Age-related decline in Digit–Symbol performance: 
Eye-movement and video analysis

R. Stephens*
School of Psychology, Keele University, Staffordshire ST5 5BG, UK
Accepted 10 August 2005

Abstract
Examines age-related decline in Digit–Symbol performance using variables obtained from a slow-motion analysis of a first person perspective video filmed during test completion, including superimposed cross-hairs indicating eye movements. Standard WAIS-3 DSCT scores and the video-derived variables were compared across two age groups (mean age 20 years vs. mean age 59 years). The older group performed more poorly overall, \( t(16) = -2.359, p = .031 \). The correlation between writing time per item and overall performance was (negatively) larger in the older group compared with the younger group, \( z = -2.180, p = .014 \). There was no difference between the groups’ correlation coefficients with respect to key search latency and overall performance, \( z = -0.064, p = .525 \). Overall these results suggest that characterisation of the age-related slowing on Digit–Symbol tests as a psychomotor deficit is appropriate.

Keywords: Eye-movements; Age-differences; Cognitive-processes; Motor-processes

Debate continues as to the psychometric properties of Digit–Symbol tests, such as the Wechsler Adult Intelligence Scale III Digit–Symbol Coding Test, or WAIS-3 DSCT (Wechsler, 1997) and its forerunner, the Wechsler Adult Intelligence Scale—Revised Digit–Symbol test, or WAIS-R DST (Wechsler, 1981). The paper-based WAIS-3 DSCT has a key area that consists of a \( 2 \times 9 \) matrix. Its first row contains the digits 1–9 numerically ordered, and its second row pairs each of these digits with a unique symbol, e.g. ‘1’ is paired with ‘–’ and 8 is paired with ‘x’. Below is a response area that consists of seven \( 2 \times 20 \) matrices. The first row of each contains the digits 1–9 in pseudo random order, with repetitions, and the second row contains empty spaces. Participants are given 120 s to enter as many symbols as possible in the spaces below the digits according to the key area pairings. The score is the number of correctly entered symbols.

A marked and robust age-related decline occurs in Digit–Symbol performance from around the age of 45 (e.g. Hoyer, Stawski, Wasylyshyn, & Verhaeghen, 2004). The Digit–Symbol task has been characterised as multifactorial, with motor response speed, cognitive speed and memory components (e.g. Joy, Kaplan, & Fein, 2004). It is unclear whether all components of Digit–Symbol performance decline similarly with age in the form of a non-specific ageing effect, or whether these performance components decline at differential rates with age. This raises a significant issue – whether the psychometric profile of the Digit–Symbol test varies depending upon the population to which it is applied – with important consequences for interpretation of test performance.

* Tel.: +44 1782 583600; fax: +44 1782 583387.
E-mail address: r.stephens@psy.keele.ac.uk.

© 2005 National Academy of Neuropsychology. Published by Elsevier Ltd. All rights reserved.

Keywords: Eye-movements; Age-differences; Cognitive-processes; Motor-processes

Debate continues as to the psychometric properties of Digit–Symbol tests, such as the Wechsler Adult Intelligence Scale III Digit–Symbol Coding Test, or WAIS-3 DSCT (Wechsler, 1997) and its forerunner, the Wechsler Adult Intelligence Scale—Revised Digit–Symbol test, or WAIS-R DST (Wechsler, 1981). The paper-based WAIS-3 DSCT has a key area that consists of a \( 2 \times 9 \) matrix. Its first row contains the digits 1–9 numerically ordered, and its second row pairs each of these digits with a unique symbol, e.g. ‘1’ is paired with ‘–’ and 8 is paired with ‘x’. Below is a response area that consists of seven \( 2 \times 20 \) matrices. The first row of each contains the digits 1–9 in pseudo random order, with repetitions, and the second row contains empty spaces. Participants are given 120 s to enter as many symbols as possible in the spaces below the digits according to the key area pairings. The score is the number of correctly entered symbols.

