A Comparison of Computer-Administered Language Assessments and Traditional Methods of Testing

Kadia Acres[1], Carolyn Bruce[2] and Caroline Newton[2]

[1] Hertfordshire Community NHS Trust, [2] Division of Psychology and Language Sciences, UCL

UCL

Introduction

Computers have successfully been used to provide therapy for people with aphasia (e.g. Mortley et al., 2004), and they have the potential to improve efficiency, objectivity and consistency in language assessment.

They have been effectively used in cognitive psychology to assess cognitive skills in adults without aphasia and in assessment of language skills when teaching English as a second language. Scores are generally found to be highly correlated to those from traditional tests, but slightly lower (e.g., Williams & McCord, 2006). Until now, little research has investigated using computers for assessment with adults with aphasia.

The aims of this study were to investigate whether computers can be used to assess receptive language in people with aphasia by comparing the equivalence of scores, the efficiency of test administration and participants' perceptions of the methods.

Methods

Participants
15 people with aphasia of varying severity were recruited from a community clinic.

Tasks
- Auditory sentence-picture matching task, using 17 items from the auditory sentence comprehension test in the Comprehensive Aphasia Test (Swinburn et al., 2004) and 13 items from the Test of Reception of Grammar (Bishop, 2003)
- Auditory grammaticality judgement task, using stimuli from McDonald (2000), e.g. "The girl *is writing/write* a letter to her mother"

Four versions of each task were constructed, either by dividing the stimuli evenly (giving 46 items in each grammaticality task) or by generating matching sentences (giving 30 items in each sentence-picture matching task).

Conditions
Each participant completed assessments in four conditions:
1. Traditionally administered tests (pen and paper)
2. Tests administered entirely by computer (computer only)
3. Tests administered by computer with a therapist in the room (computer and therapist)
4. A second version of the traditionally administered tests

Instructions were identical between conditions except for different instructions for how to respond. Start and end times of each condition were recorded by the therapist. The order of the conditions was counterbalanced.

Stimuli and instructions in the computer-administered conditions were presented on a PC and were controlled by the participant using a touchscreen. The participant could choose to repeat each stimulus once and could correct their response before moving on.

Questionnaires
Participants were asked to rate their experience of each method after completion of each condition. At the end of the experiment they completed a method preference questionnaire.

Results

Equivalence of scores between methods
Scores were highly correlated between the two pen and paper conditions and between pen and paper and computer-administered conditions for both tasks.

	Pen and paper	Computer only	Computer and therapist
Sentence comprehension	0.90**	0.94**	0.92**
Grammaticality judgement	0.75*	0.88**	0.84**

Table 1. Pearson's correlation coefficients for the relationships between pen and paper and the other conditions (*$p<.01$; **$p<.001$).

Scores for the sentence comprehension task were significantly lower in the computer alone condition than in the other two conditions.

