

Celeration of Executive Functioning while Solving the Tower of Hanoi: Two Single Case Studies Using Protocol Analysis

**Gwenny T.L. Janssen^{*1,2} Hubert R.A. De Mey² Jos I.M. Egger^{1,2,3}
and Cilia L.M. Witteman²**

¹ *Centre of Excellence for Neuropsychiatry, Vincent van Gogh Institute for Psychiatry, The Netherlands*

² *Behavioural Science Institute, Radboud University Nijmegen, The Netherlands*

³ *Pompe Institute for Forensic Mental Health, Nijmegen, The Netherlands*

ABSTRACT

The present study was designed to systematically explore individual learning abilities by an in depth analysis of the performances of two single participants on the five disk Tower of Hanoi. Verbal protocols were obtained to provide us with a detailed picture of strategy development and learning. Correct and incorrect responses were plotted on a Standard Celeration Chart in order to analyze the growth of learning across time (i.e., celeration). Protocol analysis revealed that both participants formulated rules, indicating the presence of previously learned problem-solving strategies or “schemas,” but also showed a substantial difference in adequately adapting these rules to modify their executive behaviour. Furthermore, celeration trends indicated a markedly increase of speed and accuracy level for performances of participant A, whereas performances of participant B showed no (significant) changes in behaviour accuracy. We concluded that a systematic investigation of the actual processes that underlie the acquisition of strategies using protocol analysis, can provide us with a detailed picture of individual differences on EF. Additionally, the Standard Celeration Chart proved to be a good assessment method to monitor behavioural changes and individual learning. On the critical side, further investigation of the differential effects of age, gender and pathology on individual performance seems to be warranted. *Key words:* executive function, celeration, Tower of Hanoi, protocol analysis, psychopathology, learning.

RESUMEN

El presente estudio se diseñó para explorar sistemáticamente habilidades de aprendizaje individual mediante un profundo análisis de las ejecuciones de dos participantes en los cinco discos de la Torre de Hanoi. Se obtuvieron protocolos verbales para proporcionar una detallada imagen del desarrollo de estrategias y aprendizaje. Las respuestas correctas e incorrectas se dispusieron sobre un gráfico de aceleración en orden a analizar el crecimiento del aprendizaje a lo largo del tiempo. El análisis del protocolo reveló que ambos participantes formularon reglas, indicando la presencia de estrategias de solución de problemas previamente aprendidas o “esquemas”, pero también mostraron una diferencia sustancial en adaptar adecuadamente estas reglas para modificar su ejecución. Además, las tendencias de aceleración indican un marcado incremento de la velocidad y del nivel de

* Correspondence should be addressed to Gwenny Janssen, Vincent van Gogh Institute for Psychiatry, Stationsweg 46, 5803 AC, Venray, The Netherlands. Email: gjanssen@ggznm.nl

precisión en las ejecuciones del participante A, mientras que las ejecuciones del participante B no mostraron cambios significativos en la precisión. Se concluye que la investigación sistemática de los procesos que subyacen la adquisición de estrategias usando análisis de protocolos, puede proporcionar una imagen detallada de las diferencias individuales en funcionamiento ejecutivo. Adicionalmente, el empleo de la *Standard Celeration Chart* prueba ser un buen método de evaluación para observar los cambios conductuales y el aprendizaje individual. Son necesarios nuevos estudios sobre los efectos diferenciales de la edad, el sexo y la presencia de patologías.

Palabras clave: función ejecutiva, *celeration*, Torre de Hanoi, análisis de protocolo, psuchoathology, aprendizaje.

The term executive function (EF) has been used to refer to the dimensions of complex human behaviour that are primarily involved in the control and direction of self-regulating behaviour (Cicerone, Levin, Malec, Stuss, & Whyte, 2006). EF has been associated with several (neuro-)psychiatric conditions such as autism spectrum disorder (Hughes, Russell, & Robbins, 1994; Miller & Ozonoff, 2000), attention-deficit/hyperactivity disorder (Happé, Booth, Charlton, & Hughes, 2006; Sergeant, Geurts, & Oosterlaan, 2002), obsessive-compulsive disorder (Kuelz, Hohagen, & Voderholzer, 2004), Korsakoff's syndrome (Brokate *et al.*, 2003), Alzheimer's dementia (Graham, Emery, & Hodges, 2004) and schizotypal personality disorder (Voglmaier *et al.*, 2000). Also, EF has been linked to the development of varying psychological processes, including rule following and theory of mind (Hayes, Gifford, & Ruckstuhl Jr, 1996; Hill & Frith, 2003; Zelazo, Carter, Reznick, & Frye, 1997).

EF encompasses cognitive skills such as updating of working memory representations, shifting between tasks or mental sets, and inhibition of dominant or prepotent responses (Miyake *et al.*, 2000). Barkley (2001) defined executive processes as "any act toward oneself that functions to modify one's own behaviour so as to change the future outcomes for that individual." Consequently, executive dysfunctioning is the loss of these modifying or self-regulating abilities. This view of EF emphasizes the controlling aspects of executive processes and stresses the importance of EF to human autonomy.

However, the capacity to formulate or to generate rules or strategies and to anticipate likely future outcomes, is not necessarily an "executive" skill. The loss of executive control divorces the availability of strategies from successful implementation of the intended behaviour (Royall *et al.*, 2002; Zelazo *et al.*, 1997). Therefore, executive skills are most prominently needed in novel, often unexpected situations in which automatic or well trained behaviour will not lead to effective solutions.

An example will illustrate this. Imagine your car breaking down when you are on your way to your best friend's wedding. As a matter of fact, you are his best man and treasurer of the wedding rings. Also, you always had strong feelings for his wife-to-be. Would your behaviour be the same as when it were a normal working day with a friendly and understanding boss? Or would you react in a different manner, either more or less efficient and organized, when under the influence of stress, anxiety, determination

or even despair? Dependent upon the circumstances or context, executive behaviour may differ within an individual. Although the availability of certain rules, strategies or schemas remains unchanged, implementation of these skills may vary considerably. In other words, when people under new or unusual circumstances lose their ability of executive control, they do not necessarily lose the capacity to formulate or anticipate problem solving strategies, but rather lose the capacity to successfully implement and execute goal-directed and purposeful behaviour.

There is a vast body of literature on EF but, due to inconsistency of definitions and lack of a compelling theory, a clear characterization of the underlying cognitive processes involved in EF or of the neural architecture that supports EF is lacking. Zelazo *et al.* (1997) argued that, instead of simply listing descriptions of prefrontal impairments, “a characterisation of the complex processes attributed to the executive function is needed” (p. 199). Their approach, following Luria (1973), views EF as a multifaceted framework that refers to the underlying cognitive processes involved in goal-directed problem solving. Even so, as Hayes and colleagues (1996) pointed out, by looking at the actual (observable) problem solving behaviours, EF can be understood in terms of its functional parameters, and thus form the key to definition and differential diagnosis.

With the systematic study of the performance of individuals with different EF-levels and experience in solving well-defined problems, lapses of EF in the sequence of behaviour can be located fairly precisely. This will not only aid our understanding of everyday problem solving (Williams & Noyes, 2007), but will also have great importance for understanding how problem-solving skills are acquired and how they are related to the executive (dys)abilities of individuals.

