Design and evaluation of a computerised version of the Benton visual retention test

S.B.N. Thompson b, E. Ennis a,*, T. Coffin a, S. Farman a

a Centre for the Study of Emotion Department of Psychology, University of Portsmouth, King Henry Building, King Henry 1 Street, Portsmouth PO1 2DY, UK
b Neuropsychology Assessment and Consultancy, Portsmouth, UK

Available online 22 May 2006

Abstract

The use of computers in the administration of psychological assessments is often considered standard practice. However, the evidence clearly shows that computerisation of each test needs to be evaluated independently. The current study examined the hypothesis that a computerised administration of the Benton visual retention test (BVRT) should yield comparable results to paper-and-pencil administration of this measure. Forty participants (23 females and 17 males) from a non-clinical population were assessed using a computerised version of the BVRT and the conventional paper-and-pencil administration. Parallel forms of the test were used in the two administrations in order to eliminate practice effects. Participants found the conventional method of assessment easier to use but less fun. Importantly, performances of the participants were poorer when using the computerised version, giving rise to extreme caution when using this method of assessment administration with a clinical population.

Keywords: Attitude to computers; Benton visual retention test; Computerised assessment; Computer preferences; Memory

This research was funded by the Summer Bursary Scheme, University of Portsmouth, Department of Psychology, and has been approved by the University's Ethics Committee.

* Corresponding author. Tel.: +44 23 9284 6312.
E-mail address: edel.ennis@port.ac.uk (E. Ennis).

0747-5632/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.chb.2006.03.014
1. Introduction

Computers have the potential to be used across a variety of settings, and for a variety of procedures, ranging from administration of psychological tests to test scoring and interpretation (Schulenberg & Yutrzenka, 2004; Weber et al., 2003). The current paper examines the potential use of computers within the administration of the Benton visual retention test (BVRT), which is often used in the clinical setting. Towards the end of the last century, computer-based psychological assessments had become almost mainstream in terms of clinicians’ tools for the assessment of a range of neuropsychological conditions and sequelae. However, this has not been without comment or reservations from researchers and clinical workers alike.

Since the late 1980s, growing technological sophistication has been evidenced in the psychological assessment of patient populations (Schulenberg & Yutrzenka, 2004). In particular, this has been evident in the use of computers to assess personal sensitive information amongst students (Knapp & Kirk, 2003), administer (and interpret) the MMPI in a non-clinical sample (Pinsoneault, 1996). In addition to this, within clinical samples, computers have been successfully used to assess depression and function in primary care older patients (Kurt, Bogner, Stratton, Tien, & Gallo, 2004), attention and memory in psychiatric patients (Weber et al., 2003), visual retention in neurological patients (Merten, 1999), and memory difficulties (Aharonson & Korczyn, 2004), and the use of microcomputers to assess the recovery of stroke or incompletely innervated patients (Thompson, 1987; Thompson & Coleman, 1987, 1989; Thompson, Coleman, & Yates, 1986; Thompson & Morgan, 1996).

The majority of users accept computerised psychological assessment, and agree that this holds considerable advantages over paper-and-pencil versions of assessments (e.g. speed and thus cost efficiency, consistency of administration procedures) (Weber et al., 2003). However, the implications of changes such as the removal or reduction of the human examiner from the interaction and the need for adherence to the Data Protection Acts must also be considered (Schulenberg & Yutrzenka, 2004). This has spawned guidelines for choosing suitable computerised assessments (APA, 1986; BPS, 2002; Green, Bock, Humphreys, Linn, & Reckase, 1984; ICTP, 1988; Schoenfeldt, 1989), and examinations of users’ attitudes towards computers in general (Bouman, Wolters, & Wolters-Hoff, 1989).

