ELSEVIER

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, computerized versions of self-evaluative ratings of mood have been shown to be just as reliable as the respective paper–pencil versions (Glaze and Cox, 1991; Tseng et al., 1997).

Nevertheless, there are some exceptions to the equivalence between paper and computer-based tests. For instance, Van de Vijver and Harsveld (1994) found an influence of computerization when comparing the paper-based and the computer-based versions of a general aptitude test battery. Specifically, it was found that the computerized version produced more inaccurate and faster responses as compared to the paper version of the same test, with more complex tasks being less affected by computerization. In their meta-analysis, Mead and Drasgow (1993) found an effect of presentation medium for highly speeded tests, despite finding no such effect for tests assessing ability without concern for speed.

Another study observed differences in cognitive demands (such as attention and control) for those performing both a computerized and paper–pencil version of an error correction task and various other cognitive tasks, with EEG recordings suggesting an increase in mental effort for computerized tasks as compared to paper–pencil versions (Trimmel et al., 2001). However, no significant differences were observed in the participants' ratings of mental effort, or emotional load, and the computer tasks resulted in fewer errors as compared to the paper–pencil tasks.

With respect to mood measurement, it was found that computerized self-rating scales produced a heightened negative mood as compared to the paper–pencil version, and results of mood scales correlated with computer anxiety in the computerized tests, but not in the paper versions (Tseng et al., 1997). This finding is consistent with Brosnan and Davidson's conclusion (1994), that 25–30% of the population suffers to some extent from computer anxiety, which could in turn, affect performance on computerized tests and scales.

Given the importance of demonstrating equivalency between computer-based and paper-based tests, the current study investigated: (1) the correlation between scores on the conventional and the computerized RAN, (2) the relationship between conventional and computer RAN scores with reading test scores, and (3) the test–retest reliability for both the computerized and the conventional RAN. The goal was to investigate the equivalency of the computerized RAN and the conventional RAN as, to date, reliability tests investigating medium effects have not yet been carried out for the RAN; a widely used, and important test of reading ability. To anticipate the results, computerized RAN showed strong equivalency with conventional RAN, equally high test–retest reliability, and was an equally good predictor of reading rate and comprehension.

4. Method

4.1. Participants

Sixty-two introductory psychology students at a major Canadian university were recruited to participate for course credit. All participants reported good corrected or uncorrected visual acuity, no color blindness, and that they began learning English prior to 6 years of age.

4.2. Apparatus

For the computerized RAN, stimulus grids were presented on a 17-in. color CRT monitor. E-Prime software (Schneider et al., 2002) running on a Sony VAIO desktop computer was used to present stimuli, record naming times, and to log errors.

4.3. Stimuli

4.3.1. Computerized RAN

Four 5×10 item grid stimulus displays were created for computer presentation. Each grid consisted of one of the following randomly ordered stimulus types: letters (g, k, m, r), objects (book, dog, chair, hand), colors (boxes colored yellow, red, blue, or green) and digits (2, 4, 6, 9).[1] All stimuli were approximately 1.5 cm in height, with the exception of some objects which reached 2.3 cm. Width varied, with colored boxes measuring 1.5 cm, digits 1.0 cm (approx.), letters ranging from .8 to 1.3 cm, and objects ranging from 1.5 to 2 cm. At an approximate viewing distance of 50 cm, each item subtended approximately 1.7° of visual angle. Each grid covered the majority of the 17-in. monitor, with approximately 3.0 cm between columns (measured from horizontal center to horizontal center), and approximately 4.5 cm between rows (measured from vertical center to vertical center).

