org/10.1017/S0012162206001095.
[14] B. Long, M.M. Spencer-Smith, R. Jacobs, M. Mackay, R. Leventer, C. Barnes, et al., Executive function following child stroke: the impact of lesion location, J. Child. Neurol. [Internet] 26 (3) (2011) 279–287, https://doi.org/10.1177/0883073810380049.
[15] V. Anderson, Attention deficits and the frontal lobes, in: V. Anderson, & P.RJ. Jacobs (Eds.), Executive Functions and the Frontal Lobes: A Lifespan Perspective, Taylor & Francis, New York; London, 2008, pp. 317–344.
[16] K.B. Boone, M.O. Pontón, R.L. Gorsuch, J.J. González, B.L. Miller, Factor analysis of four measures of prefrontal lobe functioning, Arch. Clin. Neuropsychol. [Internet] 13 (7) (1998) 585–595, https://doi.org/10.1016/s0887-6177(97)00074-7.
[17] P.C. Bennett, B. Ong, J. Ponsford, Assessment of executive dysfunction following traumatic brain injury: comparison of the BADS with other clinical neuropsychological measures, J. Int. Neuropsychol. Soc. [Internet] 11 (5) (2005) 606–613, https://doi.org/10.1017/S1355617705050721.
[18] A.F. Mirsky, B.J. Anthony, C.C. Duncan, M.B. Ahearn, S.G. Kellam, Analysis of the elements of attention: a neuropsychological approach, Neuropsychol. Rev. [Internet] 2 (2) (1991) 109–145, https://doi.org/10.1007/bf01109051.
[19] D.A. Pineda, V. Merchan, Executive function in young Colombian adults, Int. J. Neurosci. [Internet] 113 (3) (2003) 397–410, https://doi.org/10.1080/00207450390162164.
[20] N.D. Chiaravalloti, C. Christodoulou, H.A. Demaree, J. DeLuca, Differentiating simple versus complex processing speed: influence on new learning and memory performance, J. Clin. Exp. Neuropsychol. [Internet] 25 (4) (2003) 489–501, https://doi.org/10.1076/jcen.25.4.489.13878.
[21] D.R. Royall, E.C. Lauterbach, J.L. Cummings, A. Reeve, T.A. Rummans, D.I. Kaufer, et al., Executive control function: a review of its promise and challenges for clinical research. A report from the Committee on Research of the American Neuropsychiatric Association, J. Neuropsychiatry Clin. Neurosci. [Internet] 14 (4) (2002 Autumn) 377–405, https://doi.org/10.1176/jnp.14.4.377.
[22] A. Miyake, N.P. Friedman, M.J. Emerson, A.H. Witzki, A. Howerter, T.D. Wager, The unity and diversity of executive functions and their contributions to complex "'frontal lobe"' tasks: a latent variable analysis, Cogn. Psychol. [Internet] 41 (2000) 49–100, https://doi.org/10.1006/cogp.1999.0734.
[23] V.A. Anderson, P. Anderson, E. Northam, R. Jacobs, C. Catroppa, Development of executive functions through late childhood and adolescence in an Australian sample, Dev. Neuropsychol. [Internet] 20 (1) (2001) 385–406, https://doi.org/10.1207/S15326942DN2001_5.
[24] I.S. Baron, Neuropsychological Evaluation of the Child, Oxford University Press, New York, 2004.
[25] J.R. Best, P.H. Miller, A developmental perspective on executive function: development of executive functions, Child. Dev. [Internet] 81 (6) (2010) 1641–1660, https://doi.org/10.1111/j.1467-8624.2010.01499.x.
[26] C.R. De Luca, R.J. Leventer, Developmental trajectories of executive functions across the lifespan, in: V. Anderson, & P.RJ. Jacobs (Eds.), Executive Functions and the Frontal Lobes: A Lifespan Perspective, Taylor & Francis, New York; London, 2008, pp. 23–56.
[27] A. Diamond, Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry, in: Principles of Frontal Lobe Function, Oxford University Press, 2002, pp. 466–503.
[28] A. Diamond, Executive functions, Annu. Rev. Psychol. [Internet] 64 (1) (2013) 135–168, https://doi.org/10.1146/annurev-psych-113011-143750.
[29] P.S. Goldman-Rakic, Development of cortical circuitry and cognitive function, Child. Dev. [Internet] 58 (3) (1987) 601, https://doi.org/10.2307/1130201.
