Brain and Cognition 71 (2009) 437–451

Contents lists available at ScienceDirect

Brain and Cognition

journal homepage: www.elsevier.com/locate/b&c

The Wisconsin Card Sorting Test and the cognitive assessment of prefrontal executive functions: A critical update

Erika Nyhus [a], Francisco Barceló [b,*,1]

[a] Department of Psychology, University of Colorado at Boulder, CO, USA

[b] Clinical Neuropsychology, Institut Universitari d'Investigació en Ciències de la Salut (IUNICS), Universitat de les Illes Balears, Spain

ARTICLE INFO

Article history:
Accepted 16 March 2009
Available online 17 April 2009

Keywords:
Cognitive control
Executive functions
Frontal lobes
Information theory
Neuropsychological assessment
Task-switching
Wisconsin Card Sorting Test

ABSTRACT

For over four decades the Wisconsin Card Sorting Test (WCST) has been one of the most distinctive tests of prefrontal function. Clinical research and recent brain imaging have brought into question the validity and specificity of this test as a marker of frontal dysfunction. Clinical studies with neurological patients have confirmed that, in its traditional form, the WCST fails to discriminate between frontal and non-frontal lesions. In addition, functional brain imaging studies show rapid and widespread activation across frontal and non-frontal brain regions during WCST performance. These studies suggest that the concept of an anatomically *pure* test of prefrontal function is not only empirically unattainable, but also theoretically inaccurate. The aim of the present review is to examine the causes of these criticisms and to resolve them by incorporating new methodological and conceptual advances in order to improve the construct validity of WCST scores and their relationship to prefrontal executive functions. We conclude that these objectives can be achieved by drawing on theory-guided experimental design, and on precise spatial and temporal sampling of brain activity, and then exemplify this using an integrative model of prefrontal function [i.e., Miller, E. K. (2000). The prefrontal cortex and cognitive control. *Nature Reviews Neuroscience*, *1*, 59–65.] combined with the formal information theoretical approach to cognitive control [Koechlin, E., & Summerfield, C. (2007). An information theoretical approach to prefrontal executive function. *Trends in Cognitive Sciences*, *11*, 229–235.].

1. Introduction

In the tradition of testing thinking processes or mental set, in 1900, Ach developed the sorting task in which subjects had to sort cards with non-sense words based on common features shared by the objects the words represented. Later, in 1920, Goldstein reported the use of sorting tasks to test concrete and abstract attitudes in brain-damaged patients. Following Ach and Goldstein, the Wisconsin Card Sorting Test (WCST) was devised in 1948 by Grant and Berg as an index of abstract reasoning, concept formation, and response strategies to changing contextual contingencies (Eling, Derckx, & Maes, 2008). Years later, Milner, a neuropsychologist at the Neurological Institute of Montreal, introduced the WCST to assess prefrontal lobe dysfunction in patients with brain lesions (Milner, 1963). Currently, there are at least two different systems of administration and scoring of the WCST; the standard version by Grant and Berg (1948) with Milneŕs (1963) correction criteria and the shortened version by Heaton (Heaton, 1981; Heaton, Chelune, Talley, Kay, & Curtis, 1993). Furthermore, the test has been administered in modified versions by Nelson (1976), Delis, Squire, Bihrle, and Massman (1992), and Barceló (1999, 2003).

In its conventional form (Heaton, 1981; Heaton et al., 1993), the WCST consists of four key cards and 128 response cards with geometric figures that vary according to three perceptual dimensions (color, form, or number). The task requires subjects to find the correct classification principle by trial and error and examiner feedback. Once the subject chooses the correct rule they must maintain this sorting principle (or *set*) across changing stimulus conditions while ignoring the other – now irrelevant – stimulus dimensions. After ten consecutive correct matches, the classification principle changes without warning, demanding a flexible shift in set. The WCST is not timed and sorting continues until all cards are sorted or a maximum of six correct sorting criteria have been reached. Despite the fact that Heaton's correction norms offer sixteen different scores, due to the internal structure of the test, many authors normally rely on no more than two or three scores as an index of subject's performance, including: number of categories completed, number of perseverative errors, and number of non-perseverative errors (Barceló & Knight, 2002; Bowden et al., 1998; Greve, 1993; Greve, Bianchini, Hartley, & Adams, 1999; Greve et al., 2002).

