Journal of Neuroscience Methods 151 (2006) 30–37

JOURNAL OF NEUROSCIENCE METHODS

www.elsevier.com/locate/jneumeth

The ABCs of computerized naming: Equivalency, reliability, and predictive validity of a computerized rapid automatized naming (RAN) task

Ashley E. Howe [a], Karen M. Arnell [a,*], Raymond M. Klein [b], Marc F. Joanisse [c], Rosemary Tannock [d]

[a] *Brock University, Canada*
[b] *Dalhousie University, Canada*
[c] *University of Western Ontario, Canada*
[d] *Hospital for Sick Children, Toronto, Canada & Norwegian Centre for Advanced Study, Oslo, Norway*

Received 1 February 2005; received in revised form 27 July 2005; accepted 27 July 2005

Abstract

Population-based studies indicate dual routes to disabled reading in adolescence and adulthood: slowed acquisition of single word reading and ADHD (particularly inattention) in early childhood. Impairments in rapid serial naming may be a factor common to both problems. The gold-standard measure of this ability, the Rapid Automatized Naming Task (RAN; [Denckla MB, Rudel R. Rapid automatized naming of pictured objects, colors, letters and numbers by normal children. Cortex 1974;10:186–202]), has traditionally been administered in a paper–pencil format. Recently however, researchers [Neuhaus GF, Carlson CD, Jeng WM, Post Y, Swank PR. The reliability and validity of rapid automatized scoring software ratings for the determination of pause and articulation component durations. Educ Psychol Meas 2001;61:490–504] have begun to use computerized versions of the RAN. Here a slightly modified computerized version of the RAN was created and the equivalency between the computerized RAN and the conventional version was investigated using a university student sample. Naming times on the conventional and computerized RAN were highly correlated, overall, and for each of the four RAN stimulus types (letters, digit, colors, objects). Conventional and computerized RAN times predicted reading rate and reading comprehension scores equally well and both showed very high test–retest reliability. With our university student sample, findings indicate equivalency between the two testing mediums in all areas examined.

Keywords: Rapid automatized naming; Reading; Equivalency; Reliability; Predictive validity; Adults

Reading disability and Attention-Deficit/Hyperactivity Disorder are both highly prevalent childhood disorders that frequently co-occur in the same child (Shaywitz et al., 1999; Vogel and Holt, 2003; Willcutt and Pennington, 2000). Moreover, recent population-based studies suggest that severe inattention and delayed acquisition of single word reading in early childhood may both lead to low literacy skills in adolescence and adulthood (Decker, 1989; Korhonen, 1995; McGee et al., 2002; Rabiner and Coie, 2000). As a result of its wide-spread prevalence, poor literacy skills carry large societal and economic costs, thus it is important to investigate the cognitive deficits that may underlie reading difficulties.

1. The rapid automatized naming task

Several tasks have been created to help identify the nature of the deficits that underlie reading difficulties. One such task involves rapid serial naming (RSN) and is based on the long established link between serial naming deficits and reading disabilities. First discovered by Geschwind (1965), RSN tasks have shown those with reading disabilities to be slower than normal readers in serially naming stimuli (digits, letters, colors, and objects) (Akerman and Dykman, 1993; Catts et al., 2002; Denckla and Rudel, 1976; Katz et al., 1992; Meyer et al., 1998; Wolf et al., 1986).

* Corresponding author at: Department of Psychology, Brock University, St. Catharines, Ont., Canada L2S 3A1. Tel.: +1 905 688 5550x3225; fax: +1 905 688 6922.
E-mail address: karnell@brocku.ca (K.M. Arnell).

0165-0270/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.jneumeth.2005.07.014

One specific RSN task, which has achieved gold standard status in the assessment of reading ability, is the Rapid Automatized Naming Test (RAN), first created by Denckla and Rudel (1974). Consistently, the RAN has proven to be a reliable and valid measure for identifying different reading levels (e.g., see review by Bowers et al., 1994; Denckla and Rudel, 1976; Meyer et al., 1998; Watson and Willows, 1995), and has been used as a tool for predicting reading ability in both adults and children (Badian et al., 1990; Catts et al., 2002; Felton et al., 1990; Neuhaus and Swank, 2002; Wolf and Bowers, 1999; Wolf et al., 1986).

An important advantage of using the RAN as a tool for diagnosing potential reading problems, is that serial naming is an easy test that can be implemented at any age, even before a child is able to read. Moreover, because of its simplicity and brevity, the demands for prolonged effortful processing and sustained attention are minimal. As a result, performance on the RAN can be used not only to detect and assess reading difficulty in adult populations (Felton and Brown, 1990), but also to predict the likelihood of a future reading disorder in young children, hence enabling early detection and intervention (Schatschneider et al., 2004).

