

Assessment of executive functions: Review of instruments and identification of critical issues

Raymond C.K. Chan^{a,*}, David Shum^b, Timothea Touloupoulou^c, Eric Y.H. Chen^d

^a *Neuropsychology and Applied Cognitive Neuroscience Laboratory and Key Laboratory of Mental Health, Institute of Psychology, Chinese Academy of Sciences, China*

^b *School of Psychology and Applied Cognitive Neuroscience Research Centre, Griffith University, Brisbane, Australia*

^c *Division of Psychological Medicine, Section of General Psychiatry, Institute of Psychiatry, London, UK*

^d *Department of Psychiatry, The University of Hong Kong, Hong Kong, China*

Accepted 28 August 2007

Abstract

“Executive functions” is an umbrella term for functions such as planning, working memory, inhibition, mental flexibility, as well as the initiation and monitoring of action. The impairment of executive functions in various clinical groups is a topic of much debate, as are recent attempts to formulate the corresponding intervention and rehabilitation regimes of these dysfunctions. This article reviewed current theories of executive functions and their associated assessment instruments. In addition, it identified issues that are imperative for more accurate, sensitive, and specific assessment of various components of this construct. It is concluded that more research is needed to fractionate the executive system by assessing a wide range of functions and to verify their neuroanatomical correlates. Recently developed measurement models and technology may also facilitate a more ecologically and ethologically valid assessment for the specific needs of different individuals.

© 2007 National Academy of Neuropsychology. Published by Elsevier Ltd. All rights reserved.

Keywords: Executive function; Assessment; Rehabilitation

1. Introduction

The term “executive functions” is an umbrella term comprising a wide range of cognitive processes and behavioral competencies which include verbal reasoning, problem-solving, planning, sequencing, the ability to sustain attention, resistance to interference, utilization of feedback, multitasking, cognitive flexibility, and the ability to deal with novelty (Burgess, Veitch, de lacy Costello, & Shallice, 2000; Damasio, 1995; Grafman & Litvan, 1999; Shallice, 1988; Stuss & Benson, 1986; Stuss, Shallice, Alexander, & Picton, 1995). These functions have been called the “cold” component of executive functions because their corresponding cognitive processes tend not to involve much emotional arousal and are relatively “mechanistic” or “logically” based (Grafman & Litvan, 1999). On the other hand, those executive functions involving more “emotional”, “belief” or “desires” such as the experience of reward and punishment, regulation of one’s own social behavior, and decision-making involving emotional and personal interpretation, are regarded as “hot” components (Bechara, Damasio, Damasio, & Lee, 1999; Bechara, Tranel, Damasio, & Damasio, 1996; Damasio, 1995;

* Corresponding author at: Institute of Psychology, Chinese Academy of Sciences, 4A Datun Road, Beijing 100101, China. Tel.: +86 10 64852558; fax: +86 10 64872070.

E-mail addresses: rckchan@psych.ac.cn, rckchan2003@yahoo.com.hk (R.C.K. Chan).

Grafman & Litvan, 1999; Rolls, 1995). Studies have shown that impairments in either the “cold” or “hot” component of executive functions may have devastating effects on people’s everyday life activities, including the ability to work and attend school, function independently at home, or develop and maintain appropriate social relations (Goel, Grafman, Tajik, Gana, & Danto, 1997; Grafman et al., 1996; Green, 1996; Green, Kern, Braff, & Mintz, 2000).

The neuropsychological study of dysexecutive syndrome and its corresponding rehabilitation, however, face inherent difficulties. One of them is the accurate and valid assessment of executive functions. Executive functions are thought to comprise a series of abilities to achieve a goal (Damasio, 1995; Shallice, 1988; Stuss et al., 2005; Stuss & Benson, 1986). Therefore, failure on executive function tests may be due to many reasons, as damage to any component process is difficult to be fully ruled out following the onset of brain lesions or psychopathologies.

A further difficulty lies in the apparent fractionation of the dysexecutive syndrome. For example, a patient’s performance on one executive function test may have little or no predictive value for how he or she may perform on another test, let alone in a complex real world situation (Burgess, 1997; Burgess, Alderman, Evans, Emslie, & Wilson, 1998). In addressing this problem, there is an increasing emphasis on incorporating more complex, multifaceted and life-like challenges within performance measures, in other words, tasks that tap a number of executive domains at the same time (Schwartz, Reed, Montgomery, Plamer, & Mayer, 1991; Shallice & Burgess, 1991; Wilson, Alderman, Burgess, Emslie, & Evans, 1986).

Successful attempts in fractionating the executive system will depend to a considerable extent on the ability to develop more specific models of executive functions (Fan, McCandliss, Sommer, Raz, & Posner, 2002; Posner & Raichle, 1994; Shallice, 1988; Shallice & Burgess, 1991). The current conventional executive function tests used by clinicians and neuropsychologists tend to be crude and underspecified in terms of the cognitive processes that they engage (Burgess, 1997). These tests may not be sensitive enough to detect executive dysfunction in different clinical groups. The most recent and significant advancement in the past decade has been the attempt to isolate the specific component processes of prefrontal functions (Burgess, 1997; Chan, Chen, Cheung, Chen, & Cheung, 2006; Chan, Chen, & Law, 2006; Lin, Chan, Zheng, Yang, & Wang, 2007). Research on the role of the prefrontal lobes on human attention has given us some promising data and an example about how such advances can be achieved.

By reviewing tests that are commonly used in assessing executive functions and by evaluating the empirical evidence, this paper aimed to serve as a general review of assessment of executive functions for clinical practitioners and researchers. Specifically, this paper aimed to

- (1) identify key issues relating to the models and assessment of executive functions (e.g., sensitivity to prefrontal lobe damage, ecological validity, lack of correlation between scores and outcome measures),
- (2) discuss ways to improve the development of tests that examine executive functions (e.g., developing tasks based on theory, developing tasks that are less complex, advanced statistical procedures for item selection, virtual reality), and
- (3) discuss implications for future direction and development.

2. Models of executive functions

The models reviewed in this article are not exhaustive and are limited to the adult populations. We selected those models that have wider implications for test development and clinical application for intervention. These models will be described in chronological order, immediately followed by a description of related tests development and their clinical utility for assessing specific executive function components. The first five concerns the “cold” component of executive function, whereas the last one relates to the “hot” component. For the assessment of executive functions in children, please refer to related sources (e.g., Bush, Luu, & Posner, 2000; Das, Naglieri, & Kirby, 1994; Duncan & Owen, 2000; Kane & Engle, 2002; Royall et al., 2002; Rushworth & Owen, 1998).

2.1. Luria’s theory

According to Luria (1966, 1973), the human brain comprises three basic functional units that are interactively linked. The first unit is located mainly in the brain stem and is responsible for regulating and maintaining arousal of the cortex. The second unit is responsible for encoding, processing, and storage of information and encompasses the temporal, parietal, and occipital lobes. The third functional unit is located in the anterior region of the brain (viz., the frontal lobes)

and its functions include programming, regulating, and verifying human behavior. Within the third unit, the prefrontal cortex is considered by Luria as a superstructure that regulates or control mental activity and behavior. Damage to the frontal lobes, and in particular the prefrontal cortex, is expected to disrupt complex behavioral programmes and a person's ability to verify or regulate behavioral outcomes. Consequently, it can lead to the replacement of these complex programmes by more basic behavior or stereotypical behavior that is either illogical, irrelevant, or inappropriate.

2.1.1. Executive function tests developed from Luria's theory

The motor task developed by Luria (1973) to assess frontal lobe lesions are often used in clinical settings as screening tests. Simple finger opposition, Fist-Edge-Palm Test, and the Reciprocal Motor Programme Test are typical examples. These easy-to-administer tests are sensitive to brain dysfunction such as traumatic brain injury (Yarnell & Rossi, 1988) and schizophrenia (Chan & Chen, 2004; Chan, Chen, Cheung, & Cheung, 2004). In the Reciprocal Motor Programme Test, testees are instructed to tap their hands once when they hear two tapping sound and to tap their hands twice when they hear one tapping sound. In doing so, they are required to disinhibit their motor action and to react in an opposing way. For the Fist-Edge-Palm Test, testees are requested to place their hands in each of the postures, that is, fist, edge, and palm, in an alternate and success way as quickly as possible. These tests have been incorporated into recent batteries of the so-called "neurological soft signs" (Chen et al., 1995; Heinrichs & Buchanan, 1988) and standardized batteries for frontal-executive functions (Fox & Fox, 2001; Kaufman & Kaufman, 1983; Korkman, 1999).