4.3.2. Conventional RAN

Four, laminated, paper cards were created, each consisting of a 5×10-item grid using the same stimuli types as in the computerized RAN. The paper grids were created as recommended by the Standard RAN (Denckla and Rudel, 1974) and each stimulus subtype contained the following items which were arranged in random order: letters (a, o, s, d, p), objects (hand, dog, star, book, chair), colored boxes (green, red, blue, black yellow), and digits (2, 4, 9, 6, 7). Colored boxes were 2 cm in width and height; digits, 1.5 cm wide and 2 cm high; objects, 1.7 cm–2 cm wide and 1.2 cm–2.7 cm high; and letters, 1.3 cm wide and 1.5 cm–2 cm high. Each grid was centered on a 16.5 in. × 11 in. paper card. The distance between each row measured approximately 5.5 cm from vertical center to vertical center. The distance between each column measured approximately 4.5 cm from horizontal center to horizontal center. Participants had a viewing distance from the paper cards of approximately 40 cm such that each stimulus subtended approximately 2.1° of visual angle.

4.3.3. Nelson–Denny Reading Test (cut time version)

Reading rate and reading comprehension portions of Form D of the Nelson–Denny Reading Test (Brown et al., 1973) were

[1] Note that different letters, as compared to those used in the conventional RAN, were used to create the computerized RAN. The letters chosen follow from the computer-based RAN task used by Arnell et al. (2005), which were modified because additional tasks in their study required consonants only. Digits, colors, and objects used in the computer-based ran were the same as those used in the conventional RAN, with the exception of one less item (4) in the computerized version.

Table 1
Whole grid and per item (in bold) naming times with standard deviations (in brackets) for RAN times (in seconds) as a function of stimulus type, presentation method, and first or second session

Stimulus type	Computer RAN/ Session 1	Computer RAN/ Session 2	Paper RAN/ Session 1	Paper RAN/ Session 2	Overall computer RAN	Overall paper RAN
Letters	19.0 (2.5)	18.1 (2.7)	16.3 (2.7)	15.7 (2.8)	18.6 (2.5)	16.0 (2.6)
	.38	**.362**	**.326**	**.314**	**.372**	**.32**
Objects	31.1 (3.8)	28.4 (3.3)	27.8 (3.4)	27.0 (3.8)	29.7 (3.3)	27.4 (3.3)
	.622	**.568**	**.556**	**.54**	**.594**	**.548**
Colors	27.8 (4.6)	26.4 (4.4)	27.0 (4.3)	26.0 (4.6)	27.1 (4.3)	26.5 (4.3)
	.556	**.528**	**.54**	**.52**	**.542**	**.53**
Digits	18.5 (3.1)	17.8 (3.2)	17.3 (3.0)	16.5 (3.5)	18.1 (3.0)	16.9 (3.0)
	.37	**.356**	**.346**	**.33**	**.362**	**.338**

used to test participants' reading rate and reading comprehension. The 15-min cut-time administration was used.

4.4. Procedure and design

Participants were run individually in a single session lasting under 2 h. All participants performed tasks in a fixed order.[2] Participants first completed the computerized version of the RAN; naming the grid for letters, followed by objects, colors, and finally digits. Next, participants performed the paper version of the RAN, naming stimuli grids in the same order as in the computerized version. Upon completion of both RAN tasks (referred to as "Session 1") the participants completed the 15-min Nelson–Denny Reading Test, followed by a 15-min break. During the break, taken in the lab, participants were invited to do homework or to browse the Internet. After the break, participants completed "Session 2", consisting of the RAN tasks in the same order as in "Session 1", first the computerized then, the paper–pencil version.

4.4.1. Computerized RAN

Participants were instructed to sequentially name aloud each item in the grid from the top left item to the bottom right item as quickly as possible without errors. This was repeated for each stimulus grid (letters, objects, colors, and digits). Participants were also instructed to self-correct during naming, such that if a known error was made they were to correct themselves and continue. In order to begin the trial for each grid, the experimenter pressed the "b" key, which simultaneously began the timer, and caused the stimuli to appear. Immediately upon presentation of each grid, the participant began naming until reaching the final item at which point the "b" key was immediately pressed to end timing. Following each trial, the experimenter entered the number of naming errors made during timing. Naming time for stimulus grids was also measured manually, using a stopwatch (as in Paper RAN trials). The experimenter pressed a stopwatch button with her non-dominant hand to both begin and end the manual timing, while simultaneously pressing the "b" key with her dominant hand to begin and end the E-Prime recording.

[2] When the primary purpose of a study is correlational in nature, fixing the order of the conditions is preferable as this removes RT differences due to order as a source of unexplained variability.

