The Letter Digit Substitution Test: Demographic Influences and Regression-Based Normative Data for School-Aged Children

W. Van der Elst1,*, S. Dekker1, P. Hurks2, J. Jolles1

1Department of Educational Neuroscience, Faculty of Psychology and Education and LEARN! Institute, VU University Amsterdam, Amsterdam, The Netherlands
2Faculty of Psychology and Neuroscience, Department of Neuropsychology and Psychopharmacology, Maastricht University, The Netherlands

*Corresponding author at: Department of Educational Neuroscience, Faculty of Psychology and Education and LEARN! Institute, VU University Amsterdam, Van der Boechorststraat 1, 1081 BT Amsterdam, The Netherlands.
E-mail address: wim.vanderelst@gmail.com (W. Van der Elst).

Accepted 11 April 2012

Abstract

The Letter Digit Substitution Test (LDST) was administered to a sample of N = 296 healthy children (aged between 8.03 and 15.87). The aim of the present study was to evaluate the impact of age, gender, and parental educational level on LDST performance and to establish demographically corrected normative data. The results showed that the relationship between age and LDST performance was curvilinear (i.e., improvements in test performance were more pronounced for younger children than for older children) and was moderated by gender (i.e., the gender differences were small at younger ages but increased as a function of age, with girls outperforming boys). Moreover, children who had parents with a higher level of education outperformed their counterparts who had parents with a lower level of education. Regression-based normative LDST data were established, and an automatic scoring program was provided.

Keywords: Information processing; Continuous norms; Gender differences; Educational differences; Development

Introduction

Substitution tests are speed-dependent cognitive tasks that require the participants to match particular signs (e.g., symbols, digits, or letters) to other signs (Lezak, Howieson, & Loring, 2004). Substitution tests are mainly used to assess information processing speed, a cognitive ability that reflects the speed by which elementary cognitive operations can be performed. Information processing speed is a fundamental cognitive ability in neuropsychological theories on “normal” cognitive development and “normal” cognitive aging (Fry & Hale, 1996, 2000; Kail, 2000, 2007; Salthouse, 1996; Sheppard & Vernon, 2008). For example, developmental studies have shown that age-related improvements in processing speed result in an increased working memory span, better reasoning abilities, and more efficient long-term memory processes (Kail, 2000). Similarly, an age-related reduction in information processing speed has been associated with a deterioration in the working memory and long-term episodic memory processes of older people (Salthouse, 1996). Information processing speed is also affected in a number of clinical conditions in children and adults, such as attention deficit hyperactivity disorder, depression, pediatric traumatic brain injury, autism, Alzheimer’s disease, multiple sclerosis, and schizophrenia (Calhoun & Mayes, 2005; DeLuca & Kalmar, 2008; Donders, 2006; Lezak et al., 2004; Spreen & Strauss, 1998; Weiler, Bernstein, Bellinger, & Waber, 2000).

In view of the central place of information processing speed in different cognitive theories and its sensitivity for a variety of clinical conditions, substitution tests have become popular assessment tools in both research and clinical settings (Lezak et al., 2004). The present study focuses on the Letter Digit Substitution Test (LDST; Jolles, Houx, Van Boxtel, & Ponds, 1995), which is an adaptation of earlier substitution tests—such as the Digit Symbol Substitution Test (DSST; Wechsler, 1955, 1981) and the Symbol Digit Modalities Test (SDMT; Smith, 1982). In the DSST, digits are presented and the test participant has to respond by writing down a symbol. This relation is reversed in the SDMT, that is, symbols are presented and the test...
participant has to respond by writing down a digit. The LDST differs from the DSST and the SDMT in that the key consists of well-known signs, that is, letters and digits (Jolles et al., 1995). The SDMT and the LDST have the advantage over the DSST that responses can be given in both writing and verbally. In this way, the SDMT and the LDST can also be administered to people who have motor problems that hamper writing (which is not possible with the DSST).

Normative LDST data for adults have been established earlier (Van der Elst, Van Boxtel, Van Breukelen, & Jolles, 2006a), but norms for children were lacking until now. In the present study, we administered the LDST to a sample of $N = 296$ children (aged between 8.03 and 15.87). Previous research in adults has suggested that age, gender, and educational level affected LDST performance (Van Boxtel et al., 1997; Van der Elst et al., 2006a). Thus, we also evaluated the impact of demographic variables on the LDST performance of the children in the normative sample and derived demographically corrected norms by means of a regression-based procedure.

**Method**

**Participants**

The data were derived from the COOS (Cognitief Ontwikkelings Onderzoek bij Schoolgaande kinderen, in English: Cognitive developmental study in school-aged children), a large scale study into “normal” cognitive development. All children came from public schools. Children who (a) had not repeated or skipped a grade, (b) did not use medication known to affect cognitive performance (such as Ritalin), and (c) did not suffer from severe medical conditions known to affect cognitive performance (such as epilepsy) were eligible for study inclusion. Medication use and health status were assessed by means of a parental-report questionnaire.