The Reciprocal Motor Programme Test has been shown to have similar sensitivity to the Wisconsin Card Sorting Test in patients with frontal lobe lesions (Mirsky, 2003). Most recently, Chan and Chen (2004) also demonstrated that these neurological soft signs (e.g., motor coordination, sensory integration and disinhibition) were strongly correlated with executive functions in patients with schizophrenia. In particular, updating performance may be more predictive of motor coordination (e.g., Fist-Edge-Palm or Reciprocal Motor Programme Test), while general intellectual functioning was more predictive of sensory integration (e.g., graphesia or asteregnosis) and disinhibition (e.g., mirror movement or saccade blink inhibition). However, data from imaging studies (e.g., Chan, Rao, Chen, Ye, & Zhang, 2006; Umetsu et al., 2002) in healthy subjects indicate that this classical Lurian test of "Fist-Edge-Palm" may activate regions other than prefrontal cortex, thus questioning this task as a pure frontal or executive function task. Nevertheless, the nature and characteristics of this task and other similar neurological soft signs (e.g., shorter testing time, the language free component and its portable nature) suggest that these tests can serve as an alternate screening test of executive dysfunctions in clinical samples, particularly for the assessment of motor initiation, sequencing and inhibition components.

2.2. Supervisory attentional system (SAS) model and its variants

Norman and Shallice (1986) extended Luria's idea of frontal lobe functioning and came up with their supervisory attentional system (SAS) model. According to this model, the programming, regulating, and verifying of human actions and thoughts involve two systems, namely, contention scheduling and supervisory attentional. The former system is responsible for routine and overlearned behaviors or tasks and allows us to prioritize the order of these behaviors and tasks (e.g., making a coffee while talking on the phone). The latter system, on the other hand, is responsible for regulating non-routine and novel tasks. In particular, there are five types of situations where routine, automatic activation of behavior would not be sufficient for optimal performance (Norman & Shallice, 1986). These include situations: (1) that involve planning or decision-making; (2) that involve error correction or troubleshooting; (3) where responses are not well-learned or contain novel sequences of actions; (4) where danger is anticipated; and (5) which require the overcoming of a strong habitual response or resisting temptation.

More recently, Burgess and colleagues (Burgess, 2000; Burgess et al., 2000) extended the SAS concept to multi-tasking performance in everyday life. According to them, the eight features of multitasking behavior include: (1) many discrete tasks for an individual to complete; (2) interleaving period for an effective performance; (3) engagement in only one task at a particular time slot; (4) unforeseen interruptions and unexpected outcomes; (5) delayed intentions for the individual to return to a task which is already running and is not signaled directly by the situation; (6) tasks demanding different task characteristics; (7) self-determining targets with which the individual decides for him/herself what constitutes adequate performance; and (8) no immediate feedback on how well the individual performs. According to Burgess and his colleagues, most laboratory-based tasks do not include all of these features in assessing crucial components of multi-tasking performance in clinical cases.

2.2.1. Executive function tests developed from the SAS model and its derivatives

Evidence supporting the framework of the SAS comes from studies of patients with frontal lobe lesions associated with loss of supervisory control (Chan, 2002; Chan, Hoosain, Lee, Fan, & Fong, 2003; Shallice, 1988), attention lapses and dysexecutive behavior in healthy subjects (Chan, 2001; Reason, 1984), and reaction time costs when healthy subjects intentionally switch attention between alternative tasks (Allport, Styles, & Hsieh, 1994; Rogers & Monsell, 1995).

Several tests have been devised based on the SAS model and were specifically designed for capturing different components of SAS or executive functions (viz., planning, strategy allocation and monitoring, inhibition or suppression of semantic and action responses). These are the Six Elements Test (Shallice & Burgess, 1991), Hayling Sentence Completion Test (Burgess & Shallice, 1996a) and Brixton-Spatial Anticipation Test (Burgess & Shallice, 1996b), Sustained Attention to Response Task (Robertson, Manly, Andrade, Baddeley, & Yiend, 1997).

The Six Elements Test (SET), first described by Shallice and Burgess (1991), and subsequently incorporated into the Behavioral Assessment of Dysexecutive Syndrome (BADS) (Wilson et al., 1996), has been shown to exhibit very impressive validity and sensitivity to patients with frontal lesions and other patients with neurological and psychiatric disorders (Burgess et al., 1998; Chan & Manly, 2002; Chan et al., 2003; Chan, Chen, et al., 2004; Chan, Chen, Cheung, et al., 2006; Chan, Chen, & Law, 2006; Evans, Chua, McKenna, & Wilson, 1997). The SET consists of three types of tests (viz., simple arithmetic, written picture naming, and dictation), each of which has two sub-tasks (thus constituting a total of six sub-tasks). Testees are required to attempt at least part of each of the six sub-tasks within 10 min, following the rule that they are not allowed to switch directly from a sub-task of one type to its counterpart of the same type. To achieve good performance on this test, testees are thus required to mobilize the most appropriate schemata across the different sub-tasks consistently and optimally. Others (e.g., Levine et al., 2000; Manly, Hawkins, Evans, Woldt, & Robertson, 2002) have modified the SET and developed the Strategic Self-Regulation Test and Hotel Test to assess strategy allocation aspect of executive function in patients with traumatic brain injury and have demonstrated these SET derivatives are sensitive in detecting deficits in strategy allocation in this clinical group. The SET and its derivatives, therefore, assess the ability to maintain an optimal level of performance while sticking to the given rule across different sub-tasks.

The Hayling Sentence Completion Test (Burgess & Shallice, 1996a) primarily aims to detect difficulties in suppressing pre-potent responses and it consists of two parts. In the first part, a testee is required to complete the end of sentences with a pre-potent response to make meaningful connection. For instance, responding with the word “ship” to the sentence “the captain went down with the sinking —”. In the second part, on the contrary, the testee is required to inhibit the pre-potent response by providing irrelevant words to complete the given sentences. For instance, responding with the word “cow” to the sentence “the captain went down with the sinking —”. Therefore, the first part of the test is supposed to capture initiation whereas the second part is supposed to measure suppression or inhibition. Preliminary psychometric properties of this test in terms of test-retest reliability ($r=0.72-0.93$) and internal consistency ($\alpha=0.62-0.76$) have been demonstrated in a group of patients with a wide range of neurological disorders (Burgess & Shallice, 1997). Crawford and Henry (2005) re-analyzed the data collected by Burgess and Shallice (1997) and found the effect size for error under suppression/inhibition (0.37) is markedly larger than that for basic initiation (0.23) in the anterior lesioned cases than the posterior lesioned cases. This supports the specificity of the Hayling Sentence Completion Test for anterior lesions.

The Brixton Spatial Anticipation Test (Burgess & Shallice, 1996b) is very similar to the Wisconsin Card Sorting Test (WCST). It requires an individual to discover the rules underlying the placement of blue circles among a grid of unfilled circles. After a given pattern has been formulated the placement rule changes. However, this test is relatively less time-consuming and less stressful than the WCST and therefore more user-friendly and feasible for clinical and research purposes. Crawford and Henry (2005) found a large effect size ($r=0.5$) and a moderate effect size ($r=0.34$) between healthy controls and patients with frontal and posterior lesions respectively. Empirical evidence on the internal consistency of this test has also been shown to be modest ($\alpha=0.62$) among the healthy sample.

The Sustained Attention to Response Task (SART, Robertson et al., 1997), is a computerized test of sustained attention. In it, regularly presented non-target visual responses should be withheld for a rare target digit of “3”. Owing to the regular, rhythmic pacing and the rarity of targets, the task encourages a strategy of fast, anticipatory, automatic responding. Within the SAS framework, while the non-target response is frequently exogenously activated and elicited by the task, the activation level of the target response must be endogenously maintained close to threshold if it is to be completed successfully.

Recently, Chan, Chen, et al. (2004) have successfully demonstrated the construct validity of the Hayling Sentence Completion Test, Six Elements Test, and the SART in a group of patients with schizophrenia using factor analysis. The factors found were: (1) a “semantic inhibition” factor that comprised items inhibiting verbal and semantic responses; (2) an “action/attention inhibition” factor that comprised items on error commission and rule-breaking; and (3) an “output generation” factor that comprised items on initiation and generation of response. Subsequent studies (Chan, Chen, Cheung, et al., 2006; Chan, Chen, & Law, 2006) have found that these factors of executive functions were sensitive to discriminate patients with schizophrenia from healthy controls. Moreover, significant but modest relationships were also demonstrated between these components and negative symptoms in schizophrenia.

Finally, executive dysfunctions may be observed in routine and over-learned daily life tasks, especially when one is engaging in multiple tasks simultaneously. Wilson et al. (1996) developed a subjective rating scale, namely the Dysexecutive Questionnaires (DEX), to examine these everyday dysexecutive behaviors. Impressive psychometric properties and clinical utility of the DEX have been demonstrated in both clinical (Burgess et al., 1998; Chan & Manly, 2002) and non-clinical samples (Amieva, Phillips, & Della Sala, 2003; Chan, 2001).

2.3. *Stuss and Benson’s tripartite model*

According to Stuss and Benson (1986), there are three systems that interact together to monitor an individual’s attention and executive functions. They are the anterior reticular activating system (ARAS), the diffuse thalamic projection system, and the fronto-thalamic gating system. While the first two systems are responsible for maintaining an individual’s alertness, the third system is involved in executive attentional control. Specifically, the ARAS maintains the general arousal level of an individual, that is, tonic changes of alertness and damage of this system will result in loss of consciousness. The diffuse thalamic projection system is responsible for maintaining an individual’s alertness to external stimuli over a short period of time, that is, phasic changes of alertness and damage to this system will result in distraction by external stimuli. Finally, the fronto-thalamic gating system is responsible for higher-level cortical functioning such as planning, stimuli and response selection, and monitoring of daily performance. Damage to this system will lead to symptoms such as inattention, insight impairment, and goal-neglect behavior that is similar to those dysfunctions of the SAS described by Shallice (1988).