4.4.2. Paper RAN

Upon completion of all four computerized RAN grids, participants were asked to perform the same task, in the exact same manner, but this time using grids presented on the white, 16.5 in. × 11 in. laminated, paper cards. Each card, containing one of the four stimulus grids, was placed on the desk in front of the participant. Participants were asked to verbally and serially name aloud the items in the grid as quickly and accurately as possible and were discouraged from tracing the stimuli with their finger. Again, participants were instructed to self-correct if a known error was made. Naming time for each grid, from onset of first vocal response to end of last stimulus named, was measured using a manual stopwatch.

4.5. The Nelson–Denny Reading Test

Each participant was administered the cut-time version of the Nelson–Denny Reading Test. The 15-min test had two parts: reading rate (determined in the first minute) and reading comprehension (determined in the first minute plus the remaining 14 min of the test). To determine reading rate, the participant was instructed to silently read for one timed minute (recorded using a manual stopwatch), as quickly as possible, but still at a level of good comprehension. After 1 min the experimenter instructed the participant to point to the last word read. The number presented on that row of text was used as the reading rate score.

To determine reading comprehension, the participant was instructed to complete as much of the test booklet as possible in the remaining 14 min (instructions for the entire test were given prior to the first timed minute). Participants were asked to read each story and to answer the corresponding multiple-choice questions for each section as quickly and as accurately as possible. The experimenter also informed participants that they would not be penalized for skipping questions they found particularly challenging, but were they discouraged from skipping entire sections due to difficulty.

5. Results

Means and standard deviations for RAN times are presented in Table 1 as a function of stimulus type, method of presentation (computer or conventional paper) and session (first or second).

Table 2
Relationship between stimulus type, within and across, RAN tests

	Comp./letters	Comp./objects	Comp./colors	Comp./digits	Conv./letters	Conv./objects	Conv./colors
Comp./objects	.42**	–					
Comp./colors	.51**	.69**	–				
Comp./digits	.85**	.47**	.64**	–			
Conv./letters	**.85****	.55**	.60**	.90**	–		
Conv./objects	.42**	**.88****	.67**	.49**	.54**	–	
Conv./colors	.46**	.73**	**.90****	.55**	.55**	.72**	–
Conv./digits	.75**	.36**	.54**	**.86****	.82**	.41**	.50**

** Correlation significant at the .01 level (two-tailed).

RAN times were not log transformed because RAN times for each combination of session and medium were very close to a normal distribution (Skewness < .19 for all combinations). Because a university sample was used (containing fairly skilled readers), and participants were instructed to self-correct any recognized naming errors, no participant made an uncorrected error during naming. The correlations amongst RTs from the four stimulus types were moderate to large for conventional RAN (r's = .41–.82, see Table 2), and computer RAN (r's = .42–.85, see Table 2), justifying the creation of overall RAN scores for each RAN test version.

Naming speeds for the computerized RAN were timed with both a stopwatch, and E-Prime. RAN RTs recorded by stopwatch and RAN RTs recorded by E-Prime shared a correlation of .99 overall, and .99 for each session, suggesting substantial uniformity between the two methods. Indeed, numerically the times recorded by the two methods were strikingly equivalent, suggesting very little systematic experimenter error. Mean times for the first computer RAN session were 24,059 ms as timed with the stopwatch and 24,109 ms as timed by E-Prime. Mean times for the second computer RAN session were 22,687 ms as timed with the stopwatch and 22,666 ms as timed by E-Prime. In the analyses presented here stopwatch RTs were used, but nearly identical results were observed with E-Prime RTs.