A marked and robust age-related decline occurs in Digit–Symbol performance from around the age of 45 (e.g. Hoyer, Stawski, Wasylyshyn, & Verhaeghen, 2004). The Digit–Symbol task has been characterised as multifactorial, with motor response speed, cognitive speed and memory components (e.g. Joy, Kaplan, & Fein, 2004). It is unclear whether all components of Digit–Symbol performance decline similarly with age in the form of a non-specific ageing effect, or whether these performance components decline at differential rates with age. This raises a significant issue – whether the psychometric profile of the Digit–Symbol test varies depending upon the population to which it is applied – with important consequences for interpretation of test performance.
One common explanation for older adults’ poorer Digit–Symbol performance is that of cognitive slowing. Originally advanced by Birren (1964), Salthouse (1996) presented a case that slowed cognitive processing accounts for some of the decrement observed in measures of Type A or fluid cognition – of which Digit–Symbol is a classic example. Salthouse based his argument on data demonstrating that different speeded fluid intelligence measures have a considerable amount of shared age-related variance. However, this argument rests on the explicit assumption that the speeded fluid intelligence tests employed had sufficient cognitive components so as not merely to represent input (sensory) and output (motor) processes, but Salthouse did not present any supporting evidence that this was the case. Laux and Lane (1985) carried out a task-analysis of the Symbol Digit test. This is similar to Digit Symbol except the role of digits and symbols is reversed so that participants respond by writing digits. Adapting their analysis, the Digit–Symbol task may be de-constructed into the following sub-tasks: (a) detecting and encoding the digit; (b) finding the digit in the key area; (c) encoding the symbol paired with the digit; (d) selecting the proper response; and (e) initiating and executing that response. A priori, (a), (b), (c) and (d) comprise the more cognate (not sensory or motoric) aspects of Digit–Symbol performance. Assessing directly the speed at which any of these components are performed and comparing the magnitudes of the correlations between these measures and overall Digit–Symbol performance for older and younger samples would elucidate the importance of cognitive speed to the age-related decline in Digit–Symbol performance.

The Digit–Symbol test was initially developed to assess memory for paired associates. Indeed, incidental learning (by recall) of the digit–symbol paired associates correlates with overall performance in younger (Joy, Fein, & Kaplan, 2003) and older adults (Joy et al., 2004). However, while showing that learned paired associations are formed, there is no evidence they are actually used while completing Digit–Symbol tests. Completing an item without using the key would indicate the use of learned paired associations during Digit–Symbol tests. Indeed, the extent to which participants generally make use of the key during Digit–Symbol tests has been questioned (Piccinin & Rabbitt, 1999). Eye-movement analysis indicated that younger adults make use of the key for most items (Stephens & Sreenivasan, 2002) but there is no corresponding data for older adults. Assessing the number of items completed without using the key, and comparing the magnitudes of the correlations between this measure and overall Digit–Symbol performance for older and younger samples would elucidate the importance of paired associate learning to the age-related decline in Digit–Symbol performance.

Symbol Copy speed and Digit Symbol correlate (see Table 1). Furthermore, Joy et al. (2004) observed an increase in the magnitude of the correlation between Symbol Copy and Digit Symbol across seven decades of age in the WAIS-3 standardisation sample and a similar trend is apparent across the correlational studies listed in Table 1. As there is no key area in Symbol Copy (the task is simply to copy symbols from the upper row of each response matrix) this correlation has been interpreted as showing that motor response speed is important to Digit–Symbol performance and to the age-related decline therein. However, some cognitive processing appears a priori to be required in Symbol Copy, for example, encoding the to-be-copied symbols, begging the question as to whether non-motoric elements of Symbol Copy underlie the correlation with Digit Symbol. Employing a direct measure of motor response speed during Digit–Symbol performance and comparing the magnitudes of the correlations between this measure and overall Digit–Symbol performance for older and younger samples would elucidate the importance of motor response speed to the age-related decline in Digit–Symbol performance.

Several authors (e.g. Joy et al., 2003; Piccinin & Rabbitt, 1999) have called for analysis of eye-movements during Digit Symbol completion, and a number of pilot studies have now been carried out (Charness & Schultetus, 1998; Stephens, 1997, 2001). Stephens and Sreenivasan (2002), having first verified that wearing eye-movement measuring equipment did not affect performance, analysed eye-movements and writing time during completion of the WAIS-R...