The Tower of Hanoi puzzle (TOH) has proved to be a suitable task for studying a variety of executive processes (Welsh & Huizinga, 2005) and is used extensively both in clinical and non-clinical samples. The task has demonstrated sensitivity to prefrontal lobe dysfunction (Miyake *et al.*, 2000), and taps processes such as planning, working memory, updating and inhibition (Handley, Capon, Copp, & Harper, 2002; Welsh, Satterlee-Cartmell, & Stine, 1999; Zook, Davalos, DeLosh, & Davis, 2004). In his analysis of the TOH, Simon (1975) suggested that the constraints of this task are conducive to the spontaneous generation of several problem solving strategies that vary in effectiveness and may explain normal individual differences in performance. He stated that “different participants may in fact learn different things in the same task environment.” Through analyzing the performance of one single participant by close consideration of a verbal protocol, Anzai and Simon (1979) demonstrated that more systematic and advanced methods were used as experience increased, enabling their participant to show more deterministic and goal-directed behaviour in successfully solving the TOH.

Being interested in the actual processes that underlie the acquisition of EF strategies, we consider it necessary to complement traditional classificatory (static) diagnosis with a dynamic assessment of the process of problem solving, i.e., of strategy development and implementation during the TOH. Apart from protocol analysis, this process could also be quantified with celeration charts, a computational method which will be described in more detail in the section below.

The protocol analysis method (Ericsson & Simon, 1984) has proved considerable utility in the study of complex human behaviour. Research has focused on areas such as equivalence class formation (Wulfert, Dougher, & Greenway, 1991), and the role of private verbal behaviour on the operant performance of human adults (Cabello, Luciano, Gomez, & Barnes-Holmes, 2004). Cabello and O'Hora (2002) stated that these think aloud protocols, as used in the experimental analysis of behaviour, "provide us with a detailed analysis of the effect of specific self-instructions and strategies in the completion of certain tasks." The main feature of this technique is that participants are instructed to say aloud everything they are thinking about, allowing a moment to moment analysis of the relation between what participants are saying and what they are doing (Cabello & O'Hora, 2002; Ericsson & Simon, 1984). It is based on the assumption that the boundary between private and public speech does not introduce a fundamental difference in behaviour. During our experiment both think-aloud and retrospective reports were recorded to (a) assess the completeness of the think aloud protocol, and (b) to assure that the retrospective reports contain an actual record of the cognitive problem solving processes.

To measure not only the efficiency of problem solving behaviour, but also the progress of learning, we used Standard Celeration Charts, as developed by Ogden R. Lindsley (Potts, Eshleman, & Cooper, 1993). Competent performance on complex tasks is characterized by the fluid combination of accuracy and speed (frequency of responding). Binder (1996) has suggested that this "fluency" of performing represents a new paradigm in the analysis of complex human behaviour and the processes of learning. The Standard Celeration Chart is a display of frequency against a continuous real time line (Pennypacker, Gutiérrez, & Lindsley, 2003), which transforms learning into straight-line trends and allows for calculations and projections of (ac-, de-)celeration. Celeration is thus a measure of frequency change over time, and is used to analyze performance and monitor behavioural changes in individuals. Typically, correct and incorrect responses (moves) are counted and charted separately. This produces an individual "learning picture" (Potts *et al.*, 1993) of accuracy and speed.

In this study we aimed at a further investigation of the study reported by Anzai and Simon (1975; 1979). Therefore, we examined TOH performances of two participants, similar in educational background and (former) job occupation, but substantially different in their level of executive functioning. The issues addressed were twofold.

Firstly, like Anzai and Simon (1979), we argue that exploring strategy development and transformation by using protocol analysis in persons with varying EF levels will gain information on the learning processes that underlie the acquisition of effective problem solving skills. This will contribute to a better understanding of processes involved in EF. Considering the importance of self-regulating aspects of EF, we expected participants not only to differ in (a) strategy selection and development, and (b) strategy transformation and implementation, but also in (c) effective learning rates in response to a repeated problem solving task.

Secondly, celeration indices will be considered as an alternative measure of behavioural change, i.e., learning executive performance. As Binder (1996) stated, when learners achieve certain frequencies of accurate performance they seem to retain and

maintain what they have learned, and apply, adapt or combine what they learned in new situations, which stimulates flexible, creative and self serving behaviour. We not only wish to explore the utility of celeration as an effective measure in the study of complex human behaviour, but we also want to explore the contribution of celeration indices as an indicator of individual learning abilities.

There are, a priori, several methodological limitations to our study such as sample size (e.g. two single cases) and data collection. However, we believe that analyzing corresponding verbalizations and using (novel) measurements of behavioural change, will contribute to a better understanding of EF in both clinical and healthy populations and can, in turn, provide us with necessary knowledge on learning abilities in individuals.

METHOD

Participants

Two outpatients of a Dutch psychiatric hospital participated in partial fulfilment of a standard administration of an elaborate neuropsychological assessment battery. They were matched on educational and socio-economical backgrounds, as determined by (a) years and level of education, (b) former job occupation and (c) premorbid intelligence scores assessed by a Dutch version of the National Adult Reading Task (Nelson, 1982; Schmand, Lindeboom, & van Harskamp, 1992). To optimize the exploration of differences in strategy development and implementation during performances on the Tower of Hanoi task, both participants were selected because they varied considerably in their level of executive functioning (see Table 1).

Participant A was a male, 27 years of age, with an educational history of more than 16 years. Results indicated an (above) average level of intelligence. Executive abilities showed no significant disturbances (see Table 1). He was familiar with episodic depressive and psychotic behaviour for which he was medically treated. No signs of psychotic or mood related symptoms were present during the time of the experiment.

Participant B was a female, 61 years of age with an educational history of more than 16 years. Results indicated an average level of intelligence. Executive performances, however, showed significant disabilities characterized by planning difficulties, mental inflexibility and profound sensitivity for interference of irrelevant stimuli (see Table 1). She was previously diagnosed with a depressive mood disorder. Considering the severe cognitive decline, a dementia syndrome was suspected. She did not receive any medical treatment at the time of the experiment.

Both individuals participated on a voluntary basis and did not receive financial compensation.

Materials

The Tower of Hanoi task: The Tower of Hanoi (TOH) consists of a flat board, three evenly spaced vertical iron pegs of equal height and five alabaster disks of different sizes and colour (black/white). The TOH requires participants to rearrange the disks

Table 1. Results on the (neuro-)psychological assessment battery for participant A and B.

<i>Tests</i>	<i>Domain</i>	<i>Participant A</i>	<i>Participant B</i>
		<i>Results in percentiles*</i>	
<i>Intelligence</i>			
KAIT (Kaufman Adult- and Adolescent Intelligence Test)	Fluid intellig.	98.1	57.9
NART (National Adult Reading Test)	Crystallized intellig.	50	34.5
	Pre morbid intelligence	50-75	75-90
<i>Executive function</i>			
Stroop (Stroop Colour Word Task)	Inhibition, flexibility	66	8
WCST (Wisconsin Card Sorting Test)	Flexibility, working memory	>16 (strat.use ¹) 70 (errors) 73 (pers. err. ²)	<1 (strat.use) 3 (errors) 5 (pers.err.)
ToL (Tower of London)	Planning, working memory	50-60	5-10
<i>Self-report on EF</i>			
EFI (Executive Function Index)	Self-report questionnaire (EF)	Above average	Average
DEX (Dysexecutive Behaviour Questionnaire)	Self-report questionnaire (EF)	Average	Average
<i>Personality & psychopathology</i>			
MMPI-2 (Minnesota Multiphasic Personality Inventory)	Results showed no negative test attitudes or motivational difficulties for both participants (inspection of validity scales)		

* Percentile equivalents derived from the normal probability curve, based on standard deviation units (Lezak et al., 2004, p. 143).

¹ Strategy use in solving the Wisconsin Card sorting Test.