Because it was hypothesized that the computerized RAN and the conventional RAN are equivalent tests, measuring the same cognitive processes, it was expected that the correlation between these two testing mediums would be very high. Indeed, the results supported this expectation with a very high correlation of .95 ($p < .001$) between overall scores on the conventional RAN and overall scores on the computerized RAN. In addition, to test the interchangeability of overall computer scores with overall paper pencil scores, an intraclass correlation (ICC) was performed. Results indicated an ICC of .80 ($p < .001$), suggesting overall scores for both RAN tests to be fairly interchangeable. As expected, results from the intra-class correlation analysis were lower than those from the Pearson correlations because the ICC also incorporates mean differences between scores. Notice in Table 1 that RAN times were faster for conventional RAN, which was always performed second, than computer RAN, which was always performed first. Thus, the ICC value reflects both the strong relationship between conventional and computer RAN scores as well as the mean difference resulting from practice effects.

Not only were the overall computerized and conventional RAN times highly correlated, but very strong correlations (at least .85) were also observed between conventional and computer RAN times for each of the four stimulus types individually (see bold values in Table 2) indicating the comparability of the two mediums for each of the four stimulus types.

Overall scores (averaged across stimulus type) for both the conventional and the computerized RAN were separated by session in order to examine the test–retest reliability for each RAN medium (see Table 3). For both RAN test mediums, correlations between Session 1 and Session 2 are nearly identical and very high, indicating that reliability is not sacrificed in the computerized version of the RAN (see bolded values in Table 3). In addition, reliability is just as strong across testing mediums as within the same RAN tests. For example, conventional RAN times in Session 1 predict approximately 83% of the variability for conventional RAN times in Session 2. Similarly, these same conventional RAN times predict approximately 86% of the variability in Session 2 computerized RAN times.

Conventional RAN has been shown to explain significant variability in reading ability (e.g. Denckla and Rudel, 1976; Meyer et al., 1998; Watson and Willows, 1995), and this was also observed here where conventional RAN predicted significant variability in both reading rate and reading comprehension (see Table 4). It was expected that the computerized RAN would

Table 3
Relationship between overall scores on conventional and computerized RAN separated by session

	Computer RAN Session 1	Computer RAN Session 2	Conventional RAN Session 1
Computer RAN Session 2	**.92****		
Conventional RAN Session 1	.91**	.93**	
Conventional RAN Session 2	.86**	.92**	.91**

** Correlation significant at the .01 level (two-tailed).

Table 4
The relationship between overall computer and conventional RAN scores and reading comprehension and rate

	Reading comprehension	Reading rate	Computer RAN
Reading rate	.45**	–	
Computer RAN	−.29*	−.29*	–
Conventional RAN	−.26*	−.29*	.95**

* Correlations significant at the .05 (two-tailed).
** Correlations significant at the .01 (two-tailed).

also show comparable correlations with reading measures, and this expectation was fulfilled. Significant correlations, equivalent to those observed with conventional RAN, were indeed observed between computerized RAN and both reading comprehension and reading rate (see Table 4), demonstrating that the computerized RAN is equally predictive of reading ability.

A regression analysis was performed where overall conventional RAN scores and overall computerized RAN scores were entered simultaneously as predictors of reading rate. Although both conventional and computer RAN scores were significant predictors of reading rate scores when examining the zero-order correlations, neither explained unique variability in reading rate over and above the other (semi-partial r's = −.03 and −.06, p's > .79 and .62 for conventional and computer RAN scores, respectively). When conventional RAN scores and overall computerized RAN scores were entered simultaneously as predictors of reading comprehension the same pattern was observed (semi-partial r's = .06 and −.15, p's > .66 and .25 for conventional and computer RAN scores, respectively). Therefore, conventional RAN and computerized RAN are equivalent in that they explain the same variability in reading performance, with neither RAN task explaining any unique variability in reading performance.

6. Discussion

The present study shows strong equivalency between conventional and computerized RAN versions in a university student sample. Very high correlations between conventional and computerized RAN were observed overall, for each stimulus type, and from Session 1 to 2. This suggests that the two RAN testing modes are *equivalent* and reliable and can be used interchangeably despite differences in the method of stimulus presentation and the exact stimulus characteristics. Replicating previous research that has shown good test–retest reliability for other computer-based tests, our findings also indicate that computerized RAN is highly reliable. Finally, consistent with the conventional RAN's reputation for being predictive of reading ability, the computerized RAN was found to predict reading to the same degree as the paper version, again supporting their equivalency.