Users’ attitudes towards computers are important particularly from an ethical point of view; patients should not be exposed to procedures that they find aversive (Schulenberg & Yutrzenka, 2004; Weber et al., 2003). In addition to being important from an ethical point of view, these confounding variables such as attitudes towards computers also practically influence the process and outcome of computerised testing sessions (Schulenberg & Yutrzenka, 2004; Weber et al., 2003). Research shows that both acceptance and previous experience with computers increase user’s motivation and success of the patient–computer interaction (Schulenberg & Yutrzenka, 2004; Weber et al., 2003).

Computers may be used to either generate new psychological measures, or to adapt conventional measures into computerised formats (Schulenberg & Yutrzenka, 2004). As was outlined earlier, computerised assessment is effective for a variety of purposes, within both clinical and non-clinical settings. Moreover, computerised assessment is well accepted by patients despite prior apprehension about taking the test (Weber et al., 2003), and effective even when participants had no prior computer experience (Kurt et al., 2004). How-
ever, psychometric equivalence (i.e. the applicability of the same norms, etc.) across modalities should not be assumed (Schulenberg & Yutrzenka, 2004; Williams & McCord, in press).

Effective use of computers for administration of tests has been predominantly used for verbal tests that involve reading simple questions and answering with simple responses (Williams & McCord, in press). Also, results of equivalence across modalities (computer versus paper and pencil) have been more consistent for tests in this domain (Murphy & Davidshofer, 2005; Williams & McCord, in press). The computerization of tests involving graphics and manipulation of materials on the computer screen (e.g. Ravens progressive Matrices (RPM)) has not been either as widespread or as effective (Williams & McCord, in press). For example, although methodological differences across studies must be considered, findings has been inconsistent as to whether individuals score the same on a computerised administration of Ravens progressive Matrices and a paper-and-pencil version, or whether performance is poorer on the computerised version (Williams & McCord, in press).

Overall, it is clear that the adoption of an equivalent computer-based version of an existing psychological assessment tool should be met with some caution. Needless to say, the pre-requisites of validity, reliability, specificity, consistency, etc. should be found. Then, it is necessary to make judgements about the ease of use, the comfort of use by the administrator and examinee, the relative cost of such an assessment tool, and the comparability of the scores obtained through a computerised administration of the test to those obtained through a paper-and-pencil administration. This should be evaluated for each test independently, especially if the test is visual–spatial based. This point is even more poignant in relation to the BVRT given that the test is visuo-spatial, and participants are required to actually generate (draw) their response as opposed to simply selecting a correct alternative (e.g. as is done within the Ravens Progressive Matrices).

In order to examine equivalence across modalities and some of the above issues, a computerised version of the Benton visual retention test (BVRT) (Benton, 1992) was designed. Questionnaires were also designed to evaluate key issues such as whether users find such tools easy to use or comfortable to use, or whether they prefer conventional methods of administration. Results will examine how users perform on a parallel version of the BVRT when it is administered using a computer screen and requiring responses to be input directly to the computer.

The BVRT is regarded as a reliable and robust tool in many clinical testing situations and has been used in the early detection of Alzheimer’s disease (Thompson, MacDonald, & Coates, 2001), for dementia in the elderly (Jaqcmin-Gadda, Fabrigoule, Commenges, Letenneur, & Dartigues, 2000), and for older neurological populations (Coman et al., 2002). Computerising such a test presents a challenge. It cannot be assumed that test characteristics such as ease of administration, reliability, and accuracy will not change. Also, given that the test is most often used within sensitive populations, it is important to examine whether participant factors such as attitudes towards computers or previous experience with computers are influential in determining whether the computerised assessment is comparable to the paper-and-pencil assessment.

The current research was inspired by Merten’s (1999) work on a computer-administration of the BVRT using neurological patients. Merten’s (1999) findings showed no significant differences between a group who were administered the BVRT via a computer and a group who were administered the assessment via paper and pencil. However, the fact
remains that they used a between subjects design. While these two groups were matched with regards to age, gender and aetiology, no mention is made of whether attitudes towards computers or computer experience were examined. With regard to examination of within subjects patterns, a marked decline in patients’ performances was evident when the computerised version followed the conventional paper-and-pencil version (Merten, 1999). However, this may be attributable merely to fatigue effects, especially given that the sample was of a clinical nature.