[30] L.M. Grattan, P.J. Eslinger, Frontal lobe damage in children and adults: a comparative review, Dev. Neuropsychol. [Internet] 7 (3) (1991) 283–326, https://doi.org/10.1080/87565649109540496.
[31] H.S. Levin, K.A. Culhane, J. Hartmann, K. Evankovich, A.J. Mattson, H. Harward, et al., Developmental changes in performance on tests of purported frontal lobe functioning, Dev. Neuropsychol. [Internet] 7 (3) (1991) 377–395, https://doi.org/10.1080/87565649109540499.
[32] C.B. Romine, C.R. Reynolds, A model of the development of frontal lobe functioning: findings from a meta-analysis, Appl. Neuropsychol. [Internet] 12 (4) (2005) 190–201, https://doi.org/10.1207/s15324826an1204_2.
[33] R. Testa, P. Bennett, J. Ponsford, Factor analysis of nineteen executive function tests in a healthy adult population, Arch. Clin. Neuropsychol. [Internet] 27 (2) (2012) 213–224, https://doi.org/10.1093/arclin/acr112.
[34] K.G. Jöreskog, Statistical analysis of sets of congeneric tests, Psychometrika [Internet] 36 (2) (1971) 109–133. https://link.springer.com/article/10.1007/BF02291393.
[35] T. Raykov, Estimation of composite reliability for congeneric measures, Appl. Psychol. Meas. [Internet] 21 (2) (1997) 173–184, https://doi.org/10.1177/01466216970212006.
[36] D. Wechsler, Wechsler Abbreviated Scale of Intelligence (WASI), Psychological Corporation, San Antonio, TX, 1999.
[37] R.M. Reitan, The relation of the trail making test to organic brain damage, J. Consult. Psychol. [Internet] 19 (5) (1955) 393–394, https://doi.org/10.1037/h0044509.
[38] R. Brickencamp, Test D2: AufmerksamkeitsBelastungs-Test: Handanweisung, 1981, p. 2.
[39] M. Regard, Cognitive Rigidity and Flexibility: A Neuropsychological Study, 1981. Unpublished doctoral dissertation.
[40] P.M. Corsi, Human Memory and the Medial Temporal Region of the Brain, 1972. Doctoral dissertation.
[41] D. Wechsler, WAIS-III Administration and Scoring Manual, Psychological Corporation, San Antonio, TX, 1997.
[42] E.H. Strauss, E.M.S. Sherman, O. Spreen, A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary, Oxford University Press, New York, NY, 2006.
[43] M. Regard, E. Strauss, P. Knapp, Children's production of verbal and nonverbal fluency tasks, Perceptual. Motor. Skills [Internet] 55 (3) (1982) 839–844, https://doi.org/10.2466/pms.1982.55.3.839.
[44] K.W. Walsh, Neuropsychology: A Clinical Approach, Churchill Livingston, New York, NY, 1978.
[45] D.V. Bishop, G. Aamodt-Leeper, C. Creswell, R. McGurk, D.H. Skuse, Individual differences in cognitive planning on the Tower of Hanoi task: neuropsychological maturity or measurement error? J. Child. Psychol. Psychiatry [Internet] 42 (4) (2001) 551–556, https://doi.org/10.1017/s0021963001007247.
[46] A. Rey, L'examen psychologique dans les cas d'encephalopathie traumatique, Archiv. Psychol. [Internet] 28 (1941) 286–340.
[47] R.K. Heaton, G.J. Chelune, J.L. Talley, G.G. Kay, Wisconsin card sorting test manual: revised and expanded, in: G. Curtis (Ed.), Psychological Assessment Resources, 1993.
[48] I.H. Robertson, T. Ward, V. Ridgeway, I. Nimmo-Smith, The Test of Everyday Attention, Thames Valley Test Company, Bury St. Edmunds, 1994.
[49] B.A. Wilson, N. Alderman, P.W. Burgess, H.C. Emslie, J.J. Evans, The Behavioural Assessment of the Dysexecutive Syndrome, Thames Valley Test Company, Bury St Edmunds, UK, 1996.
[50] J.M. Orr, P.R. Sackett, C.L.Z. Dubois, Outlier detection and treatment in i/o psychology: a survey of researcher beliefs and an empirical illustration, Pers. Psychol. [Internet] 44 (3) (2006) 473–486, https://doi.org/10.1111/j.1744-6570.1991.tb02401.x.
[51] J.F. Hair, W.C. Black, B.J. Babin, R.E. Anderson, J. Hair, R.R. Anderson, et al., Harlow, UK: Pearson Education, in: Multivariate Data Analysis, Prentice-Hall, Upper Saddle River, NJ, 1998.