### Table 1

<table>
<thead>
<tr>
<th>Study</th>
<th>Participants</th>
<th>Correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Joy et al. (2003)</td>
<td>Mean age 20.1 years; n=87</td>
<td>0.59</td>
</tr>
<tr>
<td>Crowe et al. (1999)</td>
<td>Mean age 25.8 years; n=102</td>
<td>0.27</td>
</tr>
<tr>
<td>Joy et al. (2004)</td>
<td>Mean age approx 48 years; n=1167</td>
<td>0.82</td>
</tr>
<tr>
<td>Kreiner and Ryan (2001)</td>
<td>Mean age 51.7 years; n=228</td>
<td>0.82</td>
</tr>
<tr>
<td>Joy, Fein, Kaplan, and Freedman (2000)</td>
<td>Mean age 68.3 years; n=177</td>
<td>0.72</td>
</tr>
</tbody>
</table>
Digit–Symbol test in a sample of seven well-educated adults aged around 30 years. Writing time was not related to overall WAIS-R DST performance (the correlation coefficient was $-0.19$), but time spent inspecting the key area did predict overall performance (the correlation coefficient was $-0.68$). This was interpreted as indicating that speed of searching and encoding information – more cognate task requirements – were important for Digit–Symbol performance in that population.

The objective of the present study was to examine the decline in Digit–Symbol performance with increasing age using this novel methodology in a two-group (younger vs. older) design. Variables obtained from a frame-by-frame analysis of a first person perspective video filmed during completion of the second row of items of the WAIS-3 DSCT, including superimposed cross-hairs indicating eye movements, were examined. Two of these variables arguably reflect aspects of cognitive processing speed: time spent inspecting the key area per item, and the number of items where the key area was inspected two or more times. A third variable indicates utilisation of learned paired associations: the number of items completed without inspecting the key. A fourth variable indicates motor speed: response symbol writing time per item.

Older adults were predicted to obtain reliably lower scores on the WAIS-3 DSCT than younger adults. The hypothesis of non-specific ageing effects predicts similar magnitudes, across the groups, of the correlation coefficients describing the relationships between the video-derived variables and overall Digit–Symbol performance. However, if age causes Digit–Symbol performance to become more sensitive to any of the specific processes: processing speed, memory for paired associations or motor speed, then this predicts inequalities, across the groups, in the correlation coefficients describing the relationships between the video-derived variables and overall performance. On the other hand, absolute differences in the video-derived variables across the age-groups have little interpretative value from the perspectives of general versus specific ageing effects, as both positions predict some degree of inter-group difference.

1. Method

1.1. Participants

Two groups of native English speakers were recruited via Keele University campus and the Newcastle Volunteer Bureau. The Younger group comprised 9 undergraduates of mean age 20 years (S.D. 2.7), including four females. The older group comprised 9 adults of mean age 59 years (S.D. 8.8), including eight females. The groups differed on years of education, $\chi^2 (2) = 9.000, p = .011$. All younger group members were educated beyond age 18, whereas only three older group members were equivalently educated, with three educated to 18 years and three to 16 years. The groups also differed on NART-predicted WAIS Full Scale IQ scores. The older group’s mean of 119 (S.D. 4.1) was 13 points superior to that of the younger group (S.D. 3.0), $t(16) = -7.502, p < .001$. The younger participants were classified using the five-class version of the National Statistics Socio-economic Classification (Rose & Pevalin, 2003) according to the occupation of the main breadwinner in their family home, while the older participants were classified in their own right. The classifications were: managerial/professional – Yng 4, Old 3; intermediate – Yng 1, Old 2; small employers/own account workers – Yng 1, Old 0; lower supervisory/technical – Yng 3, Old 1; and semi-routine/routine – Yng 0, Old 3. The classifications did not differ across the groups, $\chi^2 (4) = 5.476, p = .242$. All participants were fit and healthy and none reported any illness likely to disrupt performance.