* Corresponding author. Fax: +34 971 172309.
E-mail address: f.barcelo@uib.es (F. Barceló).

[1] Present address: Ed. Beatriu de Pinòs #12, Universitat de les Illes Balears, Cra. Valldemossa km 7, 5, 07122 Palma de Mallorca.

0278-2626/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.bandc.2009.03.005

Several classic studies reported the sensitivity of the WCST to frontal lobe lesions (Drewe, 1974; Milner, 1963; Nelson, 1976; Robinson, Heaton, Lehman, & Stilson, 1980; Teuber, Battersby, & Bender, 1951). Many authors have later questioned the sensitivity and specificity of the WCST to frontal lobe lesion or dysfunction in neurological or psychiatric patients, respectively. Consequently, some handbooks of neuropsychological assessment advise about the risk of using WCST scores as a direct marker of frontal lobe damage without other converging evidence (Lezak, Howieson, & Loring, 2004; Strauss, Sherman, & Spreen, 2006). For example, the American Standardization warns about the use of the test as an anatomical marker of brain dysfunction (Axelrod et al., 1996).

In the following two sections we review two sources of evidence that have led to the present state of affairs. The first source is research with brain-damaged patients. Many studies have shown that damage in areas other than the frontal cortex significantly affects WCST performance. Table 1 provides details of the lesion location, paradigm, results, and conclusions for the results from the studies of brain-damaged patients reviewed below. The second source is functional neuroimaging of healthy subjects during WCST performance. These studies reveal activation in a widespread neural network of prefrontal, frontal, temporal, parieto-temporal, and parieto-occipital cortical regions during various stages of WCST performance. Table 2 provides details of the experimental design, brain imaging methods, results, and conclusions from the functional neuroimaging studies reviewed below. Although these results suggest that the WCST is not specific to frontal lobe function new experimental designs and methodologies, together with modern formal models of prefrontal executive functions (Koechlin & Summerfield, 2007; Miller, 2000) have provided new tools for understanding the cognitive processes and brain locations involved in the various component operations involved in WCST performance. For instance, it has recently been proposed that the WCST comprises task-switching demands associated with the reception of disconfirming feedback (Barceló, Escera, Corral, & Periañez, 2006; Barceló & Knight, 2002) which could be considered a component operation specific to prefrontal lobe function and can be measured more precisely using simplified task-switching paradigms (cf. Rubinstein, Meyer, & Evans, 2001; Shallice, Stuss, Picton, Alexander, & Gillingham, 2008).

2. Review of clinical studies

Milner's study, reported in the *Archives of Neurology* (1963), found that eighteen patients with epileptogenic foci in the dorsolateral prefrontal cortex (dPFC) committed more perseverative errors than patients with orbitofrontal cortex (oPFC), temporal, or parietal foci. The non-perseverative error score did not yield significant differences across clinical groups. Milner linked the fewer number of achieved categories in dPFC patients to their perseverative tendencies rather than to their tendency to being distracted (i.e., to non-perseverative errors). For years to come, these seminal findings and interpretations established the expected pattern of neuropsychological performance for patients with prefrontal lesions and, in particular, for patients with dPFC lesions. More than 40 years after Milner's original report (1963) we can verify the enormous impact her conclusions had for research and theorizing on prefrontal functions. Her conclusions influenced the interpretation of earlier and later studies. For example, the results by Teuber et al. (1951) were largely overlooked due to their poor correlation to Milner's interpretation. In addition, later studies adopted Milner's conclusions to interpret new results according to her seminal work. However, in many cases the correlation between studies is not complete. A group of prefrontal patients examined by Drewe (1974) achieved fewer categories and scored more perseverative errors than a group of patients with non-frontal lesions. However, it was the subgroup of patients with lesions in medial prefrontal cortex (mPFC), not dPFC patients, that showed the worst impairment in the number of categories achieved. In addition, there was large variability in the behavioral measures, making it difficult to classify individual cases into well-defined clinical groups. Debate over the location of WCST function has not only focused on specific areas of the frontal lobes. Teuber et al. (1951) carried out one of the first studies whose data argued against the specificity of the WCST as a test of frontal lobe function. These authors observed a larger number of total errors in subjects with lesions in posterior rather than frontal areas.