² Number of perseverative errors made.

from the initial state to a specific goal state in a minimal number of moves (Simon, 1975). Participants were told that they could (a) only move one disk at a time, (b) could not place a disk on the table while another disk is being moved and (c) could not place a larger disk on a smaller disk (Welsh & Huizinga, 1995). Internal consistency (Cronbach's alpha) is reported to be high ($\alpha = .90$) (Bishop, Aamodt-Leeper, Creswell, McGurk, & Skuse, 2001). A translation of the Dutch problem instructions is presented in Appendix. Sessions (verbal protocols and actual moves) were recorded using a USB2.01 WebCam.

The National Adult Reading Test (NART): The NART (Nelson, 1982) is a widely accepted method for estimating premorbid levels of intelligence in neuropsychological research. Participants are asked to read out loud a list of 50 irregular spelled words. Performance on this task is relatively independent of brain damage. Correlations between WAIS-III subtest performances and NART scores are considered to be high ($r = .56-.83$). Internal consistency scores (Cronbach's alpha) range from .91 to .95 (Bouma, Mulder, & Lindeboom, 1996; Schmand *et al.*, 1992).

The Kaufman Adolescent and Adult Intelligence Test (KAIT): The KAIT is a recently developed measure of fluid and crystallized intelligence, based on the ideas of Piaget (Inhelder & Piaget, 1958; Piaget, 1972), Luria (1980), and Horn and Cattell (1996).

Fluid intelligence refers to skills such as problem solving and abstract reasoning, and is generally thought not to be influenced by cultural experience or education. Crystallized intelligence refers to skills such as reading comprehension and verbatim knowledge, and is based upon facts and rooted in (educational) experience (Mulder, Dekker, & Dekker, 2005). The Dutch adaptation of the KAIT has internal consistencies (Cronbach's alpha) ranging from .87 to .93 for the core battery subtest scales. Test-retest reliability scores range from .72 to .95 (Mulder *et al.*, 2005).

The Stroop Colour Word Task: The Stroop colour word task is thought to be a classic test of selective attention and inhibition. Participants are asked to name the colour of ink instead of reading the written (incongruent coloured) words, which generates interference and thus measures the ability to suppress a habitual response in favour of an unusual one (Stroop, 1935). Test-retest reliability scores range from .86 to .91 and are considered to be high (Bouma *et al.*, 1996).

The Wisconsin Card Sorting Task: The Wisconsin Card Sorting Task (WCST) is traditionally being used to tap mental flexibility. Participants are required to sort cards on one of three possible dimensions (colour, number, and shape) according to a non-spoken rule and then shift to sort cards along a different dimension (Heaton, Chelune, Talley, Kay, & Curtiss, 1993). Reliability scores range from .39 to .71 (Bouma *et al.*, 1996).

The Tower of London: The Tower of London (TOL) is derived from the TOH and requires planning skills and working memory capacity. Individuals must move disks from a prearranged sequence on three different pegs to match a goal state determined by the experimenter. This must be done in as few moves as possible and following a number of specific rules (Shallice, 1982). Inter correlations between different EF measures and the TOL are relatively weak, whereas reliability scores are reported to be only moderate (Bishop *et al.*, 2001; Lezak, Howieson, Loring, Hannay, & Fisher, 2004).

Executive Function Index: The Executive Function Index (EFI) is a self report questionnaire developed by Spinella (2005) to measure executive functioning in daily life. The EFI consists of 27 items, divided into five subscales, named Motivational Drive (MD), Organization (ORG), Impulse Control (IC), Empathy (EM) and Strategic Planning (SP). Moderate to adequate reliability coefficients, ranging from .41 to .69, are found by Janssen, De Mey and Egger (2009).

Dysexecutive Behaviour Questionnaire: The Dysexecutive Behaviour Questionnaire (DEX), a subtest of the Behavioural Assessment of the Dysexecutive Syndrome (BADS), is a self report questionnaire consisting of 20 items divided into four categories (Emotional or Personality change, Motivational change, Behaviour change and Cognitive change) which measures difficulties in planning and organization skills (Wilson, Alderman, Burgess, Emslie, & Evans, 1997). Adequate reliability and validity scores have been reported (Wilson *et al.*, 1997).

The Minnesota Multiphasic Personality Inventory: The Minnesota Multiphasic Personality Inventory (MMPI-2) is a widely used test of adult psychopathology. It assesses symptoms of social and personal maladjustment, assists with the diagnosis of mental disorders and the selection of appropriate treatment methods, and can also be used to determine motivation and test attitude (Graham, 2006). Internal consistencies (Cronbach's alpha) range from .32 to .85 for the clinical scales and from .65 to .82 for

the content scales. Test-retest reliability coefficients range from .43 to .88 (Derksen, De Mey, Sloore, & Hellenbosch, 2006).

Procedure

Participants were given an assessment battery consisting of two intelligence tasks (NART, KAIT) and several tests (WCST, TOL, Stroop) commonly used to assess EF (Rabin, Barr, & Burton, 2005), in order to provide an estimate of baseline intelligence and executive skills. The MMPI-2 was used to screen for possible disturbing factors such as low motivation and negative test attitude. Since the correlation of measures from neuropsychological assessments with EF in daily life is reported to be very low (Alvarez & Emory, 2006), we added two EF self report scales (EFI, DEX). This was done in order to ground the construct of EF [] in more readily observable behaviours that have real world significance.

Participants were tested individually in three sessions of approximately 2 hours. During the first session, the intelligence tasks, two self report questionnaires and the personality inventory were completed. The following session was used to gather more detailed information about executive functioning by assessing inhibition, shifting and planning abilities. After completing this pre-experimental stage, participants were informed about the global objective of the experiment and were asked to sign a declaration of consent. Each participant was able to stop and withdraw from the experiment at any time. During the third session think aloud instructions were explained. A more, general instruction was followed by two “warming-up” exercises. The think aloud instruction goes as follows:

In this experiment we are interested in what you say to yourself as you perform a problem solving task. In order to do this, we will ask you to think aloud as you work on the problems. What I mean by think aloud is that I want you to say out loud everything that you are thinking to yourself silently. Just act as if you are alone in the room speaking to yourself. If you are silent for any length of time I will remind you to keep thinking aloud. Do you understand the instructions? Before we turn to the real experiment, we will start with a couple of practice problems. I want you to think aloud while you solve these problems.

The warming-up exercises, used to familiarize the participants with the think aloud procedure, consisted of mental multiplications (e.g., 9×12) and solving anagrams (e.g., OMEH = HOME) (Ericsson & Simon, 1984). At the start of the [] experiment, participants were instructed to rearrange the Tower of Hanoi disks from the initial state to a specific goal state in a minimal number of moves (see Appendix for a Dutch translation of the problem instruction). Following Anzai and Simon (1979), the last sentence of the problem instruction was added to encourage the participants to find an adequate and efficient strategy, which presumably contains executive abilities such as inhibition, planning and shifting processes.

After one practice trial with three disks, the actual experiment started. The starting peg for the problem was peg A, and the goal peg was peg C. The disks were

numbered from 1 (smallest disk) to 5 (largest disk). The shortest solution path for the five-disk Tower of Hanoi contains 31 moves. Participants completed four trials of the Tower of Hanoi. They were given as much time as was necessary in order to solve the problem. When participants returned to the original position of the task, this was to be the start considered of a new trial. Think aloud reports, total time taken and the number of moves made to successful completion of the Tower of Hanoi were recorded using a laptop computer and a USB2.01 WebCam. After administering four trials, participants were retrospectively asked to summarize their solution strategies and to comment on the structure of the Tower of Hanoi task.