Stuss et al. (1995) expanded upon how they see the relationship between the schema and the SAS might operate. They describe a schema as a network of connected neurons that can be activated by sensory input, by other schemata, or by the executive control system. The strength of this theory is that the authors identify different executive attentional components on a neural basis, including sustaining (right frontal), concentrating (cingulate area), sharing (cingulate and orbitofrontal areas), suppressing (dorsolateral prefrontal cortex), switching (dorsolateral prefrontal and medial frontal areas), preparing (dorsolateral prefrontal cortex), and goal setting (left dorsolateral prefrontal cortex).

2.3.1. *Executive function tests related to Stuss’s tripartite theory*

Stuss et al. (1995) identified several other tasks that were also supposedly developed to assess executive attentions based on the SAS. These include: (1) tasks that involve conflict, such as various versions of the Stroop, (2) tasks that demand mental switching between subtasks or categories such as WCST, Trail Making Test and verbal fluency, and (3) tests that involve higher cortical control to sustain and monitor attention over a long period of time such as conventional tests of sustained attention. Recently, tasks involving cognitive conflict have been shown to activate several brain areas thought to be involved in the executive attention network, but in functionally distinct ways (Botvnick, Nystrom, Fissell, Carter, & Cohen, 1999; Bush et al., 2000; Casey et al., 2000; Fan et al., 2002; Posner & Raichle, 1994). These tasks may therefore provide a means of fractionating the functional contributions of areas within the SAS.

Stuss and his colleagues (Stuss et al., 2005) have subsequently constructed a battery, known as the ROTman-Baycrest Battery to Investigate Attention (ROBBIA) that comprehensively assess the three categories of tasks described above. These tests consist of: (1) simple reaction time task to detect and respond to one stimulus that occurs over a prolonged period of time at a relatively infrequent rate; (2) choice reaction time task, similar to the simple reaction time task, but with the addition condition of making a second response to non-target stimuli; (3) prepare reaction time task, similar to choice reaction time task, but with the addition of a preparatory signal presented at variable time lengths preceding the stimuli; (4) concentrate task to respond to a series of choice reaction time task where responses are made to stimuli occurring at a rapid rate; (5) count task to count the number of stimuli presented at different rates; (6) divide task to respond to two separate and unrelated tasks that are occurring at the same time; (7) tap task to make a motor tapping

response at a fixed rate both with and without an external cue; (8) switch task to switch between two different tasks within the same block of stimuli; (9) NoGo task to suppress a response of a particular stimulus or a class of stimuli; (10) suppress task to suppress a response to a non-target stimulus that shares characteristics with a target stimulus; and (11) set task to establish a response mode or task setting when response requirement changes from one block of stimuli to another. This battery has been shown to have impressive preliminary sensitivity to a group of patients with focal frontal lesions (Stuss et al., 2005). In particular, energization of the schemata has been found to be sensitive to right medial superior lesions, and monitoring the occurrence of stimuli over time, in anticipation of responding more quickly to upcoming stimuli, has been found to be sensitive to right lateral lesions.

2.4. Duncan's goal-neglect theory

Duncan and colleagues (Duncan, 1986, 1995; Duncan & Owen, 2000; Duncan et al., 2000) emphasize the crucial role of a set of goals or subgoals in governing the optimal function of human behavior. In his goal-neglect theory, Duncan proposes that human behavior is goal-oriented or goal-directed and it is controlled by a list of goals or subgoals. These goals are formulated, stored and checked in mind by an individual in order to behave optimally and properly in response to environmental or internal demands. One of the main functions of the goals is to impose a structure on behavior by controlling the activation or inhibition of behavior that facilitate or prevent task completion. The involvement of the frontal lobe in goal-oriented or goal-directed behavior is illustrated by the fact that patients with damage to this area are usually disorganized and fail to achieve intended goals or what Duncan refers to as "goal-neglect". Although these patients are apparently able to remember the intended goals, they tend to lose sight of these goals and their actions may become random or stuck on one or more subgoals.

2.4.1. Executive function tests developed from the goal-neglect theory

Although no formal tests have been developed from the goal-neglect theory, Robertson (1996) put together a training package of Goal-Management Training (GMT) based on this theory. The emphasis of GMT is on helping patients to be more effective at encoding the goals and sub-goals and learning a mental checking routine that enables them to maintain task focus. The GMT trains patients in six stages which characterize the most crucial elements for maintaining goals in mind. The six stages are: (1) STOP for letting patients asking themselves what they are doing at the moment; (2) DEFINE for defining the main task out of a set of irrelevant or less prioritized tasks; (3) LIST for listing the steps of accomplishing the chosen or the most prioritized task; (4) LEARN for asking themselves whether they know the steps for achieving the chosen or prioritized task; (5) DO IT by executing the task; and (6) CHECK for monitoring the ongoing task. Therefore, the GMT targets parameters relevant to setting and selecting relevant goals and subgoals, defining the target task, keeping the steps as the task is being achieved, and monitoring task outcome. Levine et al. (2000) have demonstrated that the clinical utility and training efficacy of the GMT among a group of frontal lobe lesioned patients. However, they commented that the success of GMT depended on a number of factors, including the self-awareness and motivation to complete the training programme.

2.5. Goldman-Rakic's working memory model

Unlike the models and theories reviewed so far, Goldman-Rakic's (1992) working memory model was mainly based on animal studies. In her model, Goldman-Rakic argued that while the whole prefrontal cortex is responsible for working memory, it is divided into multiple sub-regions which are responsible for different types of working memory (e.g., spatial, featural, semantic, and mathematical knowledge). The prefrontal cortex fulfills these functions via two reciprocal pathways (viz., inhibition and excitation) that connect to the posterior brain areas. For instance, the specific prefrontal region responsible for the features of objects is directly connected to the posterior brain areas that specialize in processing the physical features of objects, excites useful information and inhibits unnecessary information, and provides up to date knowledge to the individual to function properly over time. These inhibitory or excitatory commands may be issued via neurotransmitters such as the catecholamines, especially dopamine. Working memory impairment as measured by performance on the delayed-matching task is commonly observed in brain-damaged monkeys which have reduced levels of these chemicals. Interestingly, when the level is restored to a normal level, the impairment disappears.

Because of its functional specificity, Goldman-Rakic's (1992) working memory model is not a widely adopted one (as compared to the models or theories reviewed so far). Nevertheless, the model does provide testable explanations of the role of the prefrontal cortex in working memory and executive functioning by linking working memory impairment with the dopaminergic system.

2.5.1. Executive function tests developed from the working memory model

Unlike the other theories or models reviewed, the working memory model proposed by Goldman-Rakic does not come with its own set of tests. However, there are some tests that may be particularly useful to test this model. For example, the delayed-matching task that was originally applied to animal work has now been incorporated into a computerized assessment package known as the Cambridge Neuropsychological Test Automated Battery (CANTAB). Moreover, there have been some tests developed primarily from the construct of working memory, including those proposed by Baddeley (1986) and Goldman-Rakic (1992), for experimental and clinical uses. For instance, the Letter-Number Span Test (Chan et al., in press; Gold, Carpenter, Randolph, Goldberg, & Weinberger, 1997) and the N-back Test (Callicott et al., 1998) are mostly based on this model to capture semantic and visual aspect of working memory component of the executive functions. The Letter-Number Span Test involves the auditory presentation of a mixed series of alternating numbers and letters. The testee is required to respond by first repeating the numbers in order from the smallest to the largest, followed by repeating the letters in alphabetical order. For the N-back Test, testees are instructed to recall the visual stimulus seen after "N-position" previously presented in the computer screen.

2.6. Damasio's somatic marker hypothesis

Damasio's model (Damasio, 1995; Damasio, Grabowski, Frank, Galaburda, & Damasio, 1994) emphasizes the role of the frontal lobe in emotion and social behavior, particularly in decision-making. Unlike the models and theories reviewed so far, this model tackles the "hot" component of executive functions and its impact upon the "cold" component in everyday decision-making and interpersonal relationship. Damasio (1995) put forward a somatic marker hypothesis to account for the common impairments (viz., dramatic personality change, emotional and interpersonal problems) seen after damage to the ventromedial frontal cortex in patient such as Phineas Gage. According to Damasio, emotion is mediated by the prefrontal regions, via the complicated links between the cortical and subcortical links. In particular, the cortical link involves the ventromedial cortex and the subcortical links involve the mediodorsal nucleus of the thalamus, the amygdala and hypothalamus. Patients with damage to the ventromedial frontal cortex are unable to "mark" or link inappropriate behaviors with an emotion-related somatic signal even though they may be able to understand the implications of such behaviors. Consequently, these patients will show difficulty in regulating their behaviors because they could not make use of emotion-related somatic markers.