Recently, many clinical studies of WCST performance report impairment on the WCST with frontal cortex damage (Demakis, 2003; Freedman, Black, Ebert, & Binns, 1998; Giovagnoli, 2001; Goldstein, Obrzut, John, Ledakis, & Armstrong, 2004; Igarashi et al., 2002; Leskela et al., 1999; Mukhopadhyay et al., 2008; Nelson, 1976; Robinson et al., 1980; Stuss et al., 2000). Although there have been reports of left frontal damage affecting WCST performance more than right frontal damage (Goldstein et al., 2004), others report no difference in laterality of damage in the frontal cortex (Demakis, 2003; Giovagnoli, 2001). Moreover, many clinical studies show that damage in non-frontal (Leskela et al., 1999; van den Broek, Bradshaw, & Szabadi, 1993) or diffuse damage in frontal and non-frontal regions (Anderson, Damasio, Jones, & Tranel, 1991; Axelrod et al., 1996) both affect WCST performance. More specifically, many authors have reported that damage to temporal (Corcoran & Upton, 1993; Giovagnoli, 2001; Hermann, Wyler, & Richey, 1988; Horner, Flashman, Freides, Epstein, & Bakay, 1996; Strauss, Hunter, & Wada, 1993), subcortical (Mukhopadhyay et al., 2008), hippocampal (Corcoran & Upton, 1993; Giovagnoli, 2001; Igarashi et al., 2002), and even cerebellar regions (Mukhopadhyay et al., 2008) cause similar impairments on WCST performance as those subsequent to frontal lobe lesions.

3. Review of neuroimaging studies

Modern functional neuroimaging techniques have been used in many studies to describe changes in brain activation during WCST performance. Most of these studies have focused on groups of psychiatric patients and normal controls. Here we will focus on the results from normal controls. In principle, normal subjects show a more homogeneous level of behavioral performance than clinical samples and, consequently, their functional brain imaging results are expected to show better anatomical consistency and specificity than lesion studies.

Most neuroimaging studies on WCST performance report a significant increase in metabolic or neural activity within frontal or prefrontal cortical regions (Barceló & Knight, 2002; Barceló et al., 2006; Berman et al., 1995; Catafau et al., 1994, 1998; Cicek & Nalcaci, 2001; Gonzalez-Hernandez et al., 2002, 2003; Kawasaki et al., 1993; Konishi, Jimura, Asari, & Miyashita, 2003; Konishi et al., 1998, 2002; Lie, Specht, Marshall, & Fink, 2006; Lombardi et al., 1999; Marenco, Coppola, Daniel, Zigun, & Weinberger, 1993; Mentzel et al., 1998; Monchi, Petrides, Petre, Worsley, & Dagher, 2001; Nagahama et al., 1996, 1997, 1998; Parellada et al., 1998; Ragland et al., 1998; Rogers, Andrews, Grasby, Brooks, & Robbins, 2000; Tien, Schlaepfer, Orr, & Pearlson, 1998; Volz et al., 1997; Wang, Kakigi, & Hoshiyama, 2001). In a majority of the reviewed studies the increase in activation was found in dPFC (Berman et al., 1995; Gonzalez-Hernandez et al., 2002; Kawasaki et al., 1993; Lie et al., 2006; Lombardi et al., 1999; Marenco et al., 1993; Mentzel et al., 1998; Monchi et al., 2001; Nagahama et al., 1996, 1997; Nagahama et al., 1998; Ragland et al., 1998; Rogers et al., 2000; Volz et al., 1997; Wang et al., 2001) and some studies also revealed