[52] P.M. Bentler, K.-H. Yuan, Structural equation modeling with small samples: test statistics, Multivariate Behav. Res. 34 (2) (1999 Apr) 181–197, https://doi.org/10.1207/S15327906Mb340203.

[53] R.C. MacCallum, M.W. Browne, H.M. Sugawara, Power analysis and determination of sample size for covariance structure modeling, Psychol. Methods 1 (2) (1996) 130–149, https://doi.org/10.1037/1082-989X.1.2.130.

[54] D.L. Jackson, Revisiting sample size and number of parameter estimates: some support for the N:q hypothesis. Structural equation modeling, A Multidiscip. J. 10 (1) (2003 Jan) 128–141, https://doi.org/10.1207/S15328007SEM1001_6.

[55] T. Raykov, K.F. Widaman, Issues in applied structural equation modeling research, Struct. Equ. Model.: A Multidiscip. J. 2 (4) (1995 Jan) 289–318, https://doi.org/10.1080/10705519509540017.

[56] P.M. Bentler, C.-P. Chou, Practical issues in structural modeling, Socio. Methods Res. 16 (1) (1987 Aug) 78–117, https://doi.org/10.1177/0049124187016001004.

[57] R.P. Bagozzi, Y. Yi, Specification, evaluation, and interpretation of structural equation models, J. Acad. Mark. Sci. [Internet] 40 (1) (2012) 8–34, https://doi.org/10.1007/s11747-011-0278-x.

[58] G.R. Hancock, R.O. Mueller, Rethinking construct reliability within latent variable systems, in: R. Cudeck, S. Toit, D. Sörbom (Eds.), Structural Equation Modeling: Present and Future A Festschrift in Honor of Karl Jöreskog, 2001, pp. 195–216.

[59] W. Widhiarso, H. Ravand, Estimating reliability coefficient for multidimensional measures: a pedagogical illustration [cited 2023 Feb 14], Rev. Psychol. [Internet] 21 (2) (2014) 111–121, https://hrcak.srce.hr/147109.

[60] K.G. Jöreskog, D. Sörbom, LISREL 7 User's Reference Guide, Scientific Software Inc, Chicago, 1989.

[61] D.E. Broadbent, Perceptions and Communication, Pergammon, London, 1958.

[62] A.M. Treisman, Selective attention in man, Br. Med. Bull. [Internet] 20 (1) (1964) 12–16, https://doi.org/10.1093/oxfordjournals.bmb.a070274.

[63] D.G. Mackay, Aspects of the theory of comprehension, memory and attention, Q. J. Exp. Psychol. [Internet] 25 (1) (1973) 22–40, https://doi.org/10.1080/14640747308400320.

[64] D. Wechsler, Technical Manual for the Wechsler Adult Intelligence Test, Psychological Corporation, San Antonio, TX, 1997.

[65] V. Anderson, Assessing executive functions in children: biological, psychological, and developmental considerations, Neuropsychol. Rehabil. [Internet 8 (3) (1998) 319–349, https://doi.org/10.1080/713755568.

[66] M. Sarter, B. Givens, J.P. Bruno, The cognitive neuroscience of sustained attention: where top-down meets bottom-up, Brain Res. Rev. [Internet] 35 (2) (2001) 146–160, https://doi.org/10.1016/s0165-0173(01)00044-3.

[67] M.J. Kane, D.Z. Hambrick, A.R.A. Conway, Working memory capacity and fluid intelligence are strongly related constructs: comment on Ackerman, Beier, and Boyle, Psychol. Bull. [Internet] 131 (1) (2005) 66–71, https://doi.org/10.1037/0033-2909.131.1.66.

[68] N. Garon, S.E. Bryson, I.M. Smith, Executive function in preschoolers: a review using an integrative framework, Psychol. Bull. [Internet] 134 (1) (2008) 31–60, https://doi.org/10.1037/0033-2909.134.1.31.

[69] J.C. Anderson, D.W. Gerbing, Structural equation modeling in practice: a review and recommended two-step approach, Psychol. Bull. [Internet] 103 (3) (1988) 411–423, https://doi.org/10.1037/0033-2909.103.3.411.

[70] D.R. Dajani, L.Q. Uddin, Demystifying cognitive flexibility: implications for clinical and developmental neuroscience, Trends Neurosci. [Internet] 38 (9) (2015) 571–578, https://doi.org/10.1016/j.tins.2015.07.003.