The current study attempts to further clarify this research area by investigating these questions within a normal population, and overcoming the methodological issues evident within the research of Merten (1999). Specifically, a within subjects design will be used to ensure that any differences in performance are due to the effects of mode of administration (computer versus paper and pencil). Also, the method received first (computer versus paper and pencil) will be controlled in order to ensure that any differences are not simply due to fatigue effects. Also building upon the research of Merten (1999), the current study will consider the importance of users’ attitudes towards computers and computer experience. As was outlined earlier, both these factors should be considered ethically (Bouman et al., 1989), and they are also likely to practically influence the process and outcome of computerised testing sessions (Schulenberg & Yutrzenka, 2004; Weber et al., 2003).

Overall, the present study seeks to clarify whether findings obtained administering the BVRT via a computer are comparable to those obtained administering the test via paper and pencil. Obviously, this issue has considerable practical клиническими implications in terms of time saved via reduced examiner involvement, the financial implications of this, the benefits of computers to achieving standardisation across administrations etc. Eventually, it may be possible to develop a process whereby the BVRT may be scored by means of a computer program, however it is essential to firstly demonstrate that participant’s performance on both formats is equivalent.

Although they cannot be assumed to directly translate, given the findings of Merten (1999), it is hypothesised that performance on the computerised version of the BVRT will be comparable to performance on the paper-and-pencil version of the test (H1). Secondly, given their earlier discussed importance on the success of computerised testing, performance on the computerised administration of the BVRT will be positively related to positive attitudes towards computers (H2).

2. Method

2.1. Participants

Forty individuals (23 females; 17 males) were recruited from the University’s pool of volunteer participants. The age range of this non-clinical population was 19–59 years old (mean = 30.67; SD = 12.58).

2.2. Materials

A conventional paper-and-pencil version as well as a computerised version of the BVRT was performed by each participant. Firstly, the Hospital Anxiety and Depression Scale (HADS) (Snaith & Zigmond, 1983) was administered. This comprises a 14-item questionnaire used to assess the severity of both depression and anxiety in order detect
probable caseness. Cut-off points indicate whether someone is ‘within the normal range’ (0–7), ‘mildly’ (8–10), ‘moderate’ (11–14) or ‘severely’ (15–21) depressed or anxious. Psychometric properties of the instrument are satisfactory and have been well documented elsewhere (Snaith & Zigmond, 1983).

The National Adult Reading Test (NART) (Nelson, 1982) was also administered and consists of 50 words printed in two columns on a card presented to the participant. These words are relatively short in order. Although not important in the current non-clinical sample, this is important in clinical samples in that it avoids stimulus complexity adversely affecting the reading of dementing subjects. However, the words are all ‘irregular’ with respect to the common rules of pronunciation, to avoid participants reading by phonemic decoding rather than word recognition. The participant reads each word aloud and the number of errors made in pronunciation is recorded. There is evidence that the NART shows satisfactory internal consistency, is a valid predictor of IQ, and performance on the test is unrelated to either age or social class (Nelson, 1982). Full scale Intelligence Quotient (IQ) as assessed by the Wechsler Adult Intelligence Scale Revised (WAIS-R) can be predicted from this reading error score. The scale was therefore included in the current study to both describe the IQ of the sample given that they were recruited in a university setting, and to examine whether IQ related to participants performance on the BVRT (on either computer or paper-and-pencil administration).