As the APA stipulates, equivalency must be determined prior to administering the computerized version of any conventional, paper-based test. In the present investigation, it is not particularly surprising that equivalency was indeed found, as the only large difference between the two RAN tests was whether stimuli were presented vertically on a computer screen, or horizontally on a sheet of paper. In addition, the RAN does not require manual interaction with the computer, thus eliminating the potential medium effects associated with computer anxiety (Tseng et al., 1997). Similarly, both the computer and paper RAN require the same mode of response (i.e., vocal articulation), and the initiation and termination of timing is controlled by the experimenter for both versions. Thus, the reduced equivalency found previously for speeded tasks (see Mead and Drasgow, 1993), where responses varied from the paper version (e.g., fill in the bubble) to the computer version (e.g., key presses) are not relevant to the RAN.

Extremely high correlations were found between paper and computer RAN for all stimulus types, despite using only four exemplars for each of the stimulus types instead of five as used in the conventional RAN, suggesting that there is nothing particularly important about having five exemplars. Four of the five digit, letter, and object stimuli from the conventional RAN were used in the computerized version, but none of the same letter stimuli were used (see footnote 1). Even though all of the letters used in the computerized version, were different from those used in the conventional version, the correlation between computer and conventional RAN was as large for letter stimuli as for digits, colors, and objects, where four of the stimuli were the same. This suggests that there is nothing especially useful or predictive about the exact letters used in the conventional RAN test when using a university student sample. However, letters used in the conventional RAN may lead to more errors with young children or pre-readers than letters used here for the computer RAN given that the original letters such as d and p could make alternate letters if flipped (b and q), but this is not true for letters used here with computer RAN. Such flipping may be more common in beginning readers than university students.

Test–retest reliability was also very good for the computerized RAN, as consistent with Kapes and Vansickle (1992) findings showing computer-based tests to be at least as reliable as paper-based tests. However, it should be noted that in the present study, Sessions 1 and 2 were separated by just under an hour, which may not allow for the strongest test of reliability. Thus, even though the high correlations found between Session 1 and 2 suggest that test–retest reliability would also be strong over longer intervals, this was not tested in the present investigation.

In applied and research settings, the RAN is used both with adult and child participants. However, it should be noted that the present study did not examine the equivalency of paper and computer RAN for children. Also, the conventional RAN has been shown to predict orthographic awareness (e.g. Wolf and Bowers, 1999) and phonological knowledge (eg. Wagner and Torgesen, 1987; Wolf, 1991, 1997) in children and this was also not tested here. A future study examining how computer RAN predicts reading ability, as well as, orthographic and phonological awareness in children would also be a useful addition to the literature. Also, the use of a university student sample consisting of fairly skilled readers likely provides less variability in RAN and reading measures than might be expected with a child sample or a general adult sample where poor readers would be included. It should be noted that this reduction in variability likely under-represents the correlations presented here.

In sum, based on the present findings, it was determined that computer-based and conventional RAN are equivalent, at least with a university sample. As a result, the advantages to a computerized RAN can be utilized, which include: removing the need to transcribe manually recorded data into electronic form for analysis; And more accurate stimulus presentation, since the table-top cards used in the conventional RAN allow for more variability in presenting stimulus grids. Although not used here, the use of computerized presentation also allows for the addition of audio-software that digitizes the speech stream. Such software allows for the partitioning of overall RAN times into pause time and articulation time (e.g., Neuhaus et al., 2001a,b). This latter advantage may be particularly useful given that recent data suggest that the pause time (the mean interval between naming each stimulus) is a better predictor of reading than articulation time (Neuhaus et al., 2001a,b). As a result of such advantages, and the present findings of equivalency between paper and computer RAN, we advocate the use of computerized RAN, and the collection of overall RAN times, RAN pause times, and RAN articulation times, for the purpose of estimating rapid naming ability, distinguishing different reader types, and detecting a current, or future, reading disorder.

Acknowledgement

The research was made possible with support from the Canadian Language and Literacy Research Network (CLLRNet) to KMA, MFJ and RMK, and RT, and support (to RT) from the Centre for Advanced Study with the Norwegian Academy of Science and Letters, Oslo, Norway.