Scoring and analysis

Verbal protocols: Zelazo and colleagues (1997) have described the problem solving process in terms of a cycle, consisting of stages in which the subject must (a) define and represent the problem mentally, (b) develop a solution strategy, (c) monitor his or her progress toward a goal and (d) evaluate the solution. Based on this descriptive cycle of problem solving, verbal statements were categorized in (a) problem representation (PR), containing formulations regarding (sub)goals, constraints and operators of the TOH task, (b) planning (PL), which consisted of statements regarding monitoring, decision making and the search for alternative moves, (c) execution (EX), which consisted of statements merely announcing an (intended) move and (d) evaluation (EV), which consisted of statements regarding the control and possible adaptation of behaviour (Zelazo *et al.*, 1997). Two separate categories were added, containing explicit meta statements about the solution method (META) and irrelevant comments on the self or on the environment (REST). The obtained verbal protocols were rated and classified by one single rater (first author). To secure consistency (examining similar cognitive processes across trials), concurrent and retrospective reports were compared. Content differences, imposed by different methods of solution (more or less adequate), were analyzed following Simon (1975) who described four main problem solving strategies (e.g. perceptual/trial & error strategy, sophisticated perceptual strategy, move pattern strategy or a goal-recursion strategy). Finally, differences in reversed moves and the breaking of rules between participants were calculated.

Celeration charts: Analysis of celeration and individual learning pictures were used to evaluate the change in performances, using one minute timing periods across the four TOH trials. The primary dependent variables were frequency and celeration within, and between subjects for both correct and incorrect performances. Correct responses were defined as both optimal and non-optimal moves, leading to a satisfying (although not necessary recursive) solution method. Incorrect responses were defined as those direct erroneous moves, bringing the subject further away from a satisfying solution. Moves solely made to rectify previously made mistakes were not considered to be incorrect responses.

RESULTS

Executive performances. Mean scores are shown in Table 1. On the Kaufman crystallized intelligence scales (KAIT-CIQ), the National Adult Reading Test (NART), the Executive Function Index (EFI), the Dysexecutive Behaviour Questionnaire (DEX) as well as on the validity scales of the Minnesota Multiphasic Personality Inventory (MMPI-2), results were found to be similar for both participants. On the Kaufman fluid intelligence scales (KAIT-FIQ), the Stroop Colour Word Task (Stroop), the Tower of London task (ToL) and the Wisconsin Card Sorting Test (WCST) results indicated a markedly poorer performance for participant B, suggesting impaired executive functioning.

Protocol analyses. Verbal protocols were analyzed to provide a detailed picture of strategy adaptation and transformation during repeated problem solving behaviour. Results indicated no obvious differences in the pattern of verbal statements followed by participants A and B; over 50% (52,2% and 52,7% for participant A and participant B respectively) of the total number of statements were used to describe the actual execution of a move (EX) and over 30% of the statements were used to formulate and evaluate individual (sub)goals (PR+PL+EV 33.7% and 37% for participant A and participant B, respectively). The remaining 20% contained largely the category REST-statements (9.5% and 13.6% for participant A and B, respectively) and a small part of META-statements for participant A (1.3%). Participant B did not make any statements concerning the actual structure of the problem solving task (see Table 2 for an overview).

Although patterns did not substantially differ, the content of these statements revealed essential differences between participants concerning strategy development and transformation.

Participant A. The 232 protocol statements of this participant can be divided into four solution attempts (A-D), containing a total of 85 (A; S1-S85), 54 (B; S86-S139), 48 (C; S140-S187) and 45 (D; S188-S232) verbal statements, respectively. Of these statements, 17 statements fall into the category of problem representation (PR), 36 statements into the category of planning (PL), 121 statements are bare statements to announce the execution of a move (EX) and 33 statements contain verbal evaluative comments (EV). The remaining 25 statements consist of either meta statements about

Table 2. Verbal statements participant A and B.

Verbal Statements	Participant A		Participant B	
	Number of statements	Percentages (%)	Number of statements	Percentages (%)
PR	17	7.3	25	6.7
PL	36	15.5	59	15.6
EX	121	52.2	198	52.7
EV	33	14.2	43	11.4
META	3	1.3	0	0
REST	22	9.5	51	13.6
Total	232	100	376	100

the solution method (META; 3 statements) or irrelevant comments about the self or the environment (REST; 22 statements). See Table 2 for an overview.

When we look at the actual content of the verbal statements, participant A showed a gradual progress of employing more efficient strategies with increasing task familiarity; each of the strategies used evolved from the one just preceding. Starting with trial 1, strategy use can be referred to as an active search of the problem space and the task restrictions. Conduct was characterized by trial and error learning, with some occasions of (sub)goal setting (“to move the largest disk, I should first get these smaller ones out of the way.”). Participant A determined what to do next by perceiving features of the actual problem situations, indicating a (sophisticated) perceptual approach of the problem (Simon 1975; Simon & Anzai, 1979). Statement 69 (“I am trying to rebuild this ‘tower of three’ of disks on the left peg.”) furthermore suggested that participant A was already using part of the recursive rules or strategies (deductive reasoning) that clearly emerged during trial 3 and 4.

During the second trial, participant A guided his performances by explicitly formulating (and repeating) a sequence of “hierarchical” set (sub)goals. Participant A identified the largest disk that was not yet on the target peg, identified the largest disk that was obstructing its move to the target peg, established the goal of moving the obstructer to the alternative peg (neither source or target peg) and repeated this ‘cycle’ of moves until his final goal (solving the problem) was reached. Errors were independently corrected. This second strategy resembles that of what Anzai and Simon (1975; 1979) referred to as a (perceptual based) ‘goal-peg’ or ‘move pattern’ strategy.

However, the most deliberate changes in strategy use took place between the second and third trial. During trial 3 participant A explicitly formulated and applied the recursive rule (to achieve a goal [Disk 5 on peg C], I need to achieve a subgoal [Disks 1-4 on peg B]; to achieve a subgoal [Disk 4 on peg B] I need to achieve another subgoal [Disks 1-3 on peg C]; and so on) that guided his performances on the five disk Tower of Hanoi problem (Anzai & Simon, 1979; Simon, 1975). A clear example of the application of this recursive rule can be found in several statements (e.g., “my goal is to transfer disk four to the middle peg. Therefore I should move these three smaller ones to another peg”; “I should replace each ‘tower’ of disks in order to create space for a larger disk”; “what I need is a so called ‘continuous’ strategy to solve this problem.. this is a repeating process.”). Significant changes of problem solving behaviour during trial 3 were seen in the description of “sets of disks” (versus single disks), which indicated knowledge about the pyramid subgoal structure of the Tower of Hanoi. Trial 4, which was completed in a minimal number of moves (e.g. 31), was characterized by a further refinement of the goal recursion- strategy (Simon, 1975).

Finally, retrospective reports of participant A indicated sophisticated understanding of the structure of the problem solving task; he described the recognition of moving sub-pyramids to reach the final goal state efficiently.

Participant B. The 376 protocol statements of participant B can be divided into four solution attempts (A-D), containing a total of 162 (A; S1-S162), 61 (B; S163-S223), 63 (C; S224-S286) and 90 (D; S287-S376) verbal statements, respectively. Of the 376 protocol statements, 25 statements fall into the category of problem representation (PR),

59 statements into the category of planning (PL), 198 statements are bare statements to announce the execution of a move (EX) and 43 statements contain verbal evaluative comments (EV). The remaining 51 statements consisted of irrelevant comments about the self or the environment (REST) (see Table 2 for an overview).

When we look at the content of the verbal statements, we see that participant B showed some progress in adapting an efficient problem solving strategy (strategies evolved from the one just preceded). However, she seemed unable to retain information in working memory long enough to implement and adjust behaviour across four trials.