Scores were significantly lower for the grammaticality judgement task in the computer and therapist condition than in the pen and paper condition.
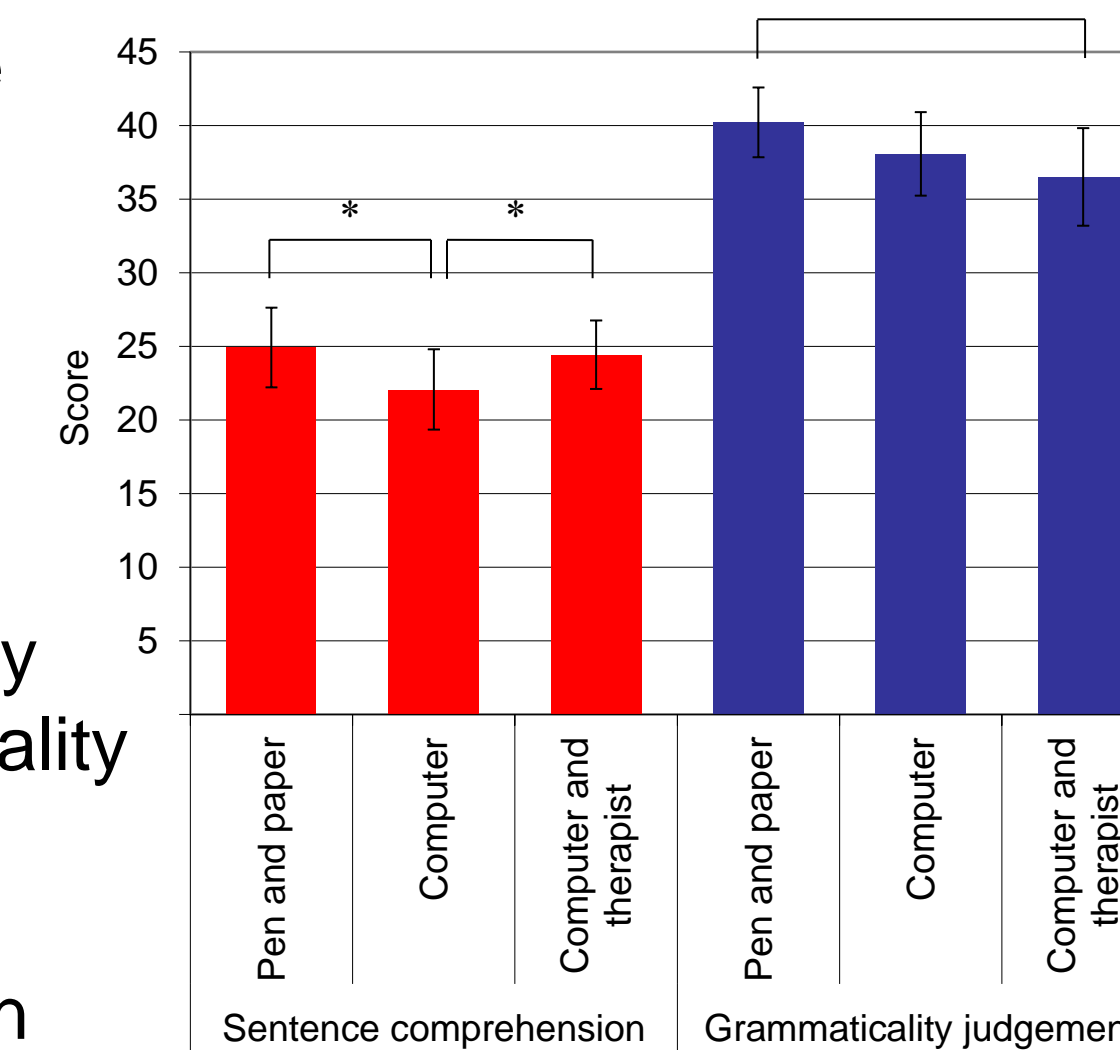

Figure 1. Mean scores for the three tasks in each conditions (error bars show 95% confidence intervals)
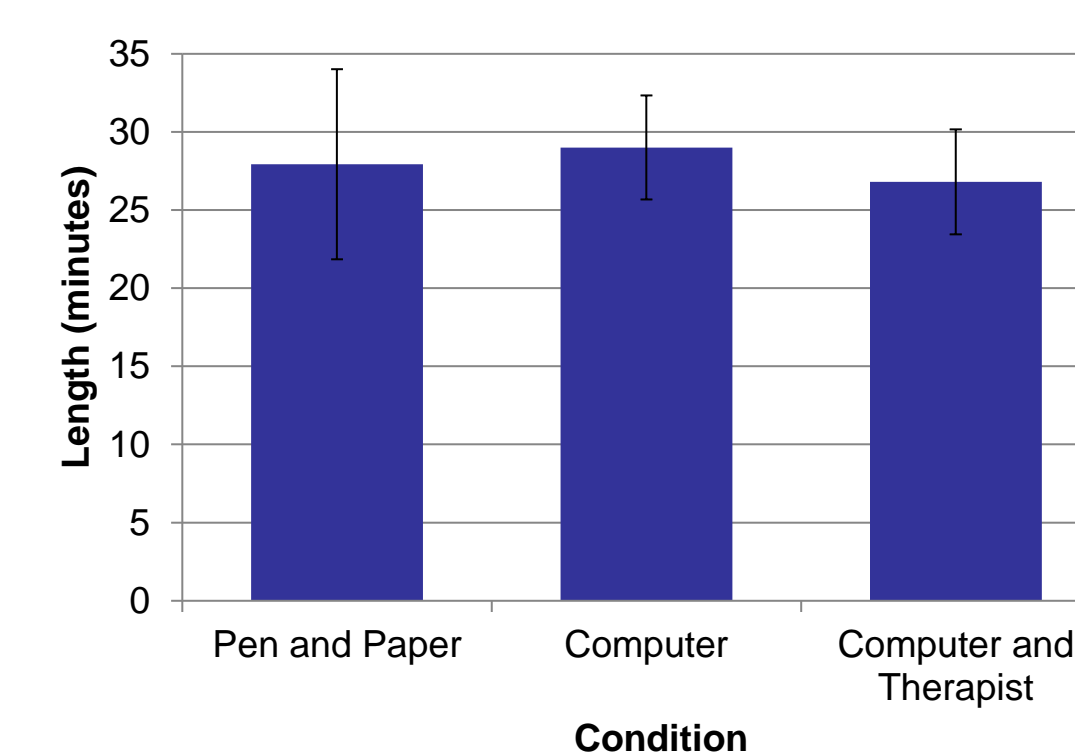
Efficiency of test administration

There were no significant differences between the times taken to administer the three conditions.