References

Akerman PT, Dykman RA. Phonological processes, confrontational naming, and immediate memory in dyslexia. J Learn Disabil 1993;26:597–609.

Alkhandher O, Clarke DD, Anderson N. Equivalence and predictive validity of paper-and-pencil and computerized adaptive formats of the Differential Aptitude Tests. J Occup Organ Psychol 1998;71:205–17.

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. Standards for educational and psychological testing. Washington, DC:American Psychological Association, 1985.

Arnell KM, Joanisse MF, Klein RM, Tannock R, Barr C. Decomposing RAN: The contribution of object recognition and attention to rapid naming and reading; 2005, submitted for publication.

Badian NA, McAnulty GB, Duffy FH, Als H. Prediction of dyslexia in kindergarten boys. Ann Dyslexia 1990;40:152–69.

Bowers PG, Golden J, Kennedy A, Young A. Limits upon orthographic knowledge due to processes indexed by naming speed. In: Beringer VW, editor. The varieties of orthographic knowledge. Theoretical and developmental issues, vol. 1. Dordrecht, Netherlands: Kluwer; 1994. p. 173–218.

Bronson MJ, Davidson MJ. Computerphobia—is it a particularly female phenomenon? The Psychol 1994;7:73–8.

Brown JI, Nelson MJ, Denny EC. The Nelson–Denny reading test—form C and D. Boston, MA: Houghton Mifflin Company; 1973.

Catts HW. Phonological processing deficits and reading disabilities. In: Kamhi A, Catts H, editors. Reading disabilities: a developmental language perspective. Boston: Allyn & Bacon; 1989. p. 101–32.

Catts HW, Gillispie M, Leonard LB, Kail RV, Miller CA. The role of speed of processing, rapid naming, and phonological awareness in reading achievement. J Learn Disabil 2002;35:510–25.

Decker SN. Cognitive processing rates among disabled and normal reading young adults: a nine-year follow-up study. Reading Writing 1989;1:123–34.

Denckla MB, Rudel R. Rapid automatized naming of pictured objects, colors, letters and numbers by normal children. Cortex 1974;10:186–202.

Denckla MB, Rudel R. Rapid “automatized” naming (R.A.N.): dyslexia differentiated from other learning disabilities. Neuropsychologia 1976;14:471–9.

Felton RH, Brown IS. Phonological processes as predictors of specific reading skills in children at risk for reading failure. Reading Writing 1990;2:39–59.

Felton RH, Naylor CE, Wood FB. Neuropsychological profile of adult dyslexics. Brain Language 1990;39:485–97.

Geschwind N. Disconnection syndrome in animals and man (part 1 & 2). Brain 1965;88(237–294):585–644.

Glaze RC, Cox JL. Validation of a computerized version of the 10-item (self-rating) Edinburgh postnatal depression scale. J Affect Disord 1991;22:73–7.

Kail R, Hall LK. Speed of processing, naming speed, and reading. Dev Psychol 1994;30:949–54.

Kail R, Hall LK, Caskey BJ. Processing speed, exposure to print, and naming speed. Appl Psycholinguist 1999;20:303–14.

Kapes JT, Vansickle TR. Comparing paper–pencil and computer-based versions of the Harrington-O’Shea career decision making system. Meas Eval Counsel Dev 1992;25:5–13.

Katz WF, Curtiss S, Tallal P. Rapid automatized naming and gesture by normal and language-impaired children. Brain Language 1992;43:623–41.

King WC, Miles EW. A quasi-experimental assessment of the effect of computerizing non-cognitive paper-and-pencil measurements: a test of measurement equivalence. J Appl Psychol 1995;80:643–51.

Klein RM. Observations on the temporal correlates of reading failure. Reading Writing: An Interdiscipl J 2002;15:207–32.

Klein RM, Farmer ME. Dyslexia and a temporal processing deficit: a reply to the commentaries. Psychol Bull Rev 1995;2:515–26.

Korhonen TT. The persistence of rapid naming problems in children with reading disabilities: A nine-year follow-up. J Learn Disabil 1995;28:232–9.