The BVRT assesses visual perception, visual memory, visuo-constructive abilities, some attention difficulties and visual neglect in both clinical and research settings. The test consists of 10 designs within a stimulus booklet, with each design containing one or more (typically three) figures. Three almost equivalent forms of the BVRT are available, and four alternative methods of administration. These vary in length of exposure time, and the delay period between exposure and reproduction of the stimuli by the participant. The participant is shown each design for a set period of time, after which the design is taken away and the participant is required to reproduce the design. ‘Number correct’ and ‘number error’ scores are calculated, and compared to the normative data. Broadly speaking, the number correct is calculated in terms of how similar the participants reproduced design is to the initial design they were shown. Number of errors is calculated in terms of how much error the participant has made within their reproduced design in comparison with the initial stimuli. For the computerised administration of the BVRT within the current study, a slide show within Microsoft PowerPoint was used to present each of the designs for the specified 10 s administration. Following this, a blank screen automatically appeared and the participant was required to reproduce the designs on an A5 ActivPenta graphics tablet (Misco, 2004) connected to the personal computer (PC). A cordless pen is used on the surface of the tablet, which communicates wirelessly with an adaptor connected to the computer’s USB port. As the participant draws on the graphics tablet, a representation of their drawing appears on the computer screen. On completion of the drawing, the participant was required to click the computer mouse for the next design to appear on the screen in front of them.

Finally, participants completed the Groningen Computer Attitude Scale (GCAS) (Bouman et al., 1989) which was translated from German, a questionnaire on method preference (adapted from Merten, 1999), and a questionnaire on computer experience that was specifically constructed for this study. The GCAS is a 16-item measure of general attitude towards computers that required participants to indicate their response using a 5-point Lickert scale (i.e. ranging from ‘totally agree’ to ‘totally disagree’). The possible score
range is 16–80, with scores above 48 considered to represent a positive attitude towards computers (Bouman et al., 1989). This instrument has satisfactory psychometric properties (Weber, Fritze, Schneider, Kühner, & Maurer, 2002). The 8-item measure of method preference asked of participants which method (paper-and-pencil versus computerised version) they could best concentrate on, which one was most pleasant, which they would prefer for further tasks of a similar nature, which one they found easier, which one they would prefer if they had to carry out such tasks for a job interview, which gave them the most fun, which they found most comfortable and finally, with which method did the time go fastest. Response options were ‘paper and pencil’, ‘computer’ or ‘no difference’.

The newly constructed measure of computer experience consisted of 4 items, each of which was subdivided into two sections. Item 1 enquired whether or not participants had ever used a computer before, and if so how much experience they had of using a computer (limited, intermediate or extensive experience). Item 2 enquired whether participants used a computer regularly in their everyday life, such as for work or study, and if so, how often they used the computer (occasionally, frequently, or everyday). Item 3 asked whether participants were comfortable using computers or computerised technology, and if so, how comfortable they were (fairly, moderately or very comfortable). Item 4 asked whether participants had ever used a computerised pad to input information into a computer before, and if so, how much experience they had of using it (limited, intermediate or extensive experience). Exact answer alternatives were not given. Although there are limitations with this approach (i.e. individuals interpreting the terms differently), the aim of the questionnaire was to record the participants’ subjective feelings regarding their experience and comfort with computers. It is participants’ own feelings about their interactions with the computer that are most likely to be influential.

2.3. Procedure

The HADS was administered first to ensure that no mild levels of anxiety or depression were present within the sample. Scores above 15 (on both the anxiety and depression components of the questionnaire) indicate such affective disturbances. No participants were found in this category, and thus all proceeded to completing the NART. A code of procedure involving forms of consent and payment at each stage was strictly adhered to in this study in line with University’s Ethics approval and the British Psychological Society’s guidelines (BPS, 2004).