2.6.1. Executive function test developed from the somatic marker's hypothesis

Damasio's team (Bechara, Damasio, Damasio, & Anderson, 1994; Bechara, Damasio, Tranel, & Damasio, 1997; Bechara, Tranel, & Damasio, 2000; Bechara et al., 1996; Tranel, Bechara, & Damasio, 2000) developed the Iowa gambling task to test the somatic marker's hypothesis. In this task, individuals have to choose between decks of cards which yield high immediate gain but larger future loss, and decks which yield lower immediate gain but a smaller future loss. The goal in this task is to optimize profit on a loan of play money. Individuals are asked to sit in front of four decks of cards (A, B, C, and D) equal in appearance and size and are given a \$2000 loan of play money. They are told that the game requires a series of card selections, one card at a time, from any of the four decks, until they are told to stop. However, the individuals' decision to choose from one deck versus another is largely influenced by various schedules of immediate reward and future punishments, which have been pre-programmed and are not known to the individuals. In short, Decks A and B are "disadvantageous" because they cost the most in the long run, whereas decks C and D are "advantageous" because they result in an overall gain in the long run.

The Iowa gambling task has been shown to have impressive sensitivity in patients suffering from ventromedial prefrontal cortex lesions. In an initial study done by Bechara et al. (1994), patients with ventromedial prefrontal lesions performed significantly worse than healthy and other brain-damaged controls on this task. The patients with ventromedial prefrontal lesions failed to demonstrate a shift in their selection towards the advantageous decks throughout the trial blocks of the gambling task, while the healthy and other brain-damaged controls gradually shifted their selections towards the advantageous decks in the latter trial blocks. Convergent evidence from behavioral data and psychophysi-

ecological data were obtained from subsequent studies that used different permutations of the gambling task and larger groups of participants (Bechara et al., 1996, 2000).

2.7. A summary of existing tests of executive functions

In recent years, most executive functions tests were developed based on theories or models and most were shown to be sensitive to the effects of injury to the prefrontal cortex (e.g., Wilson et al., 1996; Burgess & Shallice, 1996a, 1996b; Robertson et al., 1997; Gold et al., 1997). Table 1 summarizes the range of theory-based tests of executive functions that are usually used for clinical assessment and experimental study in different clinical groups. The specific components of executive functions are also highlighted in this table according to the three levels of assessment at which they are devised to obtain the necessary functional ability information. For the impairment level, tests may include the Cambridge Neurological Inventory (Chen et al., 1995), Design Fluency Test (Jones-Gotman & Milner, 1977), and various types of Tower tests (Humes, Welsh, Retzlaff, & Cookson, 1997; Newell & Simon, 1972; Shallice, 1982). For the disability level, tests may include more laboratory-based tests with relatively constraint environments such as Greenwich Test (Burgess et al., 2000) and Hotel Test (Manly et al., 2002). For the handicap and participation level, tests administration may extend to a more open and naturalistic environment for assessment such as the Naturalistic Action Test (Schwartz, Segal, Veramonti, Ferraro, & Buxbaum, 2002), Assessment of Motor and Process Skills (Fisher, 1997), and may include everyday life questionnaires to capture subjective complaints (e.g., Chan, 2001; Chiavaralloti & DeLuca, 2003; Grace, Stout, & Malloy, 1999; Wilson et al., 1996). For some of these tests, there is also evidence from neuroimaging studies that implicate the involvement of the prefrontal cortex in some of the executive functions. Despite these encouraging results, however, we argue that there are still a number of issues that need to be dealt with before these tests can effectively and successfully used in clinical practice, especially for planning rehabilitation programme.

3. Three issues of validity—functionality, ecological validity, and ethological validity

The tests described above have basically demonstrated impressive psychometric properties and clinical utilities. However, as commented by a lot of researchers (e.g., Burgess, 1997; Sbordone, 1996; Sbordone, Seyranian, & Ruff, 1998), most of the clinical or conventional tests of executive functions may be limited by their own test–retest reliabilities. Theoretically, only novel tasks can pick up deficits of executive functions, and because these tasks can only be novel once, the corresponding test–retest reliabilities are usually relatively low (Rabbitt, 1997). Moreover, there are a lot of discrepancies between experimental tasks and naturalistic tasks encountered in everyday life. In this section we would like discuss three concepts concerning the issue of validity. The first relates to functionality level, which is derived from the International Classification of Impairment, Disability and Handicap (ICIDH) (WHO, 1980) and the International Classification of Functioning, Disability and Health (WHO, 2000). According to these classifications, there are four levels of terminologies to describe how a disease may affect a person. The first two levels (viz., pathology and impairment) describe the illness within the person, whereas the latter two levels (viz., disability and handicap) describe the illness in terms of its external consequences. The latter two levels have recently been rephrased to activity and participation, and they affect a patient's functionality.

Participation reduction refers to the social, familial, educational, vocational, or other role disadvantages associated with a disability (e.g., failure in school or loss of a job because of a communication impairment). The majority of studies on executive functions, however, are primarily concentrated on the use of experimental paradigms. This has been criticized because these tasks mainly capture performance at either the pathological or impairment level (Whyte et al., 1996). The real functional status of a patient, as manifested in terms of executive dysfunctions, has not been fully acknowledged in everyday life scenarios. In terms of rehabilitation and ultimate evaluation of outcome measures, tests that are specifically designed to capture this “functionality” should be developed and evaluated.

The second issue concerns ecological validity. The lack of ecological validity has been another criticism for experimental tasks and traditional neuropsychological tests (Goldstein, 1996; Sbordone, 1996). Although quite a lot of patients with frontal lobe lesions have been found to perform equally well as controls on traditional neuropsychological tests, they still experienced a lot of difficulties in everyday life activities (Shallice & Burgess, 1991). It is suggested that the purposes of assessment require various tests of attention, memory, language, and executive functions, which may exist within a hierarchy of levels of analysis, to span along the continuum of disability (Goldstein, 1996; Whyte

Table 1
Classification of executive functions tests and levels of assessment properties in rehabilitation programme

Tests	Components	Theory	Level of assessment
Clinical rating/bedside assessment			
Cambridge Neurological Inventory (Chen et al., 1995)	Motor initiation, sequencing, and inhibition	Luria's model of mental process	Impairment
Lab-based/constraint environment			
WCST (Heaton, Chelune, Talley, Kay, & Curtiss, 1993; modified version Nelson, 1976)	Switching, perseveration	No	Impairment
Verbal Fluency Test	Verbal production	No	Impairment
Design Fluency Test (Jones-Gotman & Milner, 1977)	Nonverbal switching	No	Impairment
Stroop Test	Verbal inhibition	No	Impairment
Hayling Sentence Completion Test (Burgess & Shallice, 1996a)	Verbal inhibition	Supervisory attentional system	Impairment
Brixton Spatial Anticipation Test (Burgess & Shallice, 1996b)	Rule-detection and impulsivity	Supervisory attentional system	Impairment
Tower of London (Shallice, 1982) and Tower of Hanoi (Humes et al., 1997; Newell & Simon, 1972)	Planning	Supervisory attentional system	Impairment
Sustained Attention to Response Task (Robertson et al., 1997)	Sustained attention and inhibition	Supervisory attentional system	Impairment
N-back (Callicott et al., 1998)	Online monitoring and updating	Goldman-Rakic's working memory model	Impairment
Letter-Number Span test (Chinese version, Chan et al., in press; Gold et al., 1997)	Online monitoring and sequencing	Working Memory Model	Impairment
Six Elements Test (Wilson et al., 1996)	Planning, strategy allocation	Supervisory attentional system	Disability
Greenwich Test (Burgess et al., 2000)	Executive memory, planning and intentionality	Supervisory attentional system	Disability
Multiple Errands Test (Shallice & Burgess, 1991)	Strategy allocation, planning	Supervisory attentional system	Disability
Hotel Test (Manly et al., 2002)	Planning, strategy allocation	Supervisory attentional system	Disability
Naturalistic environment			
Gambling Task (Damasio et al., 1994)	Emotion and decision-making	Damasio's somatic marker hypothesis	Handicap and participation
Assessment of Motor and Process Skills (Fisher, 1997)	Motor and process skills	No	Handicap and participation
Naturalistic Action Test (Schwartz et al., 2002)	Planning, sequencing and strategy allocation	Supervisory attentional system/Schwartz theory	Handicap and participation
Cognitive Failures Questionnaires (Broadbent, Cooper, FitzGerald, & Parkes, 1982; Chinese version, Chan, 1999)	General everyday life cognitive failures	No	Handicap and participation
Dysexecutive Questionnaires (Wilson et al., 1996; Chinese version Chan et al., 2001a)	Intentionality, inhibition, executive memory, positive and negative affect	Supervisory attentional system	Handicap and participation
Frontal Systems Behavior Scale (Chiiaravalloti & DeLuca, 2003; Grace et al., 1999)	Initiation and disinhibition	No	Handicap and participation

& Rosenthal, 1993). It potentially starts from pathology (e.g., abnormal electrical brain potentials) to impairment (e.g., failures in sustained attention) to disability (e.g., poor performance in real-world tasks), and finally to handicap or social participation level (e.g., social roles dysfunction). Conventional experimental tasks demand relatively simple responses to single events. More complex multi-step tasks in daily life may require more complicated series of responses to achieve, including goal and sub-goals setting, prioritization of sub-goals, triggering prospective memory to initiate sub-tasks when the conditions for them become ripe, and inhibition of irrelevant and inappropriate actions to different sub-tasks. Therefore, most of the conventional experimental tasks only tackle issues at the impairment level, but cannot reflect a true picture beyond the levels of disability and handicap. In spite of its theoretical significance, it is not until recently that the construct of the SAS had been incorporated in clinical and experimental measurements of executive functions.