During trial 1 and 2, her behaviour was characterized by a rather impulsive and automatic start; at nine occasions (seven times during trial 1 and two times during trial 2) participant B reversed a move just made or moved the same disk twice in succession. Rules (moving two disks simultaneously) were broken at several times during trial 1, but not during trial 2. Strategy may be described as a simple perceptual, primarily stimulus driven, search strategy (Anzai & Simon, 1979; Simon, 1975). Although verbal statements consisted of problem representations (formulating restrictions and rules, e.g., "I should transfer the largest peg to the utmost right peg"; "I cannot place a larger disk on a smaller disk.") and planning behaviour (formulating subgoals and defining the problem state, e.g., "in order to move the largest disk, first I need to clear the target peg by moving these disks."), participant B was unable to efficiently follow these self formulated strategies, producing a repeated cycle of (ineffective) moves. During the third trial participant B did identify the largest disk, the obstructer of this disk and established the goal of moving the obstructer to the alternative peg, but noticed that she was moving the disks back and forth. After persevering a little longer, she decided to return to the original configuration, ending trial 3 prematurely.

Trial 4 was, like the first two trials, characterized by trial and error responding. However, by repeatedly recalling previously described rules and strategies participant B was able to gradually monitor her behaviour; she did generate some adequate rules and strategies (goal setting and evaluating), but remained unsuccessful in implementing these strategies adequately (e.g., "to move disk four, I should replace disk three...disk three can only be moved when disk two is replaces...where can I put this middle disk?").

Inspection of the retrospective report indicated no recognition of the pyramid (or hierarchical) structure of the problem solving task.

Overall learning effects. Participant A solved trial 1 to 4 of the five disk Tower of Hanoi successfully in 81, 55, 38 and 31 (optimal solution) moves, respectively (see Figure 1) in an average time of 3.16 minutes per trial. Participant B solved trial 1, 2 and 4 of the five disk Tower of Hanoi successfully in 85, 53 and 48 moves, respectively (see Figure 1) in an average time of 7.36 minutes per trial. Trial 3 was ended by participant B after 37 moves. Results indicated an overall learning effect for both participants, showing decreasing trends for both total time taken and total number of (correct and incorrect) moves per trial.

Frequency. Overall performances of participant A showed an accelerating trend for the number of correct responses per minute and a slightly decreasing trend for incorrect number of responses per minute (see Table 3). Frequencies of correct number of actions per minute increased from an average of 5.42 per minute (trial 1) to an average

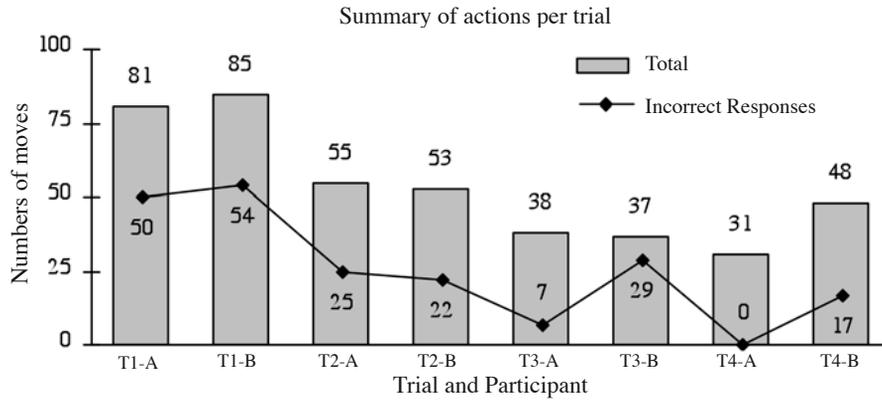


Figure 1. Summary of actions per trial for participant A and B.

of 13.30 per minute (trial 4). Number of incorrect responses decreased from an average of 4.72 per minute (trial 1) to zero over the course of four trials.

Results across trials for participant B showed an accelerating trend for the number of correct responses per minute and a maintaining trend for the number of incorrect actions per minute (see Table 3). Frequencies of correct number of moves increased from 2.01 per minute (trial 1) to 6.16 per minute (trial 4). Incorrect responses varied from an average of 1.75 (trial 1) to an average of 2.93 (trial 2) per minute. Accuracy ratios (e.g. frequency correct responses / frequency incorrect responses) were determined in order to measure the quality of behaviour (see Table 3), and are illustrated in Figure 2-5. Results indicate an increase of speed and accuracy (overall quality of behaviour increasing) for participant A, whereas participant B shows slightly increasing speed, but no (significant) change in accuracy.

Celeration. To determine the overall change of behaviour over the four trials (learning rate), celeration lines were drawn using the Quarter Intersect Method (Pennypacker *et al.*, 2003). Numerical values were added using the Celeration Finder as described by Pennypacker and colleagues (2003). For a detailed illustration see Figure 2-5. Results indicate a steady increase of correct responses for participant A, showing an overall (ac)celeration of X 1.4 across trials. Number of errors increased, maintained

Table 3. Frequencies per minute of correct and incorrect responses over trial 1-4.

Trial	Participant A			Participant B		
	Correct	Incorrect	Accuracy ratio	Correct	Incorrect	Accuracy ratio
1	5.42	4.72	X1.15	2.01	1.75	X1.15
2	10.80	4.53	X2.38	7.56	2.93	X2.58
3	14.62	1.89	X7.74	1.38*	2.07*	+0.67*
4	13.30	0.00	X13.30	6.16	1.79	X3.44
Celeration	CL correct	CL incorrect	AIM	CL correct	CL incorrect	AIM
	X1.4	+1.4	X2.0	X1.2	X1.0	X1.2

* Participant B ended trial 3 prematurely after 37 actions

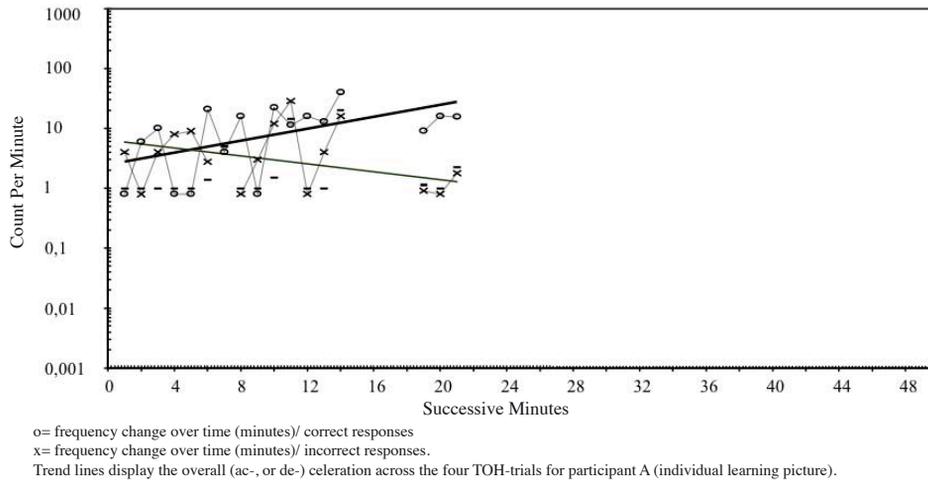


Figure 2. Overall (ac-, de-)celeration measures for participant A.