Figure 2. The average time taken to complete each condition

Participants' Perceptions
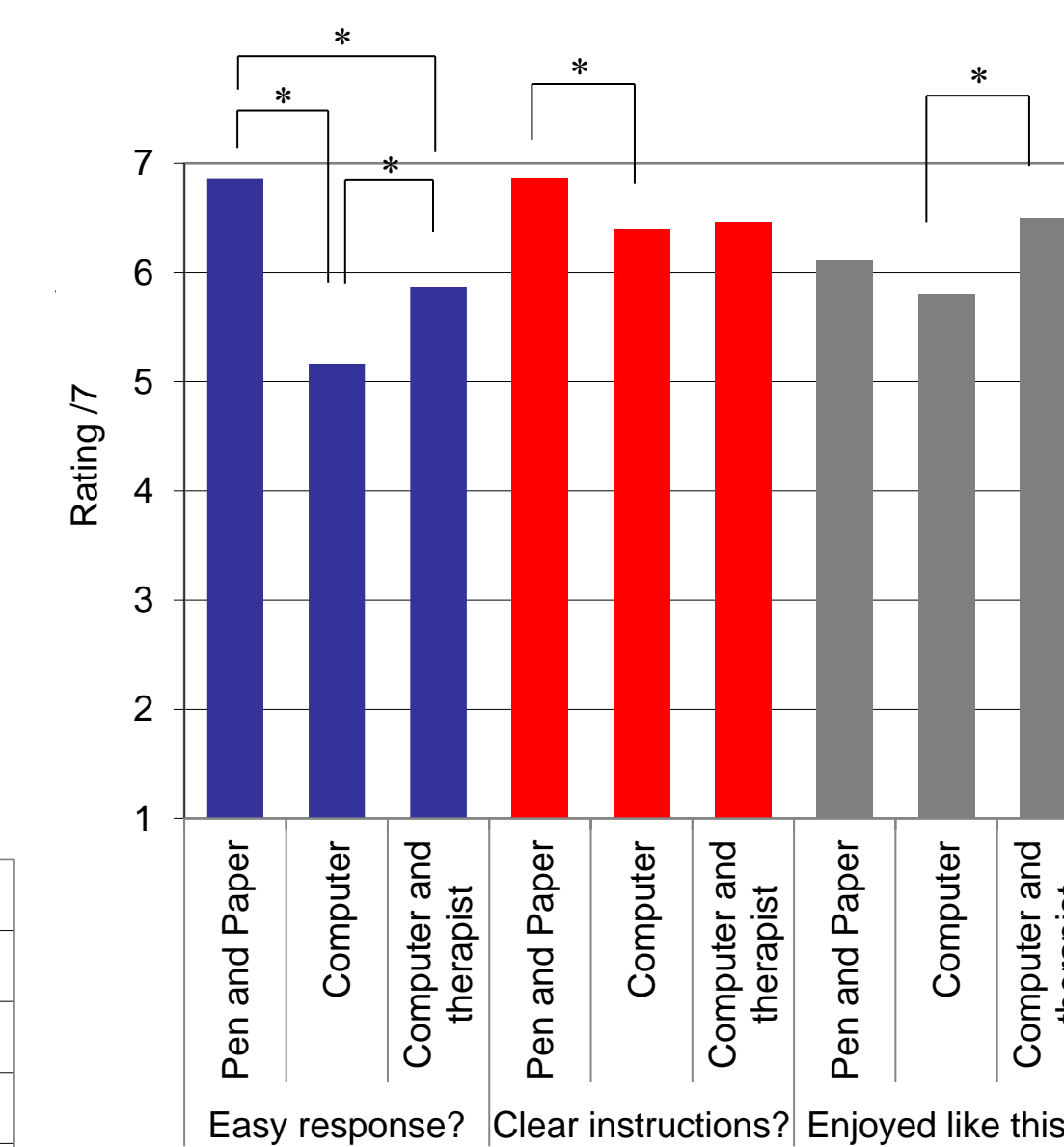
Figure 3: Average ratings (/7) for the participants' perceptions of the conditions

Participants rated ease of response significantly higher for the pen and paper condition than both computer-administered conditions. They also rated ease of response higher in the computer and therapist than the computer condition.

Clarity of instructions was also rated significantly higher for the pen and paper than the computer alone condition despite identical wording.

Enjoyment was rated higher in the computer and therapist condition than the computer alone condition.