The final issue is what we termed ethological validity. From an evolutionary perspective, the functional capacity of the human mind is related to an environment that has existed for the last one million years, before the rapid progress of human culture. Evolutionary psychology sought to understand how the physical and social environment that existed could have been relevant to human cognitive abilities (Bradshaw, 1997). This approach is still in its infancy but has important implications. It offers an alternative approach to the current information-processing models that are largely based on consideration of cognitive abilities at an abstract level (as exemplified by the use of abstract shape, letters and numbers, which are designed to be relatively content-free). Such models are driven based on an ad hoc consideration of information-processing steps and flow, often driven by our understanding in computer systems. It is likely that this approach does not neatly delineate the natural functional modules of the mind (Beaman, 2002; Flombaum, Santos, & Hauer, 2002).

Consideration of potential purposes for which cognitive function evolved could significantly enrich neuropsychology and provide a framework for reflecting on the appropriateness of our current demarcations of cognitive abilities. These considerations are not merely theoretical. For instance, in searching for genetic underpinning of cognitive abilities, the ability to define and evaluate natural cognitive modules may be crucial to progress (Beaman, 2002; Fodor, 2000, 2002). The fact that there is a significant empirical difference between the processing of abstract symbols and the processing of meaningful material has been demonstrated in a number of studies. Perhaps the best-known example is the study using the Wason-selection task (Wason, 1966). This task is a logical problem presented in two logically equivalent forms, in one the stimulus is in letters and numbers, in the second, the stimulus refers to a real-life situation. It was found that normal people had more difficulty in performing the first task in comparison to the second. A series of control tasks has ruled out other explanations, demonstrating that it was the informational content which alters the performance (George, 1991; Hiraishi & Hasegawa, 2001). Interpreting these findings as evidence that functional modules may not be activated sufficiently by abstract stimulus. These all highlight the need for neuropsychology to include tasks, which are content-specific, in terms of evolutionary important functions of cognition. Prominent themes would include social alliances and cooperation, parenting, mate selection, and avoidance of predators. The ability to function in large social groups appears to particularly involve highly complex cognitive processes and may have driven the development of some of the most complex human cognitive function, including executive functions.

Fig. 1 illustrates the current trend of assessment of executive functions in research and clinical settings. The *x*-axis represents the number of studies done on ecological validity, the *y*-axis on functionality, and the *z*-axis on ethological validity, respectively. Therefore, a bigger star along the axis indicates a larger amount of work done on this issue. For example, the majority of the tests have primarily been focused on functionality and ecological validity. Most of the tests used, however, are laboratory-based. Very few of them could cater for all the functionality, ecological validity and ethological validity. As we argue here, a holistic consideration of these dimensions could provide us with a more accurate, sensitive, and specific information upon a specific component of executive functions.

4. Main challenges and opportunities for rehabilitation of executive functions

Rehabilitation programmes are concerned with reducing disability and handicap and/or improving social participation. For example, they aim to reduce the everyday problems of people with traumatic brain injury or schizophrenia, enable them to cope with their most appropriate environments, and/or reduce limitations and increasing activities and participation (Wilson, Evans, & Keobane, 2002). Neuropsychological tests typically measure impairments (e.g., deficits at the level of cognitive functioning). The relationship between neuropsychological test scores and everyday life problems is far from straightforward, and the aim of rehabilitation should not be just to improve test scores. Even

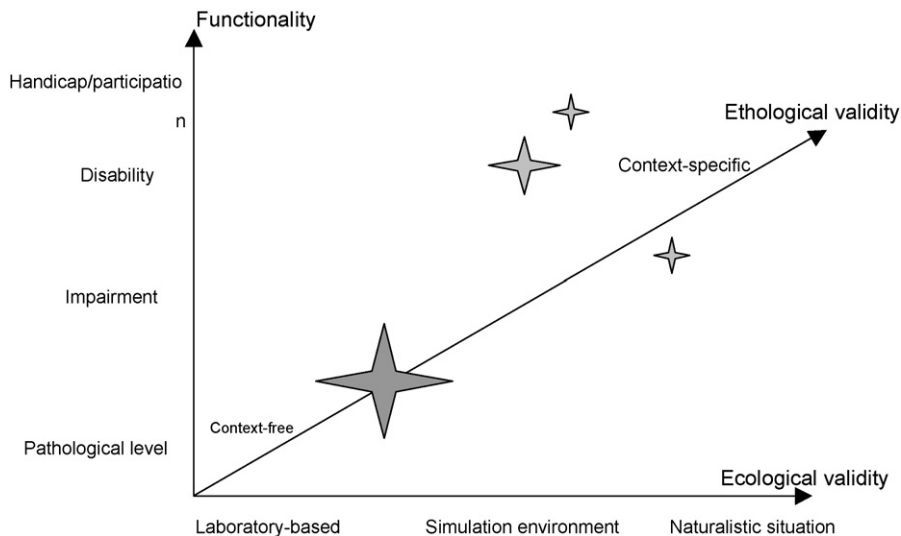


Fig. 1. Current trends of executive functions assessment in clinical settings.

measures such as return to work or independent living are problematic when used to evaluate the effectiveness of rehabilitation. The chance of a patient obtaining and maintaining employment is obviously influenced by economic factors and beyond the control of the patient and his/her rehabilitation team. Although most independent living scales differentiate between living at home and living in an institution, they are less clear cut at distinguishing between those families who are only just coping with their injuries or illness before rehabilitation, and therefore, do not reflect the complete rehabilitation outcome of the afflicted individuals.

4.1. Social cognitive neuroscience as a feasible link between laboratory-based test and everyday life functioning

Executive functions have been roughly classified into a “cold” and a “hot” component. The dorsolateral prefrontal cortex is important in mediating the “cold” one such as mechanistic planning, problem-solving, or verbal reasoning, whereas the ventromedial or orbitofrontal prefrontal cortex mediates the “hot” functions such as interpersonal and social behavior, and the interpretation of complex emotions during social interaction.

Recent advancement of social cognitive neuroscience paves a platform to link the social behaviors of the brain in terms of these “cold” and “hot” functions in everyday life functioning at a theoretical basis (e.g., Grafman & Litvan, 1999). Social cognitive neuroscience (Ochsner & Lieberman, 2001) is an emerging interdisciplinary field of research that seeks to understand phenomena in terms of interactions between three levels: the social level, which is concerned with the motivational and social factors that influence behavior and experience; the cognitive level, which is concerned with information-processing mechanisms that give rise to social-level phenomena; and the neural level, which is concerned with brain mechanisms that instantiate cognitive-level processes. The work on the Iowa gambling task and its derivatives (Bechara et al., 1996; Elliott, Friston, & Dolan, 2000; O’Doherty, Kringelbach, Hornak, Andrews, & Rolls, 2001) may serve as the examples on testing the interplay between emotion and decision-making in patients with lesions in ventromedial prefrontal cortex and amygdala.

4.2. Advanced statistics for refining, equating, and interpreting executive function performance

Despite the inherent difficulties in capturing complex neurocognitive deficits, controlled quantification remains a desirable goal for both clinicians and researchers. A principal approach to examining the dysexecutive syndrome has therefore been to develop tasks that maximally tax one or other hypothesized impairment. As mentioned above, emphasis has recently been paid on the articulation between laboratory-based test and real-life activities. However, traditional methods of equating different tests presumably assessing the same constructs of neurocognitive performance poses some analytical problems that have made it difficult to compare and interpret results. First, the equivalence of

scores of is heavily dependent on the characteristics of the population from which the groups are drawn. Therefore, the conversion scale for any given group may not be altogether accurate for the other groups. Most probably, the assumption of normal score distribution is not true in many situations. Second, the traditional models are norm-referenced ones. That means the distribution of a specific test score will be ultimately determined by the ability of a specific referenced group at a particular time point. In other words, the score distribution of a specific test may be unstable across time periods.