1.2. Stimuli

The WAIS-3 DSCT and the National Adult Reading Test (NART, Nelson & Wilson, 1991) were employed. The latter was included to enable group-comparisons on intellectual functioning (Crawford, 1992).

1.3. Apparatus

The Applied Science Laboratories Model 501 Eye Tracking System was used. This comprises a control box and a headband with two cameras, one trained on the left eye via a reflective visor, and the other (the ‘scene camera’) trained outwards, recording the participant’s field of view. Utilising the double Purkinje method, the moving image from the scene camera is outputted with a superimposed cross hair indicating the participants’ real-time direction of gaze within the visual field. Accuracy was to within 1 degree (Applied Science Laboratories, 2001), translating to a distance of...
8 mm at the test paper surface. The output was video-recorded at resolution 25 frames-per-second and digitised for analysis on PC using the Adobe Premier software package (version 6.0).

1.4. Design

The study should be conceptualised as being in two parts. For the first part, a quasi-experimental between-subjects design was applied in which the standard WAIS-3 DSCT scores of two independent groups made up of younger and older individuals were compared. For the second part, a correlational design was employed. The correlations between the video-derived performance indicators and the standard WAIS-3 DSCT scores were calculated for each of the two independent age groups, and the magnitudes of these correlations were compared across the groups.

1.5. Procedure

Participants were tested individually in the standardised conditions of a windowless psychology research laboratory. First, participants provided demographic data and completed the NART. Next, the eye tracker device was set-up, video recording commenced and participants completed the WAIS-3 DSCT following the standard instructions. For each item in the second row of items of the WAIS-3 DSCT, specific eye-movements and writing time were recorded via frame-by-frame analysis of the eye-tracker video output. Key area inspection latency was the sum of the whole number of frames that direction of gaze was within the confines of the $2 \times 9$ key matrix. Zero and multiple key area checks were defined by the number of incidences of eye-gaze direction entering and exiting the key matrix per item. Writing time was the sum of the whole number of frames that the pencil was in contact with the paper and producing a mark that contributed towards production of a symbol. The study was approved by the Keele University Psychology Department Research Ethics Committee.

2. Results

Tables 2 and 3 summarise the data. The groups’ overall WAIS-3 DSCT scores were compared using a $t$-test. Differences in the magnitudes of the correlations were compared across the groups using Fisher’s $z$ transformation of
The older group completed reliably fewer WAIS-3 DSCT items overall than the younger group, \( t(16) = -2.359, p = .031 \). The correlation between writing time per item and overall performance was (negatively) larger in the older group, \( z = -2.180, p = .014, q = -1.259 \). However, the groups did not differ in the magnitudes of their respective correlation coefficients for key inspection latency and overall performance, \( z = -0.064, p = .525, q = -0.037 \), for the number of items where the key area was searched two or more times and overall performance, \( z = -1.290, p = .098, q = -0.745 \), or for the number of items with zero key area checks and overall performance, \( z = 0.571, p = .716, q = 0.330 \).

3. Discussion

The older group performed the WAIS-3 DSCT more poorly than the younger group as predicted by age-related slowing. Where a quasi-experimental design is applied, an understanding of the distribution across experimental groups of variables known to covary with the dependant variable is necessary. The older group’s poorer performance could be attributed to their poorer educational level relative to the younger group. However, any effect of educational level would be countered by the older groups’ superior intellectual functioning according to its NART-derived WAIS Full Scale IQ scores. Furthermore, there were no differences in health status between the groups and no differences in socio-economic grouping. Therefore, the groups appear to be broadly balanced with respect to background factors.

The older group contained a greater proportion of females, which would predict improved performance in that group. That age effects were present despite this inequivalence demonstrates the robustness of the age-related decrement generally observed on Symbol–Digit tests.