Participants chose their favourite and least favourite conditions. More chose conditions with the therapist in the room as best than as worst, and more chose the computer only condition as worst.
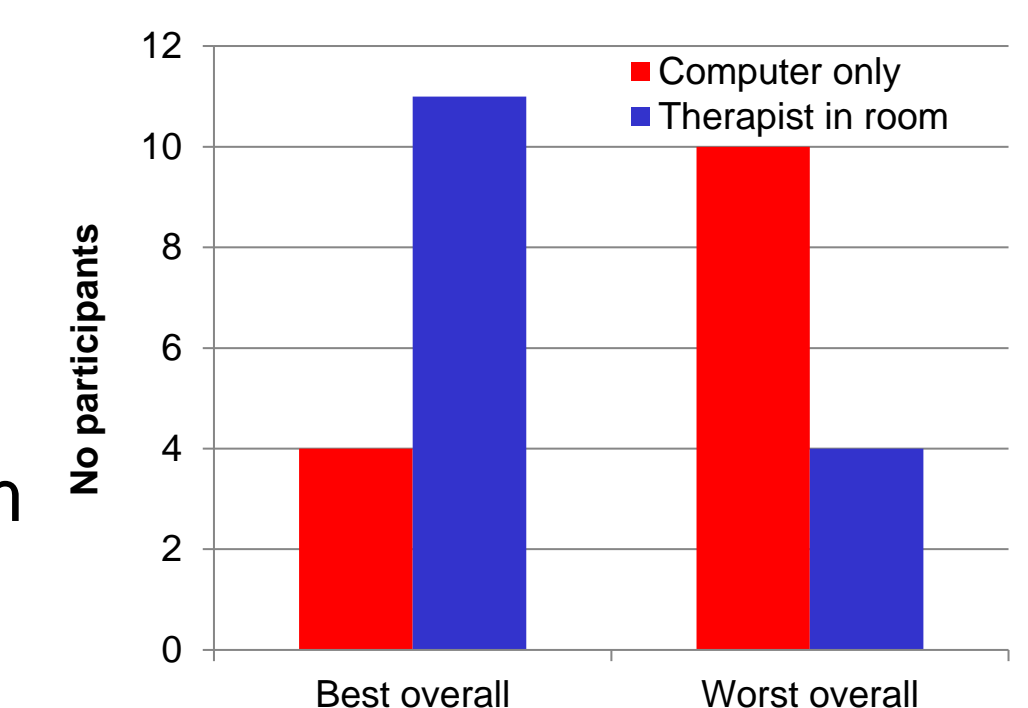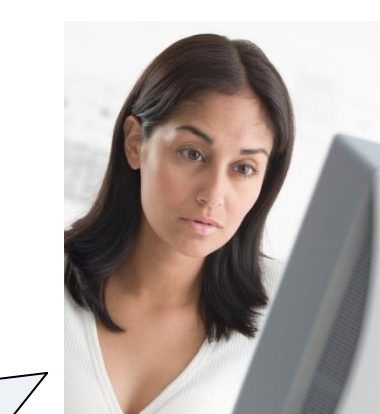

Figure 4: The effect of computer delivery on participants choosing the conditions as best or worst

Participants' Comments

I like human contact . . . looking at the person in front of me . . . the computer is impersonal . .

I don't have a computer . . . scared to repeat or spoil sometimes . . . Never-ending . . . makes me feel nervous because I don't normally use them.

You've always got a little bit of help in case something goes wrong. . . Computers don't always work right . . . If you're on your own you don't even know if you're doing it right.

I'm used to the computer . . .[prompt – why does that make it better?] . . . I'm in control, I understand it.

It's my time alone. I can do it fast or slow, you know. When I'm working I just want free time.

Conclusions

- Computers can successfully be used to assess language, but caution should be used as scores may not be equivalent to those gained in traditional administration. This may impact comparison to published norms.

- Using computers is not quicker, but it is potentially more efficient as when the test is computer-delivered the therapist could be doing other tasks while inside or outside the test room.

- Participants varied in their reaction to the computer-administered tests, with evidence that they enjoyed them less, found them harder to follow and found the therapist's presence reassuring. Therefore, computer-administered assessment may be more suitable for some clients than others.

References
Bishop, D. V. M., 2003. Test for Reception of Grammar, London, The Psychological Corporation.
Mcdonald, J. L., 2000. Grammaticality judgments in a second language: Influences of age of acquisition and native language. Applied Psycholinguistics, 21, 395-423.
Mortley, J., Wade, J., Enderby, P. & Hughes, A. (2004). Effectiveness of computerised rehabilitation for long-term aphasia: a case series study. British Journal of General Practice, 54, 856-857

Swinburn, K., Porter, G. & Howard, D., 2004. Comprehensive Aphasia Test, Hove, Psychology Press.
Williams, J. E. & McCord, D. M., 2006. Equivalence of standard and computerized versions of the Raven Progressive Matrices Test. Computers in Human Behavior, 22, 791-800.