All participants subsequently performed both the computerised administration of the Benton visual retention test and the paper-and-pencil version of the test. As was noted earlier, there are four alternative methods of administering the BVRT. These all vary the length of time for which the participant is exposed to the stimuli, and the length of time between withdrawing the stimulus and the participant generating a reproduction. In the current study, administration A was adopted for both versions (i.e. the computerised and the paper and pencil). In this method of administration, stimuli were displayed for 10 s and then immediately reproduced by the participant. To eliminate order effects, participants were randomly allocated to either the BVRT or paper-and-pencil version to complete first (i.e. ‘paper-and-pencil’ versus ‘computerised’ versions).

In addition to this, Form C of the BVRT was always used within the computerised administration of the test, whilst Form D was always used within the paper-and-pencil administration. The different parallel versions of the BVRT are identical in format (see
earlier materials section for more detailed description of test format and content), but simply use different drawing designs. The use of different design drawings means that practice or memory effects across the two performances (the computer administration and the paper-and-pencil administration) are eliminated, but participants two performances can still be compared. Finally, participants completed questionnaires on computer experience, attitude to computers and method preference. These measures were completed last in order to prevent participants being too fatigued when completing the BVRT, and also to avoid the content of these questionnaires making the objectives of the study too apparent.

3. Results

Mean response scores for anxiety, depression, full IQ and the general attitudes towards computers questionnaire were 6.42 (range 1–12; SD = 2.71); 2.52 (range 0–7; SD = 1.74); 113.52 (range 102–126; SD = 4.82); and 61.6 (range 43–73; SD = 7.56), respectively. The depression, anxiety and IQ measures were included to provide information on the nature of the sample. As can be seen, participants’ depression and anxiety levels were largely within a non-clinical range. IQ scores are also within a normal range. IQ scores bore no significant relation to either the number of errors or the number of correct responses made on the BVRT. Correlations ranged between $-0.2$ and 0.2 (all ns) regardless of the method of administration (computer versus paper and pencil). Table 1 illustrates participant’s responses to the computer experience questionnaire. As can be seen from this table, the majority of individuals had some form of computer experience, used computers regularly in their everyday lives, and were comfortable using computers.

Of those who had computer experience, 10% rated it as limited, 45% rated it as intermediate, 37.5% rated it as extensive, and 7.5% neglected to indicate their level of experience. Of those who used computers regularly in their everyday lives, 2.5% rated their use of computers as occasional, 22.5% rated it as frequent, 62.5% rated it as everyday, and 12.5% neglected to rate their degree of usage. Of those who said they were comfortable using computers, 12.5% said they were fairly comfortable, 30% were moderately comfortable, and 47.5% were very comfortable and 10.0% did not say. Few individuals had experience of using computerised pads, and those who did rated their experience as either limited or intermediate.

To test possible differences in performance on the computerised versus paper-and-pencil version, a repeated measures ANOVA utilizing a 2 (type of administration: number correct within computerised administration versus number correct within paper-and-pencil administration) × 2 (order of administration: computerised first versus paper-and-pencil first) mixed design was employed. The number correct within both forms of administra-

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes (%)</th>
<th>No (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Have you ever used a computer before?</td>
<td>92.5</td>
<td>5.0</td>
</tr>
<tr>
<td>Do you use the computer regularly in your everyday life?</td>
<td>85.0</td>
<td>12.5</td>
</tr>
<tr>
<td>Are you comfortable using computers?</td>
<td>90.0</td>
<td>10.0</td>
</tr>
<tr>
<td>Have you ever used a computerised pad before?</td>
<td>17.5</td>
<td>82.5</td>
</tr>
</tbody>
</table>
tion was a within subjects factor, whilst the order of administration was a *between* subjects factor. Results showed that participants provided significantly more correct responses using the paper-and-pencil version of the test \((M = 8.35; SE = .25)\) compared to the computer version of the test \((M = 6.22; SE = .33)\) \((F(1,38) = 67.19, p < .001, \eta^2 = .64)\). The order of administration did not exert any interactive effect \((F(1,38) = .08; ns)\).