[71] B.E. Goldstein, Cognitive Psychology Connecting Mind, Research, and Everyday Experience, Thomson Wadsworth, Belmont, CA, 2008.

[72] R.W. Engle, M.J. Kane, S.W. Tuholski, Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex, in: A. Miyake, P. Shah (Eds.), Models of Working Memory, Cambridge University Press, Cambridge, 1999, pp. 102–134.

[73] R.W. Engle, S.W. Tuholski, J.E. Laughlin, A.R.A. Conway, Working memory, short-term memory, and general fluid intelligence: a latent variable approach, J. Exp. Psychol.: General [Internet] 128 (3) (1999) 309–331, https://doi.org/10.1037/0096-3445.128.3.309.

[74] A.R.A. Conway, N. Cowan, M.F. Bunting, D.J. Therriault, S.R.B. Minkoff, A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence, Intelligence 30 (2) (2002 Mar) 163–183, https://doi.org/10.1016/S0160-2896(01)00096-4.

[75] R.V. Kail, C.A. Miller, Developmental change in processing speed: domain specificity and stability during childhood and adolescence, J. Cogn. Dev. [Internet] 7 (1) (2006) 119–137, https://doi.org/10.1207/s15327647jcd0701_6.

[76] L. Pires, O. Moura, C. Guerrini, I. Buekenhout, M.R. Simões, J. Leitão, Confirmatory factor analysis of neurocognitive measures in healthy young adults: the relation of executive functions with other neurocognitive functions, Arch. Clin. Neuropsychol. [Internet] 34 (3) (2019) 350–365, https://doi.org/10.1093/arclin/acy040.

[77] N.J. Cepeda, K.A. Blackwell, Y. Munakata, Speed isn't everything: complex processing speed measures mask individual differences and developmental changes in executive control, Dev. Sci. 16 (2) (2013 Feb 23) 269–286, https://doi.org/10.1111/desc.12024.

[78] A.F. Fry, S. Hale, Processing speed, working memory, and fluid intelligence: evidence for a developmental cascade, Psychol. Sci. 7 (4) (1996 Jul) 237–241, https://doi.org/10.1111/j.1467-9280.1996.tb00366.x.

[79] S. Hale, J. Jansen, Global processing-time coefficients characterize individual and group differences in cognitive speed, Psychol. Sci. [Internet] 5 (6) (1994) 384–389, https://doi.org/10.1111/j.1467-9280.1994.tb00290.x.

[80] B. Peter, M. Matsushita, W.H. Raskind, Global processing speed in children with low reading ability and in children and adults with typical reading ability: Exploratory factor analytic models, J. Speech, Lang., Hear. Res. [Internet] 54 (3) (2011) 885–899, https://doi.org/10.1044/1092-4388(2010/10-0135.

[81] A.F. Fry, S. Hale, Relationships among processing speed, working memory, and fluid intelligence in children, Biol. Psychol. [Internet] 54 (1–3) (2000) 1–34. https://www.sciencedirect.com/science/article/pii/S030105110000051X.

[82] T. McAuley, D.A. White, A latent variables examination of processing speed, response inhibition, and working memory during typical development, J. Exp. Child Psychol. [Internet] 108 (3) (2011) 453–468, https://doi.org/10.1016/j.jecp.2010.08.009.

[83] R.H. Pietrzak, P. Maruff, L.C. Mayes, S.A. Roman, J.A. Sosa, P.J. Snyder, An examination of the construct validity and factor structure of the Groton Maze Learning Test, a new measure of spatial working memory, learning efficiency, and error monitoring, Arch. Clin. Neuropsychol. [Internet] 23 (4) (2008) 433–445, https://doi.org/10.1016/j.acn.2008.03.002.

[84] R. Bray, S. McDonald, Austin Maze, Australian Society for the Study of Brain Impairment, Sydney, NSW, 2010.

[85] K.W. Walsh, A. Darby, Neurosychology: A Clinical Approach, Churchill Livingston, 1994.

[86] S.J. Handley, A. Capon, C. Copp, C. Harper, Conditional reasoning and the Tower of Hanoi: the role of spatial and verbal working memory, Br. J. Psychol. [Internet] 93 (4) (2002) 501–518, https://doi.org/10.1348/000712602761381376.

[87] V. Goel, S.D. Pullara, J. Grafman, A computational model of frontal lobe dysfunction: working memory and the Tower of Hanoi task, Cogn. Sci. [Internet] 25 (2) (2001) 287–313, https://doi.org/10.1016/S0364-0213(01)00037-4.