The relationship between serial naming and reading has been established for more than three decades, to the point where the RAN has even been nominated as the "second core deficit" in reading disability after phonology (Wolf and Bowers, 1999). However, there continues to be no agreement as to why the RAN predicts reading. In light of this, researchers are now trying to isolate which cognitive processes are shared between RAN and reading ability (e.g., Arnell et al., 2005).

Several hypotheses have been put forward to explain the RAN's ability to predict reading. For example, Denckla and Rudel (1974, 1976) initially explained the relation between the RAN and reading to be phonological in nature; a hypothesis that has since received some empirical support (Catts, 1989; Wagner and Torgesen, 1987; Wolf, 1991, 1997). On the other hand, others believe the relationship between the RAN and reading to be reflective of variability associated with orthographic knowledge (Mannis et al., 2000; Wolf and Bowers, 1999). Further still, others have shown that it is a global speed of processing that determines performance in both the RAN and reading (Kail and Hall, 1994; Kail et al., 1999; Wolf, 1997), with poor readers suffering from a general, temporal processing deficit, as compared to normal readers (Klein and Farmer, 1995). Recently, Klein (2002) suggested that RAN performance predicted reading ability even after the variance attributable to phonological awareness has been removed because the RAN might uniquely tap the efficacy of neural pathways connecting visual pattern recognition with verbal output modules, pathways required when we read.

2. Computer-based testing

Regardless of why the RAN is predictive of reading, it remains a useful and reliable measure for detecting reading difficulties (Akerman and Dykman, 1993; Felton and Brown, 1990; Klein, 2002; Meyer et al., 1998; Wolf et al., 1986). In its original form, the RAN task (Denckla and Rudel, 1974) measures rapid serial naming using stimulus grids presented on white, chart paper laid opposite the participant on a tabletop. Despite the usefulness of this conventional paper version of the RAN, recent technological advances have made computerized presentation and/or scoring of the RAN a desirable option. For example, Neuhaus et al. (2001) have used voice activated software to separate overall RAN naming times into per item pause time (the time between ending an articulation for one grid item and beginning the articulation for the next grid item) and per item articulation time (the time from the start to the end of a vocalized articulation for a grid item).

Computer versions of traditional paper-based tests are now used in many areas of psychology. For example, as discussed by Mead and Drasgow (1993), computerized versions are now used for personality scales, job attitude surveys, cognitive ability tests, aptitude tests, and clinical instruments. Computer-based tests have a number of advantages, including: precise response time scoring, more options for presenting complex and changing stimuli, faster results and instantaneous statistical analysis, easy administration with less chance of human error, the ability to reach a broad sample (as with on-line surveys for example), allowance for adaptive-style tests, and, in some testing situations computerized versions can even minimize cheating (Mead and Drasgow, 1993).

This widespread conversion of paper–pencil tests to computer-based tests does not come without cost however, as the APA has expressed concern for equivalency when converting a conventionally administered test to a computerized version. They recommend that researchers be aware of the possible inequalities resulting from the different testing mediums and modes of administration. The APA also warns that the equivalency between tests should be established rather than assumed prior to drawing any statistical conclusions (APA, 1985). Not surprisingly then, many studies have focused on establishing equivalency between computerized and conventional tests. Most of these support equivalency, however, some studies caution against using computer-based tests (Mead and Drasgow, 1993; Trimmel et al., 2001; Van de Vijver and Harsveld, 1994). Given that computerized scoring and/or presentation of the RAN has now begun (e.g., Neuhaus et al., 2001a,b), it seems wise to test the equivalency between paper and computerized versions of the RAN prior to widespread use of computerized RAN testing. Testing this equivalency is the focus of the present paper.

3. Computer-based versus paper–pencil tests

There is good reason to expect that computer-based RAN is equivalent to paper-based RAN. Equivalency has been found when converting many traditional paper-based tests to computer-based versions across many areas of psychology. To name just a few examples, equivalency between the paper–pencil and computer-based modes of administration has been observed for; the Harrington-O'Shea Career Decision-Making System (Kapes and Vansickle, 1992), the Differential Aptitude Tests (Alkhandher et al., 1998), four work-related non-cognitive psychological measures (King and Miles, 1995), The Eysenck Personality Questionniare and the Carroll Rating Scale for Depression (Merten and Ruch, 1996). In addition,