The recent advancement of the Item Response Theory and Rasch Analysis (Rasch, 1960; Wright & Masters, 1982; Wright & Stone, 1979) may solve parts of these problems. The Rasch model can create a scale of task performance so that both item difficulty and person ability are calibrated on the same scale. The model produces scale-free person measures and sample-free item difficulties (Wright & Masters, 1982; Wright & Stone, 1979). Parameters are estimated for each person (B_n) and item (D_i) and their differences ($B_n - D_i$) are compared against unidimensional measurement model predictions. This approach represents an important shift from specific sample characteristics common in conventional “score” analyses to differences between person and item parameters and their objective magnitude. When quantitative properties of these differences conform to the one-parameter logistic, person and item estimates are measures on an absolute scale. A related property of this difference ($B_n - D_i$) is a systematic accumulation of scale units across the interval scale further supporting scale linearity (Wright & Masters, 1982; Wright & Stone, 1979). Moreover, the rater’s bias towards an individual’s performance can also be resolved by using this type of statistics (Linacre, 1989). The problem of leniency in rating the same individual by different raters on clinical symptoms of executive behavior as well as performance on functional ability tests could be eliminated.

Although the specific models being used may vary, this methodology has the potential to substantially advance the field of neuropsychological assessment and intervention (Bode, Cella, Lai, & Heinemann, 2003; Bode, Heinemann, & Semik, 2000; Fisher, 1997). For instance, Chan and Bode (2004, 2008) applied the Rasch model to examine the dysexecutive behavior patients with traumatic brain injury and found that there were differential item functioning being reported by patients and significant others, that is, discrepancies in reporting individual items of dysexecutive behavior. These were distractibility, temporal processing problems, poor decision-making ability, knowing–doing dissociation, and lack of concern. Chan and Bode (2008) suggest that simple subtraction of items between patients and significant others may not truly reflect the phenomenon of insight deficits in this sample because the discrepancy may simply imply a differential item functioning in a statistical sense. Chan, Linacre, and Wright (2004) also provide preliminary evidence to equate different executive function test performances among a group of patients with traumatic brain injury using the Rasch Model. In so doing, these authors suggest that we could compute the specific executive functioning performance along the same yardstick (or construct in psychological term) when sufficiently large sample has been collected. Although these propositions of using Rasch Model to establish a common metrics or yardstick in computing specific executive functioning requires more rigorous methodology and cross-validation, this sophisticated statistical technique may serve as a potential tool for researcher to refine the measurement of executive functions in the near future.

4.3. Advanced technology

The application of virtual reality in the assessment and rehabilitation of executive functions may help to tackle the issues of ecological and ethological validity discussed earlier. Preliminary work has been launched by some researchers in assessing executive functions using virtual reality. For example, Morris and colleagues (Miotto, Bullock, Polkey, & Morris, 1996; Morris et al., 1998) designed the Executive Golf Task to capture the online updating component of executive functioning. This test involves asking individuals to search around a set of specific locations and subsequently remembering not to return to successful positions. The task simulates a game of golf whereby subjects are asked to putt a ball, according to a specified set of rules, into each golf hole presented on a touch-sensitive computer screen. The advantage of this task is that it can be incorporated with imaging technique to study brain activation in situations (in virtual reality scenario) that is relevant and meaningful to a particular patient. This may help to improve the ecological and ethological validity of the assessment instrument.

However, this advanced technology (and other related equipments) should be carefully implemented in clinical setting. It is particularly true for those patients who are not familiar with computerized tests or who are anxious in undertaking computerized test or being tested in a semi-enclosed environment (Browndyke et al., 2002; Wiechmann & Ryan, 2003). Therefore, clinicians should be sensitive enough to screen out these cases before implementing these

advanced techniques. Moreover, it should be noted that while virtual reality tests may have excellent face validity, its ecological validity may be no better than standard neuropsychological instruments. Clinicians should not be completely biased toward new technologies and be totally dependent on the results generated from them. They should take a more balanced or cautionary perspective between technology and conventional neuropsychological instruments.

5. Conclusions

Although there is no gold standard for assessment of executive functions, many tests are now available to capture specific components of this complex construct. With the advancement of theories and models of executive functions, the second-generation tests focus more on the theoretical account of the underlying deficits of specific components. These major advancements can facilitate a more accurate, sensitive and specific assessment of executive function performance. However, most tests or tasks could only address the functionality and ecological validity dimensions. Very few of them were specifically designed to capture executive functions on a context-specific situation. As we argue previously, a holistic consideration of functionality level, ecological validity and ethological validity could lead to more accurate, sensitive, and specific information about a specific component of executive functions. Moreover, the theoretical link between social, cognitive, and neuroscience factors could contribute to the formulation of a better understanding of executive dysfunctions manifestation in the context of different cultures. The integration of these factors is, therefore, important for the development of neuropsychological tests, particularly executive functions tests, in the future.

Acknowledgements

This study was supported partially by the Research Initiation Fund of the 100-Scholar Programme (O7CX031003) and the Research Fund (KSCX2-YW-R-131) from the Institute of Psychology, Chinese Academy of Sciences to Raymond Chan.

References

- Allport, D. A., Styles, E. A., & Hsieh, S. (1994). Shifting intention set: Exploring the dynamic control of tasks. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance XV* (pp. 421–452). Cambridge, MA: MIT Press.
- Amieva, H., Phillips, L., & Della Sala, S. (2003). Behavioural dysexecutive symptoms in normal aging. *Brain and Cognition*, *53*(2), 129–132.
- Baddeley, A. D. (1986). *Working memory*. Oxford: Clarendon Press.
- Beaman, C. P. (2002). Why are we good at detecting cheaters? A reply to Fodor. *Cognition*, *83*(2), 215–220.
- Bechara, D., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*, 7–12.
- Bechara, D., Damasio, H., Damasio, A. R., & Lee, G. P. (1999). Different contributions of the human amygdale and ventromedial prefrontal cortex to decision-making. *Journal of Neuroscience*, *19*, 5473–5481.
- Bechara, D., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science*, *275*, 1293–1295.
- Bechara, D., Tranel, D., & Damasio, H. (2000). Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. *Brain*, *123*, 2189–2202.
- Bechara, D., Tranel, D., Damasio, H., & Damasio, A. R. (1996). Failure to respond automatically to anticipated future outcomes following damage to prefrontal cortex. *Cerebral Cortex*, *6*, 215–225.
- Bode, R. K., Cella, D., Lai, J. S., & Heinemann, A. W. (2003). Developing an initial physical function item bank from existing sources. *Journal of Applied Measurement*, *4*(2), 1–23.
- Bode, R. K., Heinemann, A. W., & Semik, P. (2000). Measurement properties of the Galveston Orientation and Amnesia Test (GOAT) and improvement patterns during inpatient rehabilitation. *Journal of Head Trauma Rehabilitation*, *15*(1), 637–655.
- Botvnick, M., Nystrom, L. E., Fissell, K., Carter, C. S., & Cohen, J. D. (1999). Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature*, *402*, 179–181.
- Bradshaw, J. L. (1997). *Human evolution: A neuropsychological perspective*. East Essex: Psychology Press.
- Broadbent, D. B., Cooper, P. F., FitzGerald, P., & Parkes, K. R. (1982). The Cognitive Failures Questionnaire (CFQ) and its correlates. *British Journal of Clinical Psychology*, *21*, 1–16.
- Brownndyke, J. N., Albert, A. L., Malone, W., Schartz, P., Paul, R. H., Cohen, R. A., et al. (2002). Computer-related anxiety: Examining the impact of technology-specific affect on the performance of a computerized neuropsychological assessment measure. *Applied Neuropsychology*, *9*(4), 210–218.
- Burgess, P. W. (1997). Theory and methodology in executive function research. In P. Rabbitt (Ed.), *Methodology of frontal and executive function* (pp. 81–116). East Sussex, UK: Psychology Press Ltd.