To test for differences in error responses, a second repeated measures ANOVA was employed utilizing a 2 (type of administration: number incorrect within computerised administration versus number incorrect within paper-and-pencil administration) \(\times\) 2 (order of administration: computerised first versus paper-and-pencil first) mixed design was employed. The number incorrect within both forms of administration was a within subjects factor, whilst the order of administration was a *between* subjects factor. Results showed that participants provided significantly less error responses using the paper-and-pencil version of the test \((M = 1.75; SE = .28)\) compared to using the computer version of the test \((M = 5.25, SE = .57)\) \((F(1,38) = 57.92, p < .001, \eta^2 = .60)\). Again, order of administration did not exert any interactive effect on this main effect \((F(1,38) = 0.29; ns)\). Thus, results did not support our hypothesis (H1), in that the performance on the computerised version of the BVRT was significantly different to performance on the paper-and-pencil version of the test. Specifically, performance on the computer version of the test was significantly poorer than performance on the paper-and-pencil version of the test.

Contrary to Hypothesis 2, Pearson’s correlation coefficient showed that the GCAS bore no significant relation to either the number of correct responses using the computer administration \((r = .15; ns)\) or the number of error responses using the computer administration \((r = -.27; ns)\). It should be noted that this analysis should be accepted with caution due to a lack of variability. Specifically, whilst scores on the GCAS scale within the current study ranged between 43 and 73; the majority of participants in the current study had positive attitudes towards computers. Scores on the GCAS scale can theoretically range between 16 and 80, with scores above 48 indicating positive attitudes towards computers (Weber et al., 2003).

Results indicated that computerised administration was the most ‘fun’ (see Table 2). However, when asked about practical factors of the preferred method, the paper-and-pencil version was easiest to concentrate on, preferred for future tasks, easier generally, and the most comfortable. Opinions appeared to be divided as to which method, if any, the time went faster with.

<table>
<thead>
<tr>
<th>Method Preference Questionnaire</th>
<th>Paper and Pencil</th>
<th>Computerised</th>
<th>No Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Easiest to concentrate on</td>
<td>67.5</td>
<td>7.5</td>
<td>25</td>
</tr>
<tr>
<td>Most pleasant</td>
<td>77.5</td>
<td>12.5</td>
<td>10</td>
</tr>
<tr>
<td>Preferable for similar tasks</td>
<td>75</td>
<td>15</td>
<td>10</td>
</tr>
<tr>
<td>Easiest</td>
<td>92.5</td>
<td>5</td>
<td>2.5</td>
</tr>
<tr>
<td>Preferable for job interview tasks</td>
<td>80</td>
<td>7.5</td>
<td>12.5</td>
</tr>
<tr>
<td>Most fun</td>
<td>15</td>
<td>77.5</td>
<td>7.5</td>
</tr>
<tr>
<td>Most comfortable</td>
<td>77.5</td>
<td>2.5</td>
<td>20</td>
</tr>
<tr>
<td>Time went faster</td>
<td>45</td>
<td>32.5</td>
<td>22.5</td>
</tr>
</tbody>
</table>
4. Discussion

The majority of participants (70%) indicated that they were able to concentrate better on the paper-and-pencil (PAP) version of the BVRT. This was a surprising result given that computer technology is now a part of most people’s lives, both in the home and work environment. Accordingly, the majority of participants in the current study indicated that they were quite comfortable using computers and had experience of using computer technology. However, attention should be drawn to the fact that this refers to ‘subjective’ rather than ‘objective’ computer experience, and different individuals may have interpreted terms such as ‘comfortable’ differently.

The above findings are consistent with Merten’s (1999) study of Dutch speaking unselected neurological patients. Compared to the computerised version, the majority of participants rated the PAP version as more pleasant to use, easier to use, more comfortable to use, and preferable for use in future tasks of a similar nature. However, consistent with existing research in the area (Kurt et al., 2004; Weber et al., 2003), the computerised version of the conventional test received favourable evaluations in other areas. It was interesting to note that the computer version was typically considered as being more fun to use than the PAP.