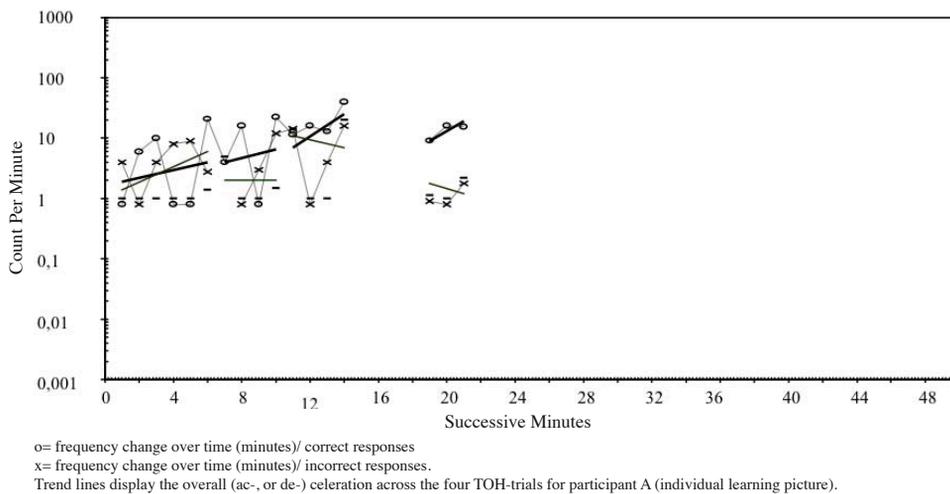


Figure 3. (ac-, de-) Celeration measures per trial (1-4) for participant A.

and decreased during trials 1 to 4, showing an overall (de)celeration of $\div 1.4$ (see Table 3 and Figure 2). Participant B showed a slight overall (ac)celeration of number of correct responses (X 1.2) across trials 1, 2 and 4. Trial 3 was ended before successful completion. However, overall number of errors maintained at an X 1.0 level (see Table 3 and Figure 4).

To compare behavioural change over trials, Accuracy Improvement Measures (AIMs), applications of celeration multipliers using concurrent correct and incorrect celerations (Pennypacker *et al.*, 2003; Koorland & MacLeod 2005), were calculated (see Table 3). Comparing AIMs across trials for participant A showed an overall learning picture of increasing speed and accuracy. Although speed increased during trial 1 to 4,

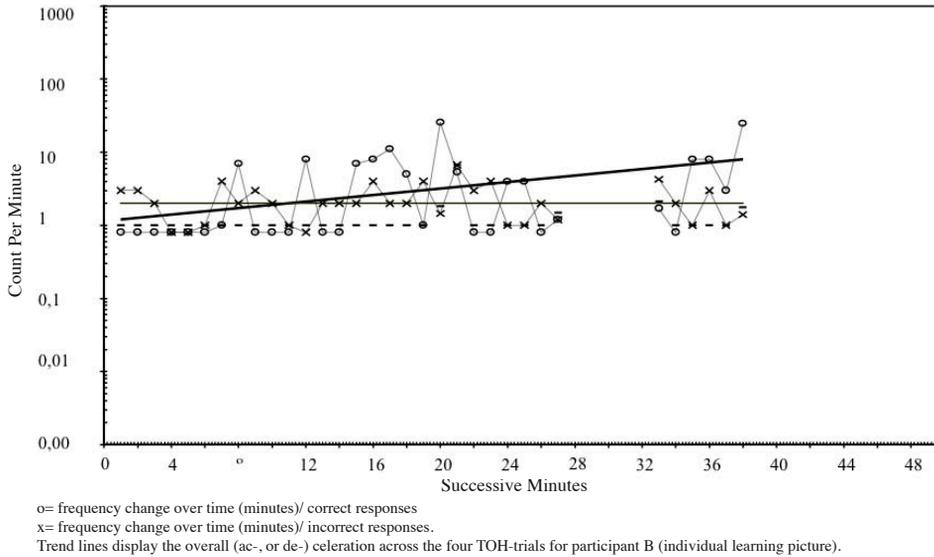


Figure 4. Overall (ac-, de-) celeration measures for participant B.

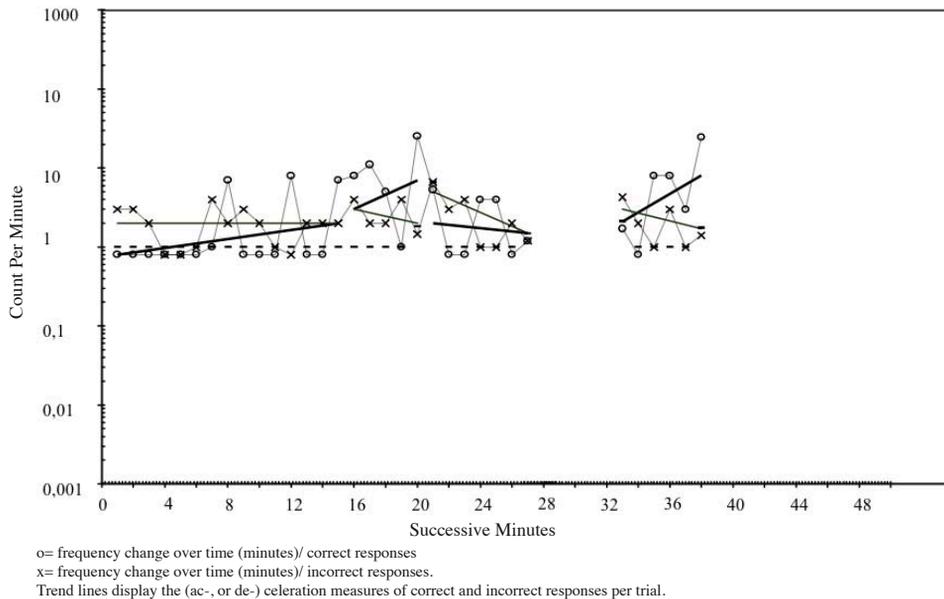


Figure 5. (ac-, de-) Celeration measures per trial (1-4) for participant B.

accuracy (errors) decreased during trial 1 and maintained during trial 2. The third and fourth trials did show both a significant acceleration of speed and accuracy (see Figure 2 & 3) Participant B, however, showed an overall learning picture of increasing speed, but no significant change in accuracy levels (see Figure 4 & 5).

In summary, results indicate an overall change of behaviour (accelerating speed and accuracy) for participant A. Although participant B did “learn by doing” (Anzai & Simon, 1979), level of behaviour change was markedly poorer when compared to participant A.

DISCUSSION

The main purpose of this study was to replicate (and refine) results on problem solving behaviour as previously described by Anzai and Simon (1979). We aimed (a) at gaining a better understanding of the different mechanisms underlying performances of the Tower of Hanoi, using an in depth analysis of the verbal statements gathered from participants varying in their EF level, and (b) at considering the contribution or utility of celeration charts and learning rates as an efficient measure of these individual executive performances.

Addressing the first issue, several remarks can be made. First, results confirm the notion that several executive subfunctions, such as updating, working memory and inhibition, are important mediators to success in solving the Tower of Hanoi problem (Welsh & Huizinga, 1995). Furthermore, results are convergent with the notion that, as stated by Zook and colleagues (2004), fluid intelligence (KAIT-FIQ) is an important contributor to executive functioning and novel problem solving ability. Duncan, Burgess, and Emslie (1995) asserted that fluid intelligence tasks rely on the integrity of the frontal lobes and underscore the utility of fluid intelligence as a measure of EF.

Secondly, in depth analysis of the recorded verbal statements showed important differences concerning strategy development and transformation between participants. Participant A showed a gradual progress of employing more efficient strategies with increasing task familiarity, indicating an sophisticated understanding of the sub pyramid (recursive) goal structure of the tower of Hanoi. However, strategy use of participant B may be described as a simple perceptual search strategy, characterized by rather impulsive and automatic behaviour. Although she did formulate some rules and strategies, she seemed unable to inhibit prepotent responses (repeating moves and rule breaking) or retain information long enough in working memory to successfully implement or adjust previously learned behaviour across trials. Participants basically differed in their (a) avoidance of repetitive or recurrent moves, (b) ability of sequencing disks or moves and (c) evaluative skills needed in order to modify and adapt their behaviour and create opportunities for new strategy solutions (Anzai and Simon, 1979).