- Burgess, P. W. (2000). Real-world multitasking from a cognitive neuroscience perspective. In S. Monsell & J. Driver (Eds.), *Control of cognitive processes: Attention and performance XVIII* (pp. 465–472). Massachusetts: The MIT Press.
- Burgess, P. W., Alderman, N., Evans, J., Emslie, H., & Wilson, B. A. (1998). The ecological validity of tests of executive function. *The Journal of the International Neuropsychological Society*, 4(6), 547–558.
- Burgess, P. W., & Shallice, T. (1996a). Response suppression, initiation, and strategy use following frontal lobe lesions. *Neuropsychologia*, 34, 263–273.
- Burgess, P. W., & Shallice, T. (1996b). Bizarre responses, rule detection, and frontal lobe lesions. *Cortex*, 32, 1–19.
- Burgess, P. W., & Shallice, T. (1997). The relationship between prospective and retrospective memory: Neuropsychological evidence. In M. A. Conway (Ed.), *Cognitive models of memory* (pp. 74–90). Hove: Psychology Press.
- Burgess, P. W., Veitch, E., de lacy Costello, A., & Shallice, T. (2000). The cognitive and neuroanatomical correlates of multi-tasking. *Neuropsychologia*, 38, 848–863.
- Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences*, 4(6), 215–222.
- Callicott, J. H., Ramsey, N. F., Tallent, K., Bertolino, A., Knable, M. B., Coppola, R., et al. (1998). Functional magnetic resonance imaging brain mapping in psychiatry: Methodological issues illustrated in a study of working memory in schizophrenia. *Neuropsychopharmacology*, 18, 186–196.
- Casey, B. J., Thomas, K. M., Welsh, T. F., Badgalyan, R., Eccard, C. H., Jennings, J. R., et al. (2000). Dissociation of response conflict, attentional selection, and expectancy with functional magnetic resonance imaging (fMRI). *Proceedings of the National Academy of Sciences of the United States of America*, 97, 8728–8733.
- Chan, R. C. K. (1999). What does Cognitive Failures Questionnaire measure: General failures or specific domain deficits? *Archives of Clinical Neuropsychology*, 14, 735–736.
- Chan, R. C. K. (2001). Dysexecutive symptoms among the non-clinical sample: A study with the use of the Dysexecutive Questionnaire. *British Journal of Psychology*, 92(3), 551–565.
- Chan, R. C. K. (2002). Attentional deficits in patients with persisting postconcussive complaints: General deficit or specific component deficits? *Journal of Clinical and Experimental Neuropsychology*, 24(8), 1081–1093.
- Chan, R. C. K., & Bode, R. (2004). Analysis of patient and proxy ratings on the Dysexecutive Questionnaire. *Brain Impairment*, 5(Supplement), 141.
- Chan, R. C. K., & Bode, R. (2008). Analysis of patient and proxy ratings on the dysexecutive questionnaire: An application of Rasch analysis. *Journal of Neurology, Neurosurgery, and Psychiatry*, 79, 86–88.
- Chan, R. C. K., & Chen, E. Y. H. (2004). Executive dysfunctions and neurological manifestations in schizophrenia. *Hong Kong Journal of Psychiatry*, 14(3), 2–6.
- Chan, R. C. K., Chen, E. Y. H., Cheung, E. F. C., Chen, R. Y. L., & Cheung, H. K. (2006). The components of executive functioning in a cohort of patients with chronic schizophrenia: A multiple single-case study design. *Schizophrenia Research*, 81(2–3), 173–189.
- Chan, R. C. K., Chen, E. Y. H., Cheung, E. F. C., & Cheung, H. K. (2004). Executive dysfunctions in schizophrenia: Relationships to clinical manifestations. *European Archives of Psychiatry and Clinical Neuroscience*, 254, 256–262.
- Chan, R. C. K., Chen, E. Y. H., & Law, C. W. (2006). Specific executive dysfunction in patients with first-episode medication-naïve schizophrenia. *Schizophrenia Research*, 82(1), 51–64.
- Chan, R. C. K., Hoosain, R., Lee, T. M. C., Fan, Y. W., & Fong, D. T. S. (2003). Are there subtypes of attentional deficits in patients with persisting postconcussive complaints? A cluster analytical study. *Brain Injury*, 17(2), 131–148.
- Chan, R. C. K., & Manly, T. (2002). The application of “dysexecutive syndrome” measures across cultures: Performance and checklist assessment in neurologically healthy and traumatic brain-injured Hong Kong Chinese volunteers. *Journal of the International Neuropsychological Society*, 8, 778–780.
- Chan, R. C. K., Linacre, J. M., & Wright, B. D. (2004). Using Rasch statistics to rescale neuropsychological performance in patients with postconcussion symptoms. *Brain Impairment*, 5(Supplement), 175–176.
- Chan, R. C. K., Rao, H., Chen, E. Y. H., Ye, B., & Zhang, C. (2006). The neural basis of motor sequencing: An fMRI study of healthy subjects. *Neuroscience Letters*, 398, 189–194.
- Chan, R. C. K., Wang, Y., Deng, Y., Zhang, Y., Yao, X., & Zhang, C. (in press). The development of a Chinese equivalence version of Letter-Number Span Test. *Clinical Neuropsychologist*.
- Chen, E. Y. H., Shapleske, J., Luque, R., McKenna, P. J., Hodges, J. R., Calloway, S. P., et al. (1995). The Cambridge Neurological Inventory: A clinical instrument for soft neurological signs and the further neurological examination for psychiatric patients. *Psychiatry Research*, 56, 183–202.
- Chiavarallotti, N. D., & DeLuca, J. (2003). Assessing the behavioral consequences of multiple sclerosis: An application of the Frontal Systems Behavior Scale (FrSBs). *Cognitive and Behavioral Neurology*, 16, 54–67.
- Crawford, J. R., & Henry, J. D. (2005). Assessment of executive function. In P. W. Halligan & D. T. Wade (Eds.), *Effectiveness of rehabilitation for cognitive deficits* (pp. 233–246). Oxford, UK: Oxford University Press.
- Damasio, A. R. (1995). Toward a neurobiology of emotion and feeling: Operational concepts and hypotheses. *The Neuroscientist*, 1(1), 19–25.
- Damasio, H., Grabowski, T., Frank, R., Galaburda, A. M., & Damasio, A. R. (1994). The return of Phineas Gage: Clues about the brain from the skull of a famous patient. *Science*, 264, 1102–1105.
- Das, J. P., Naglieri, J. A., & Kirby, J. R. (1994). *Assessment of cognitive processes: The PASS theory of intelligence*. Boston: Allyn and Bacon.
- Duncan, J. (1986). Disorganization of behaviour after frontal lobe damage. *Cognitive Neuropsychology*, 2, 271–290.
- Duncan, J. (1995). Attention, intelligence and the frontal lobes. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 721–733). Cambridge, MA: MIT Press.

- Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neuroscience*, 23, 475–483.
- Duncan, J., Seitz, R. J., Kolodny, J., Bor, D., Herzog, H., Ahmed, A., et al. (2000). A neural basis for general intelligence. *Science*, 289, 457–460.
- Elliott, R., Friston, K. J., & Dolan, R. J. (2000). Dissociable neural responses in human reward systems. *Journal of Neuroscience*, 20, 6159–6165.
- Evans, J. J., Chua, S. E., McKenna, P. J., & Wilson, B. A. (1997). Assessment of the dysexecutive syndrome in schizophrenia. *Psychological Medicine*, 27, 635–646.
- Fan, J., McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience*, 14(3), 340–347.
- Fisher, W. P., Jr. (1997). Physical disability construct convergence across instruments: Towards a universal metric. *Journal of Outcome Measurement*, 1(2), 87–113.
- Flombaum, J. I., Santos, L. R., & Hauer, M. D. (2002). Neuroecology and psychological modularity. *Trends in Cognitive Sciences*, 6(3), 106–108.
- Fodor, J. (2000). Why we are so good at catching cheaters? *Cognition*, 75(1), 29–32.
- Fodor, J. (2002). A reply. *Cognition*, 83(2), 221.
- Fox, G., & Fox, A. M. (2001). The effects of brain damage on the performance of hand movement sequences. *Brain Impairment*, 2, 140–144.
- George, C. (1991). Facilitation in the Wason selection task with a consequent referring to an unsatisfactory outcome. *British Journal of Psychology*, 82(1), 463–472.
- Goel, V., Grafman, J., Tajik, J., Gana, S., & Danto, D. (1997). A study of the performance of patients with frontal lobe lesions in a financial planning task. *Brain*, 120, 1805–1822.
- Gold, J. M., Carpenter, C., Randolph, C., Goldberg, T. E., & Weinberger, D. R. (1997). Auditory working memory and Wisconsin Card Sorting Test performance in schizophrenia. *Archives of General Psychiatry*, 54(2), 159–165.
- Goldman-Rakic, P. S. (1992). Working memory and the mind. *Scientific American*, 267(3), 73–79.
- Goldstein, G. (1996). Functional considerations in neuropsychology. In R. J. Sbordone & C. J. Long (Eds.), *Ecological validity of neuropsychological testing* (pp. 75–89). Delray Beach, Florida: GR Press/St. Lucie Press.
- Grace, J., Stout, J. C., & Malloy, P. F. (1999). Assessing frontal lobe behavioral syndromes with the Frontal Lobe Personality Scale. *Assessment*, 6, 269–284.
- Grafman, J., & Litvan, I. (1999). Importance of deficits in executive functions. *The Lancet*, 354, 1921–1923.
- Grafman, J., Schwab, K., Warden, D., Pridgen, A., Grown, H. R., & Salazar, A. M. (1996). Frontal lobe injuries, violence, and aggression: A report of the Vietnam Head Injury Study. *Neurology*, 46, 1231–1238.
- Green, M. F. (1996). What are the functional consequences of neurocognitive deficits in schizophrenia? *American Journal of Psychiatry*, 153(3), 321–330.
- Green, M. F., Kern, R. S., Braff, D. L., & Mintz, J. (2000). Neurocognitive deficits and functional outcome in schizophrenia: Are we measuring the “right stuff”? *Schizophrenia Bulletin*, 26(1), 119–136.
- Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtiss, G. (1981, 1993). *Wisconsin Card Sorting Test Manual*. Odessa: Psychological Assessment Resources.
- Heinrichs, D. W., & Buchanan, R. W. (1988). Significance and meaning of neurological signs in schizophrenia. *American Journal of Psychiatry*, 145, 11–18.
- Hiraishi, K., & Hasegawa, T. (2001). Sharing-rule and detection of free-riders in cooperative groups: Evolutionarily important deontic reasoning in the Wason selection task. *Thinking and Reasoning*, 7(3), 255–294.
- Humes, G. E., Welsh, M. C., Retzlaff, P., & Cookson, N. (1997). Towers of Hanoi and London: Reliability and validity of two executive function tasks. *Assessment*, 4, 249–257.
- Jones-Gotman, M., & Milner, B. (1977). Design fluency: The invention of nonsense drawings after focal cortical lesions. *Neuropsychologia*, 15, 653–674.
- Kane, M., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin Review*, 9(4), 637–671.
- Kaufman, A., & Kaufman, N. (1983). *Kaufman Assessment Battery for Children*. Circle Pines, MN: American Guidance Service.
- Korkman, M. (1999). Applying Luria’s diagnostic principles in the neuropsychological assessment of children. *Neuropsychology Review*, 9, 89–105.
- Lin, H., Chan, R. C. K., Zheng, L., Yang, T., & Wang, Y. (2007). Executive functioning in healthy elderly Chinese people. *Archives of Clinical Neuropsychology*, 22, 501–511.
- Levine, B., Robertson, I. H., Clare, L., Carter, G., Hong, J., Wilson, B. A., et al. (2000). Rehabilitation of executive functioning: An experimental-clinical validation of Goal Management Training. *Journal of the International Neuropsychological Society*, 6, 299–312.
- Linacre, J. M. (1989). *Many-Facet Rasch Measurement*. Chicago: MESA Press.
- Luria, A. R. (1966). *Higher cortical functions in man*. New York: Basic Books.
- Luria, A. R. (1973). *The working brain*. London: Penguin.
- Manly, T., Hawkins, K., Evans, J., Woldt, K., & Robertson, I. H. (2002). Rehabilitation of executive function: Facilitation of effective goal management on complex tasks using periodic auditory alerts. *Neuropsychologia*, 40, 271–281.
- Miotto, E. C., Bullock, P., Polkey, C. E., & Morris, R. G. (1996). Spatial working memory and strategy formation in patients with frontal lobe excisions. *Cortex*, 32, 613–630.
- Mirsky, A. F. (2003). Cross-cultural assessment of attention-related cognitive functions: Studies in Canada, Ecuador, Ireland, Israel, and the United States. *An international conference on neuropsychology: Recent advances in the East and West*, December 9–12, 2003, Hong Kong.
- Morris, R. G., Downes, J. J., Sahakian, B. J., Evenden, J. L., Heald, A., & Robbins, T. W. (1998). Planning and spatial working memory in Parkinson’s disease. *Journal of Neurology, Neurosurgery, and Psychiatry*, 51, 757–766.
- Nelson, H. (1976). A modified card sorting test sensitive to frontal lobe deficits. *Cortex*, 12, 313–324.