There was a reliable negative correlation between key inspection latency and the standard WAIS-3 DSCT score for all participants suggesting that cognitive speed plays a role in Digit–Symbol test performance regardless of age. This conclusion is supported by the analysis of the difference in the magnitude of the correlation coefficients for this relationship for each of the age groups. No reliable difference was observed and the effect size indicated that any group difference was very small, tending towards zero. Taken together, these two findings are consistent with the notion that the Digit–Symbol test discriminates between participants of all ages at the level of cognitive processing speed without there being any pronounced slowing that is especially characteristic of the age-related slowing of Digit–Symbol performance.

There was a negative correlation between writing time and the standard WAIS-3 DSCT score for all participants. However, this was largely due to the influence of the older participants who exhibited a reliable and marked negative relationship between writing time and overall WAIS-3 DSCT performance. In contrast, there was no corresponding correlation for the younger group and there was a reliable difference in the magnitude of the groups’ correlation coefficients with respect to this relationship. These results appear unambiguously to support a significant role for slowed motor speed in the age-related decline in Digit Symbol performance, and support the interpretation of Symbol Copy as primarily a motoric task (e.g., Joy et al., 2004).

Piccinin and Rabbitt (1999) queried whether excessive carefulness slows down older individuals on Digit–Symbol tests. One could interpret re-inspecting the key as a manifestation of carefulness (i.e. double-checking, or ‘making sure’). However, there was no significant correlation between the number of items requiring 2+ key inspections and overall WAIS-3 DSCT performance and no inter-group difference in the magnitudes of the correlation coefficients for this relationship. Therefore, there was no evidence linking excessive carefulness with the age-related decline in Digit–Symbol performance. It was assumed that the number of items completed without inspection of the key area would serve as an indicator of the use of memory for the paired associations during completion of Digit–Symbol tests. However, there was no significant correlation between the number of items completed without inspection of the key area and overall WAIS-3 DSCT performance and no inter-group difference in the magnitudes of the correlation coefficients for this relationship. Therefore, there was no evidence that differential use of memory for paired associates in younger and older individuals contributes to the age-related decrement in Digit–Symbol performance.

Digit–Symbol tests are speeded, require information from the visual environment to be encoded and processed, and set goals in terms of the fast turnover of items requiring completion. Just and Carpenter (1976) argued that under such conditions, eye movement patterns are likely to reflect ongoing cognitive processes. To ensure that what is a labour-intensive method of data acquisition was manageable, only the second row of 20 items was analysed. This was
justified on the basis of research showing that the rate of performance of older adults does not show a warm-up pattern of improvement, even as the task continues (Paulo & Ryan, 1994).

Measuring eye-movements has yielded information concerning how people go about completing Digit–Symbol tests. Overall, most responses (93%) were made with reference to the key, perhaps supporting the theoretical argument made by Joy et al. (2003) that serial memory search processes render use of memory for the pairings detrimental to performance in comparison with referencing the coding key. Piccinin and Rabbit (1999) suggested that eye-tracker data could elucidate whether the time spent searching the key is small relative to the time required to look between the key and the response area, a contention first raised by Salthouse (1978). Although not formally analysed, in the current study the time spent searching the key added up to substantive periods of activity, whereas the time required to make a saccade from the key area back to the response area was very rapid, taking only one or two frames (i.e. 0.08 s or under). On the basis of these observations, therefore, the above contention appears to be incorrect.

It has been previously acknowledged that different underlying abilities may be important to Digit–Symbol performance in different populations (e.g. Laux & Lane, 1985, p. 134), that is, that the cognitive specificity of the Digit–Symbol test may vary according to the population. These data support such a change in task specificity with age. Older individuals performing Digit–Symbol tests most successfully are likely to exhibit faster motor response speed. On the other hand, based on our previous work (Stephens & Sreenivasan, 2002), younger adults performing Digit–Symbol tests most successfully are likely to be the fastest information processors, with motor speed not being a comparable discriminatory factor. Taken as a whole, the results of both studies suggest that characterisation of the Digit–Symbol task as a test of psychomotor speed is appropriate for older but not for younger participants.

Acknowledgments

This study was supported by a grant from the Wellcome Trust, Ref: VS/03/KEE/2/SL/TH/FH. Thanks to Gillian Mackay for data collection.

References