Third, results indicated an overall learning effect for both participants. However, examination of frequency and celeration scores revealed a marked increase of speed as well as accuracy level (e.g., the overall quality of behaviour) for participant A, whereas performance of participant B showed slightly increasing speed, but no (significant) change in behaviour accuracy. Prominent changes in problem solving behaviour of participant A took place between the second and third trial, whereas performances of participant B showed a much more gradual (and random) increase of correct moves.

In summary, although both our participants did “learn by doing,” involving task

independent knowledge (availability of rules and strategies) and task dependent knowledge (efficient transformation or implementation of these rules and strategies), they varied considerably in their degree of awareness (Anzai and Simon, 1979). Results suggested that although both participants formulated rules, indicating the presence of strategies or schemas, and acquired (evaluative) knowledge about the effectiveness of prior or intended moves, they differed substantially in adequately adapting or implementing these rules to modify their executive behaviour in a novel or non-habituated situation. In particular, participant B seemed unable to efficiently process prior information gathered from the immediate preceding TOH trials, which, as Simon (1975) stated, plays a major role in generating and transforming to new (effective) problem solving strategies. This indicates that problem solving processes not only involve the availability of efficient rules or strategies, but also involve the flexible implementation or adaptation of these rules and strategies in different contextual frameworks, stipulating the controlling or self regulating aspects of EF.

As to the second issue, the present paper does provide us with some preliminary evidence on the utility of Standard Celeration Charts as an effective assessment tool in the study of complex human behaviour. Through their presentation of both frequency and celeration in standard graphic and quantitative units, the charts differentiated between individual performance levels and growth in learning across time. They provided us with information on individual change, which was consistent with findings from protocol analysis. Furthermore, additional information on the ability to learn in novel or conflict situations, represented by both accuracy levels and speed of performing, was obtained.

As stated by Binder (1996), competent or adaptive executive performance is not only characterized by the use of efficient rules or strategies (verbal statements), but is also dependent on the fluid combination of both accuracy and speed. By using celeration as an (additionally) assessment method of executive functioning, we can not only visualize differences between individual learning abilities, but also monitor the effectiveness of interventions when solving a novel and complex problem situation. We therefore argue that the use of Standard Celeration Charts, after further refinement, improves and contributes to the detailed analysis of human complex behaviour, which, in time, could offer a broader window on performances and trainability of EF.

Several limitations of the present pilot study need to be addressed. There have long been well argued objections against the use of verbal reports as data. These concern the accessibility of unconscious behaviour, the accuracy and completeness of verbal reports and whether or not think aloud protocols will influence task performance (Hayes, White, *et al.* 1998). Moreover, although research on protocol analysis is available, the number of empirical studies is quite small, and there has been no systematic attempt to analyze verbal behaviour (Cabello & O'Hara, 2002). However, recent studies did provide strong evidence that concurrent verbalization reflects the actual thought processes that occur during problem solving, indicating protocol analysis to be a research methodology which bears promises in the study of human complex behaviour (Cabello *et al.*, 2004; Hayes, White, & Bisett, 1998; Wulfert *et al.*, 1991).

Secondly, difficulties in single-case research arise in evaluating which precise

factor(s) contribute to (significant) behaviour change and how these factors could represent performances of a specific population (Kazdin, 1981). Although the generality of our results is indeed questionable, our main objective was to obtain detailed information on specific variables that warrant further investigation.

Third, participants in this study were not matched on age or gender, but were primarily chosen on their (varying) executive abilities. It is unclear to what extent these factors, could have influenced the results. Furthermore, by recording their performances while solving the Tower of Hanoi, participants could have experienced feelings of discomfort or tension, which, in turn could have altered their behaviour. Finally, it is unclear to what extent medical interventions influenced internal motivation and/or test attitude.

Our main purpose of this article was to contribute to a general understanding of executive (dys)functioning, leading to a more adequate conceptualization on the nature of EF and contributing to the development of effective assessment and intervention methods for dysexecutive symptoms. In order to form new and more effective rules or strategies, people extract knowledge about problem structures from previous solution attempts, needing the abilities to update and monitor working memory representations, shift between tasks or mental sets and inhibit dominant or prepotent responses. These are all examples of fundamental executive skills that mediate complex human behaviour (Miyake *et al.*, 2000). Our results indicate, that although repetition on problem solving tasks leads to general improvement of rules and strategy use, participants differ substantially in their ability to learn.

Future research should consider these differences in learning abilities on problem solving behaviour, and interventions should be based on a precise assessment of (executive) dysfunctional behaviour of individuals within the context of their living space. The question is not only to what extent rules are available, but to what extent these rules can be transformed and implemented. Is an individual able to generate, apply or adapt these rules or strategies in order to solve a new problem within a specific context? Research should focus on training and extending fluency in acquiring these rules and strategies in order to influence participants in their (more efficient) choice of strategies. This type of training-testing methods constitute a dynamic assessment procedure from which the implications for treatment can be more directly derived (Egger *et al.*, 2007).

REFERENCES

- Alvarez JA & Emory E (2006). Executive function and the Frontal Lobes: A meta-analytic review. *Neuropsychology Review*, *16*, 17-42.
- Anzai Y & Simon HA (1979). The Theory of Learning by Doing. *Psychological Review*, *86*, 124-140.
- Barkley RA (2001). The executive functions and self-regulation: An evolutionary neuropsychological perspective. *Neuropsychology Review*, *11*, 1-29.

- Binder C (1996). Behavioral fluency: Evolution of a new paradigm. *The Behavior Analyst*, 19, 163-197.
- Bishop DVM, Aamodt-Leeper G, Creswell C, McGurk R, & Skuse DH (2001). Individual differences in cognitive planning on the Tower of Hanoi task: Neuropsychological maturity or measurement error? *Journal of Child Psychology and Psychiatry*, 42, 551-556.
- Bouma A, Mulder J & Lindeboom J (1996). *Neuropsychologische Diagnostiek: Handboek*. Lisse: Swets & Zeitlinger.
- Brokate B, Hildebrandt H, Eling P, Fichtner H, Runge K, & Timm C (2003). Frontal lobe dysfunctions in Korsakoff's syndrome and chronic alcoholism: continuity or discontinuity? *Neuropsychology*, 17, 420-428.
- Cabello F, Luciano C, Gomez I, & Barnes-Holmes D (2004). Human schedule performance, protocol analysis and the 'silent dog' methodology. *The Psychological Record*, 54, 405-422.
- Cabello F & O'Hara D (2002). Addressing the limitations of protocol analysis in the study of complex human behavior. *International Journal of Psychology and Psychological Therapy*, 2, 115-130.
- Cicerone K, Levin HS, Malec J, Stuss D, & Whyte J (2006). Cognitive rehabilitation interventions for executive function: Moving from bench to bedside in patients with traumatic brain injury. *Journal of Cognitive Neuroscience*, 18, 1212-1222.
- Derksen JLL, De Mey HRA, Sloore H, & Hellenbosch G (2006). *Minnesota Multiphasic Personality Inventory-2: Handleiding voor afname, scoring en interpretatie*. Nijmegen: PEN Psychodiagnosics.
- Duncan J, Burgess P, & Emslie H (1995). Fluid intelligence after frontal lobe lesions. *Neuropsychologia*, 3, 261-268.
- Egger J, De Mey H, & Janssen G (2007). Assessment of executive functioning in psychiatric disorders: functional diagnosis as the overture of treatment. *Clinical Neuropsychiatry*, 4, 111-116.
- Ericsson KA & Simon HA (1984). Techniques of protocol analysis: Examples. In KA Ericsson & HA Simon (Eds.) *Protocol analysis: Verbal reports as data* (pp. 313-378). Cambridge: The MIT Press.
- Graham JR (2006). *MMPI-2: Assessing personality and psychopathology*. Oxford: University Press.
- Graham NL, Emery T, & Hodges JR (2004). Distinctive cognitive profiles in Alzheimer's disease and subcortical vascular dementia. *Journal of Neurology Neurosurgery and Psychiatry*, 75, 61-71.
- Handley SJ, Capon A, Copp C, & Harper C (2002). Conditional reasoning and the Tower of Hanoi: The role of spatial and verbal working memory. *British Journal of Psychology*, 93, 501-518.
- Happe F, Booth R, Charlton R, & Hughes C (2006). Executive function deficits in autism spectrum disorders and attention-deficit/hyperactivity disorder: Examining profiles across domains and ages. *Brain and Cognition*, 61, 25-39.
- Hayes SC, Gifford EV, & Ruckstuhl Jr LE (1996). Relational frame theory and executive function: A behavioural approach. In Lyon GR & NA Krasnegor (Eds.) *Attention, Memory and Executive function* (pp. 279-305). Baltimore: Paul Brooks Publisher.
- Hayes SC, White D & Bisett RT (1998). Protocol analysis and the 'silent dog' method of analyzing the impact of self-generated rules. *The Analysis of Verbal Behavior*, 15, 57-63.
- Heaton RK, Chelune JL, Talley G, Kay G, & Curtiss C (1993). *Wisconsin Card Sorting Test Manual: Revised and expanded*. Odessa, FL: Psychological Assessment Resources.
- Hill EL & Frith U (2003). Understanding autism: Insights from mind and brain. In U Frith & EL Hill (Eds.) *Autism: Mind and Brain* (pp. 1-20). Oxford: University Press.