- Newell, A., & Simon, H. A. (1972). *Human problem solving*. New York: Prentice Hall.
- Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behaviour. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research and theory* (pp. 1–18). New York: Plenum.
- Ochsner, K. N., & Lieberman, M. D. (2001). The emergence of social cognitive neuroscience. *American Psychologist*, 56(9), 717–734.
- O'Doherty, J., Kringsbach, M. I., Hornak, J., Andrews, C., & Rolls, E. T. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neuroscience*, 4, 95–102.
- Posner, M. I., & Raichle, M. E. (1994). *Images of mind*. New York: Scientific American Library.
- Rabbitt, P. (1997). Introduction: Methodologies and models in the study of executive function. In P. Rabbitt (Ed.), *Methodology of frontal and executive function* (pp. 1–38). East Sussex, UK: Psychology Press Ltd.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Demarks Paedagogiske Institut. (Reprint in Wright, B. D. (1980). With foreword and backwords. *Probabilistic models for some intelligence and attainment tests*. Chicago: University of Chicago Press.).
- Reason, J. (1984). Lapses of attention in everyday life. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 515–549). London: Academic Press Inc.
- Robertson, I. H. (1996). *Goal management training: A clinical manual*. Cambridge, UK: PsyConsult.
- Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. T., & Yiend, J. (1997). Oops! Performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia*, 35(6), 747–758.
- Rogers, R. D., & Monsell, S. (1995). Costs of predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, 124, 207–231.
- Rolls, E. T. (1995). A theory of emotion and consciousness and its application to understanding the neural basis of emotion. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences*. Cambridge, MA: MIT Press.
- Royall, D. R., Lauterbach, E. C., Cummings, J. L., Reeve, A., Rummans, T. A., Kaufer, D. I., et al. (2002). Executive control function: A review of its promise and challenges for clinical research. *Journal of Neuropsychiatry and Clinical Neurosciences*, 14, 377–405.
- Rushworth, M. F. S., & Owen, A. M. (1998). The functional organization of the lateral frontal cortex: Conjecture or conjuncture in the electrophysiology literature? *Trends in Cognitive Science*, 2(2), 46–53.
- Sbordone, R. J. (1996). Ecological validity: Some critical issues for neuropsychologist. In R. J. Sbordone & C. J. Long (Eds.), *Ecological validity of neuropsychological testing* (pp. 15–41). Delray Beach, Florida: GR Press/St. Lucie Press.
- Sbordone, R. J., Seyranian, G. D., & Ruff, R. M. (1998). Are the subjective complaints of traumatically brain injured patients reliable? *Brain Injury*, 12(6), 505–515.
- Schwartz, M. F., Reed, E. S., Montgomery, M., Plamer, C., & Mayer, H. H. (1991). The quantitative description of action disorganization after brain damage: A case study. *Cognitive Neuropsychology*, 8, 381–414.
- Schwartz, M. F., Segal, M., Veramonti, T., Ferraro, M., & Buxbaum, L. J. (2002). The Naturalistic Action Test: A standardized assessment for everyday action impairment. *Neuropsychological Rehabilitation*, 12(4), 311–339.
- Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London*, 298, 199–209.
- Shallice, T. (1988). *From neuropsychology to mental structure*. Cambridge, UK: Cambridge University Press.
- Shallice, T., & Burgess, P. W. (1991). Deficits in strategy application following frontal lobe damage in man. *Brain*, 114, 727–741.
- Stuss, D. T., Alexander, M. P., Shallice, T., Picton, T. W., Binns, M. A., Macdonald, R., et al. (2005). Multiple frontal systems controlling response speed. *Neuropsychologia*, 43, 396–417.
- Stuss, D. T., & Benson, D. F. (1986). *The frontal lobes*. New York: Raven Press.
- Stuss, D. T., Shallice, T., Alexander, M. P., & Picton, T. W. (1995). A multidisciplinary approach to anterior attentional functions. *Annals of the New York Academy of Sciences*, 769, 191–209.
- Tranel, D., Bechara, A., & Damasio, A. R. (2000). Decision making and the somatic marker hypothesis. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences* (pp. 1047–1061). Cambridge, MA: MIT Press.
- Umetsu, A., Okuda, J., Fuji, T., Tsukiura, T., Nagasaka, T., Yanagawa, I., et al. (2002). Brain activation during the fist-edge-palm test: A functional MRI study. *Neuroimage*, 12, 385–392.
- Wason, P. C. (1966). Reasoning. In B. Foss (Ed.), *New horizons in psychology* (pp. 135–151). London: Penguin.
- Whyte, J., Polansky, M., Cavallucci, C., Fleming, M., Lhulier, J., & Coslett, B. (1996). Inattentive behaviour after traumatic brain injury. *Journal of the International Neuropsychological Society*, 2, 274–281.
- Whyte, J., & Rosenthal, M. (1993). Rehabilitation of the patient with traumatic brain injury. In J. A. DeLisa (Ed.), *Rehabilitation medicine: Principles and practice* (pp. 825–860). Philadelphia: J.B. Lippincott.
- Wiechmann, D., & Ryan, A. M. (2003). Reactions to computerized testing in selection contexts. *International Journal of Selection and Assessment*, 11(2–3), 215–229.
- Wilson, B. A., Alderman, N., Burgess, P., Emslie, H., & Evans, J. (1996). *Behavioral assessment of the dysexecutive syndrome*. Bury St. Edmunds, Suffolk: Thames Valley Test Company.
- Wilson, B. A., Evans, J. J., & Keobane, C. (2002). Cognitive rehabilitation: A goal-planning approach. *Journal of Head Trauma Rehabilitation*, 17(6), 542–555.
- World Health Organization. (1980). *International Classification of Impairments, Disabilities and Handicaps. A manual of classification relating to consequences of disease*. Geneva: WHO.
- World Health Organization. (2000). *International Classification of Functioning, Disability and Health*. Geneva, Switzerland: WHO.
- Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis—Rasch measurement*. Chicago: MESA Press.
- Wright, B. D., & Stone, M. H. (1979). *Best test design*. Chicago: MESA Press.
- Yarnell, P. R., & Rossi, G. V. (1988). Minor whiplash head injury with major debilitation. *Brain injury*, 2, 255–258.