Mannis FR, Doi LM, Bhada B. Naming speed, phonological awareness, and orthographic knowledge in second graders. J Learn Disabil 2000;33:325–33.

McGee R, Prior M, Williams S, Smart D, Sanson A. The long-term significance of teacher-rated hyperactivity and reading ability in childhood: findings from two longitudinal studies. J Child Psychol Psychiatr 2002;43:1004–17.

Mead AD, Drasgow F. Equivalence of computer and paper-and-pencil cognitive ability tests: a meta-analysis. Psychol Bull 1993;114:449–58.

Merten T, Ruch W. A comparison of computerized and conventional administration of the German versions of the Eysenck Personality Questionnaire and the Carroll Rating Scale for Depression. Pers Indiv Differ 1996;20:281–91.

Meyer MS, Wood FB, Hart LA, Felton RH. Selective predictive value of rapid automatized naming in poor readers. J Learn Disabil 1998;31:106–17.

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 2001a;61:490–504.

Neuhaus G, Foorman BR, Francis DJ, Carlson CD. Measures of information processing in rapid automatized naming (RAN) and their relation to reading. J Exp Child Psychol 2001b;78:359–73.

Neuhaus GE, Swank PR. Understanding the relations between RAN letter subtest components and word reading in first-grade students. J Learn Disabil 2002;35:158–74.

Rabiner D, Coie JD. Early attention problems and children’s reading achievement: a longitudinal investigation. The Conduct Problems Prevention Research Group. J Am Acad Child Adolesc Psychiatr 2000;39:859–67.

Schatschneider C, Fletcher JM, Francis DJ, Carlson CD, Coleen D, Foorman BR. Kindergarten prediction of reading skills: a longitudinal comparative analysis. J Educ Psychol 2004;96:265–82.

Schneider W, Eschman A, Zuccolotto A. E-Prime User's Guide. Pittsburgh: Psychology Software Tools Inc; 2002.

Shaywitz SE, Fletcher JM, Holahan JM, Schneider AE, Marchione KE, Stuebing KK, Francis DJ, Pugh KR, Shaywitz BA. Persistence of dyslexia: The Connecticut longitudinal study at adolescence. Pediatrics 1999;104:1351–9.

Trimmel M, Strassler F, Knerer K. Brain DC potential changes of computerized tasks and paper–pencil tasks. Int J Psychophysiol 2001;40:187–94.

Tseng HM, Macleod HA, Wright P. Computer anxiety and measurement of mood change. Comput Hum Behav 1997;13:305–16.

Van de Vijver FJR, Harsveld M. The incomplete equivalence of the paper-and-pencil and computerized versions of the General Aptitude Test Battery. J Appl Psychol 1994;79:852–9.

Vogel SA, Holt JK. A comparative study of adults with and without self-reported learning disabilities in six English speaking populations. What have we learned? Dyslexia 2003;9:193–228.

Wagner RK, Torgesen JK. The nature of phonological processes and its causal role in the acquisition of reading skills. Psychol Bull 1987;101:112–212.

Watson C, Willows DM. Information-processing patterns in specific reading disability. J Learn Disabil 1995;28:216–31.

Willcutt EG, Pennington BF. Comorbidity of reading disability and attention-deficit/hyperactivity disorder: differences by gender and subtype. J Learn Disabil 2000;33:179–91.

Wolf M. Naming speed and reading: The contribution of the cognitive neurosciences. Read Res Quart 1991;26:123–41.

Wolf M. A provisional, integrative account of phonological and naming speed deficits in dyslexia: Implications for diagnosis and intervention. In: Blachman B, editor. Foundations of reading acquisition and dyslexia: implications for early interventions. Marhwah: Earlbaum; 1997. p. 67–92.

Wolf M, Bally H, Morris R. Automaticity, retrieval processes, and reading: a longitudinal study in average and impaired readers. Child Dev 1986;57:988–1000.

Wolf M, Bowers PG. The double-deficit hypothesis for the developmental dyslexias. J Educ Psychol 1999;91:415–38.