- Horn JL & Catell RB (1996). Refinement and test of the theory of fluid and crystallized general intelligence. *Journal of Educational Psychology*, *57*, 253-270.
- Hughes C, Russell J, & Robbins TW (1994). Evidence for executive dysfunction in autism. *Neuropsychologia*, *32*, 477-92.
- Inhelder B & Piaget J (1958). *The growth of logical thinking from childhood to adolescence*. New York: Basic Books.
- Janssen G, De Mey H, & Egger J (2009). Executive functioning in college students: Evaluation of the Dutch Executive Function Index (EFI-NL). *International Journal of Neuroscience*, *119*, 792-805.
- Kazdin AE (1981). External validity and single-case experimentation: Issues and limitations (a response to JS Birnbrauer). *Analysis and Intervention in Developmental Disabilities*, *1*, 133-143.
- Koorland MA & MacLeod S (2005). The accuracy improvement measure: A tool for assessing the effectiveness of teacher preparation programs. *Journal of Precision Teaching and Celeration*, *21*, 13-18.
- Kuelz AK, Hohagen F, & Voderholzer U (2004). Neuropsychological performance in obsessive-compulsive disorder: A critical review. *Biological Psychology*, *65*, 185-236.
- Lezak MD, Howieson DB, Loring DW, Hannay HJ & Fisher JS (2004). *Neuropsychological assessment*. Oxford: University Press.
- Luria AR (1980). *Higher Cortical Functions in Man*. New York: Basic Books.
- Miller JN & Ozonoff S (2000). The external validity of Asperger disorder: lack of evidence from the domain of neuropsychology. *Journal of Abnormal Psychology*, *109*, 227-38.
- Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, & Wager TD (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49-100.
- Mulder JL, Dekker R, & Dekker PH (2005). *Nederlandse bewerking van de Kaufman-adolescent and adult intelligence test (KAIT)*. Leiden: PITS.
- Nelson HE (1982). *National Adult Reading Test Manual*. London: National Hospital.
- Pennypacker HS, Gutiérrez A, & Lindsley OR (2003). *Handbook of the Standard Celeration Chart*. Cambridge: Concord.
- Piaget J (1972). Intellectual evolution from adolescence to adulthood. *Human Development*, *15*, 1-12.
- Potts L, Eshleman JW, & Cooper JO (1993). Ogden R Lindsley and the historical development of precision teaching. *The Behavior Analyst*, *16*, 177-189.
- Rabin LA, Barr WB, & Burton LA (2005). Assessment practices of clinical neuropsychologists in the United States and Canada: A survey of INS, NAN and APA Divisions 40 members. *Archives of Clinical Neuropsychology*, *20*, 33-65.
- Royall DR, Lauterbach EC, Cummings JL, Reeve A, Rummans TA, Kaufer DI, LaFrance Jr WC, & Coffey CE (2002). 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. *Journal of Neuropsychiatry and Clinical Neuroscience*, *14*, 377-405.
- Schmand B, Lindeboom L, & van Harskamp F (1992). *De Nederlandse leestest voor volwassenen: Handleiding*. Lisse: Swets & Zeitlinger.
- Sergeant JA, Geurts H, & Oosterlaan J (2002). How specific is a deficit of executive functioning for attention-deficit/hyperactivity disorder? *Behavioural Brain Research*, *130*, 3-28.
- Shallice T (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London*, *298*, 199-209.

- Simon HA (1975). The functional equivalence of problem solving skills. *Cognitive Psychology*, 7, 268-288.
- Spinella M (2005). Self-rated executive function: Development of the Executive Function Index. *International Journal of Neuroscience*, 115, 649-667.
- Stroop JR (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643-662.
- Voglmaier MM, Seidman LJ, Niznikiewicz MA, Dickey CC, Shenton ME, & McCarley RW (2000). Verbal and nonverbal neuropsychological test performance in subjects with schizotypal personality disorder. *American Journal of Psychiatry*, 157, 787-93.
- Welsh MC & Huizinga M (2005). Tower of Hanoi disk-transfer task: influences of strategy knowledge and learning on performance. *Learning and Individual Differences*, 15, 283-298.
- Welsh MC, Satterlee-Cartmell T, & Stine M (1999). Towers of Hanoi and London: Contribution of working memory and inhibition to performance. *Brain and Cognition*, 41, 231-242.
- Williams DJ & Noyes JM (2007). Effect of experience and mode of presentation on problem solving. *Computers in Human Behavior*, 23, 258-274.
- Wilson BA, Alderman N, Burgess PW, Emslie H, & Evans JJ (1997). *Behavioural assessment of the dysexecutive syndrome (BADS): Handleiding*. Bury St. Edmunds: Thames Valley Test Company.
- Wulfert E, Dougher MJ, & Greenway DE (1991). Protocol analysis of the correspondence of verbal behavior and equivalence class formation. *Journal of the Experimental analysis of Behavior*, 56, 489-504.
- Zelazo PD, Carter A, Reznick JS, & Frye D (1997). Early development of executive function: a problem-solving framework. *Review of General Psychology*, 1, 198-226.
- Zook NA, Davalos DB, DeLosh EL, & Davis HP (2004). Working memory, inhibition and fluid intelligence as predictors of performance on the Tower of Hanoi and London tasks. *Brain and Cognition*, 56, 286-292.

Received, November 19, 2008

Final acceptance, November 30, 2009

APPENDIX*Translation of the Dutch instructions for the five disk Tower of Hanoi problem task*

There are three iron pegs on the board, which are evenly spaced and of equal height. These pegs are named peg A, peg B and peg C from the left to the right. Furthermore, there are five disks of different size and colour (e.g. black and white) on peg A. These disks are arranged in such a way that each disk lies on top of the disks that are bigger than it is.

Your task is to transfer all of those disks in order to create a similar pyramid-shaped configuration of disks on peg C. Doing this, you should consider the following rules:

- a) you can only move one disk at a time
- b) disks should at all times be placed on another peg (do not put a disk on the table or do not hold one in your hand while moving another disk)
- c) you can not place a larger disk on a smaller disk

Try to find a good and well fit solution procedure for the problem presented to you.

Good luck!