In terms of test performance, despite controlling for which version was administered first, participants’ performances were significantly better on the PAP than on the computer version of the same test. This poorer performance on the computerised version of the BVRT is despite participants having previous experience with computers and having positive attitudes towards computers. Additionally, although education has been shown to interact with the success of the patient–computer interaction (Weber et al., 2003), this should not have been an influential factor within the current study although it was IQ that was assessed as opposed to educational attainment. It is possible that participants find the computer version less user-friendly in terms of the display, and more confusing to use in terms of the graphics tablet and mouse. This is also consistent with Merten’s (1999) study of neurological patients which is interesting since it seems that between these two studies both a non-clinical and a clinical population share performances, i.e. both populations declined when the computer version was used. Also, as the order of administration of test was randomised (i.e. computerised BVRT first versus paper-and-pencil version first), this confirms that the decline in performance on the computerised version is not simply attributable to fatigue effects.

Results of number correct and number error responses on the BVRT were examined for both versions of administration showing that the computer version yielded poorer performances. Hence, the computer version of the BVRT is less reliable than the PAP in a non-clinical population. Therefore, extreme caution should be exercised if it were to be used in a clinical population. In fact, the authors would advise against such use in its present format. This finding is in line with the opinion that performance on a computerised version of the test on visuospatial tests such as the BVRT is often significantly poorer than performance on a paper-and-pencil version of a test (Merten, 1999; Williams & McCord, in press). This reinforces the suggestion that computerisation of each test should be evaluated independently.

It may be that the method of input of participants’ responses via a graphics tablet is the key to the poor performance evidenced in this study. It is suggested that responding to presented stimuli by “drawing” directly onto a special computer touch screen, as was used by Weber et al. (2003), may be better as participants are able to see in front of them their immediate response rather than via a horizontal graphics tablet. However, the cost of this sort of display is approximately double the price of the device used in this study.
It should also be noted that the format used in the current computerised version of the test is comparable to that used in the conventional paper-and-pencil administration. Specifically, with the use of the graphics tablet, the drawing tablet is placed horizontally in front of the participant. This is comparable to within the paper-and-pencil administration, where participants draw onto a piece of paper in front of them. Moreover, although a representation of the drawing appears on the computer screen in front of the participant, the participant can also see what they are drawing when they look down at the graphics tablet that they are drawing on to. Thus, this is comparable to what the participant would see when looking at their piece of paper within the paper-and-pencil administration. In future examinations, giving participants a break should be considered, particularly in view of the fact that participants found the computer version more difficult to concentrate on. Weber et al. (2003) who allowed interruptions on demand during testing, found the computerised version of the test comparable to the paper-and-pencil version of the test.

To aid scoring participants’ responses, a computer program could be written to encode directly scores and hold in a database results from such testing. Although easier to administer for the examiner and more accurate to administer in terms of accurate timing of stimuli-display, it must be born in mind that impaired and perhaps older participants such as those found in a clinical population may find using such computer displays less easier to understand because of their lower computer literacy generally. It is suggested that this generation gap of computer literacy will fade as more and more people are familiar with technological advances in everyday situations. However, even in years to come, the particular characteristics and capabilities of each population will need to be borne in mind, especially when assessing within clinical situations.

Finally, it is suggested that whilst computer technology is no doubt a welcomed advance to most lifestyles, the human element should never be completely replaced in clinical situations where reassurance and the “bedside manner” provides empathy for the clinically challenged individual.

5. Conclusions

Evidence was found to suggest that participants from a non-clinical population perform less well on a computerised version of the Benton visual retention test as compared with the paper-and-pencil version. Participants rated the computer version as more fun but less easy to use. Hence, the authors concluded that the conventional administration is a more reliable and more accurate method of assessment administration.

Acknowledgement

The authors are grateful to Dr. G. Simons for her translation skills.

References