The LDST was individually administered to a sample of 296 children (at school). None of the children had significant motor or sensorial impairments that hampered administration of the LDST. Basic demographic data for the participants were provided in Table 1. The children were aged between 8.03 and 15.87. Age was used as a continuous variable in the normative analyses (see below), but it was categorized into bands of 2 years in Table 1 (for descriptive purposes). The educational level of the children’s parents (or caregivers) was measured with a commonly used Dutch educational 8-point rating scale which ranges from primary school to university degree (De Bie, 1987). Mean Level of Parental Education (MLPE) was dichotomized into Low and High groups (after a median split) for parents who had MLPE values that were <5 and ≥5 on the 8-point scale, respectively (with 5 = at most junior vocational education). These two levels of education corresponded with a mean (SD) of about 9.88 (2.59) and 14.68 (3.30) years of full-time education, respectively. All children were native Dutch speakers, and all parents (or caregivers) of the children gave consent for their child to participate in the research. The Ethics Committee of the Faculty of Psychology and Neuroscience of Maastricht University approved the study protocol.

**Procedure and Instruments**

The LDST sheet presents the substitution key on the top of the page. The key shows the numbers 1–9, which are all paired with a different letter. The test items are printed beneath the key. The children were instructed to write down the appropriate digits in the blank spaces under the letters. The first 10 items were used as practice items (to ensure that the children understood the test instructions). After completion of the practice items, the children were instructed to fill in the remaining items as quickly and as accurately as possible. It was not allowed to skip items. After 60 s, the test was stopped. The number of correct substitutions that were made in 60 s was the outcome variable.

<table>
<thead>
<tr>
<th>Age range (years)</th>
<th>$N$</th>
<th>Age ($M$ [SD])</th>
<th>Gender</th>
<th>Mean level of parental education</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;8.03 and ≤10</td>
<td>60</td>
<td>8.67 (0.57)</td>
<td>24</td>
<td>31</td>
</tr>
<tr>
<td>≥10 and ≤12</td>
<td>70</td>
<td>10.66 (0.53)</td>
<td>45</td>
<td>39</td>
</tr>
<tr>
<td>≥12 and ≤14</td>
<td>109</td>
<td>12.97 (0.54)</td>
<td>52</td>
<td>50</td>
</tr>
<tr>
<td>≥14 and ≤15.87</td>
<td>57</td>
<td>14.52 (0.35)</td>
<td>27</td>
<td>38</td>
</tr>
<tr>
<td>Total</td>
<td>296</td>
<td>11.85 (2.12)</td>
<td>148</td>
<td>158</td>
</tr>
</tbody>
</table>
Statistical Analyses

The LDST data were analyzed by means of a multiple linear regression model. Age, Age^2, Sex, MLPE, and all two-way interactions between these variables were used as predictors in the full regression model. Age was centered (=calendar age – mean age in the sample [=12 years]) before computing the quadratic age term (to avoid multicollinearity; Aiken & West, 1991). Both Age and Age^2 were added as predictors in the model to allow for a curvilinear relationship between age and the LDST scores. When only a linear age term is included in the model, the relation between age and the outcome is assumed to be linear over the entire age range. This is a strong assumption that may not be fulfilled, for example, it is possible that the age-related increase in LDST performance is larger for younger children than for older children. MLPE was dummy coded as 1 = high and 0 = low. Sex was dummy coded as 1 = men and 0 = women. The full regression model was reduced by means of a hierarchical step-down procedure in which the non-significant predictors were removed from the model (starting with the higher-order terms). The assumptions of regression analysis were tested by conducting a Kolmogorov–Smirnov test (to evaluate the normality of the residuals), by grouping the participants into quartiles of the predicted scores and applying the Levene test (to evaluate homoscedasticity), by calculating Variance Inflation Factors (VIFs) and condition indices (to identify multicollinearity), and by computing Cook’s distances (to identify influential cases). The Kolmogorov–Smirnov and the Levene tests should not be significant. The VIF and the condition index of a predictor should not exceed 10 and 15, respectively (Belsley, Kuh, & Welsch, 1980; Kutner, Nachtsheim, Neter, & Li, 2005). The maximum Cook’s distance value was related to an F(p, n-p) distribution (with p = the number of predictors in the regression model and n = the total number of observations). When the corresponding percentile value is about 50 or higher, the observation is influential (Kutner et al., 2005).

The normative data were established by means of a four-step procedure (Testa, Winicki, Pearlson, Gordon, & Schretlen, 2009; Van Breukelen & Vlaeyen, 2005; Van der Elst et al., 2006a, in press; Van der Elst, Van Boxtel, Van Breukelen, & Jolles, 2006b). First, the expected LDST score of a test participant was computed as based on the final multiple regression model that was established in the normative sample, that is, Expected test score = B_0 + B_1X_{1i} + ... + B_kX_{ki} (with B_0 = the intercept, B_k = the regression weights for the demographic variables, and X_{ki} = the values of the demographic variables for test participant i). Second, the residual was calculated, that is, e_i = observed test score - expected test score. Third, the residual was standardized, that is, Z_i = e_i/SD(residual). If heteroscedasticity occurred, the SD(residual) values were estimated by regressing the squared residuals (e_i^2) on the predicted values, and taking the square root of the obtained value (Kutner et al., 2005). Fourth, the standardized residuals were converted into percentile values via the standard normal cumulative distribution function.

The analyses were conducted with the “R” software package (version 2.11.1 for Linux), using a α level of 0.05.

Results

The LDST Regression Model

The final regression model for the LDST score is presented in Table 2. The Levene test suggested that the homoscedasticity assumption was violated (Levene statistic = 3.45, p = .02). The estimated SD(residual) function equaled \(\sqrt{(-6.09 + (0.962 \times \text{predicted score}_i))}\). The assumption of the normal distribution of the standardized residuals was not violated (Kolmogorov–Smirnov Z = 0.789, p = .56), multicollinearity was not a problem (the VIFs and the condition indices of the predictors had a maximum value of 2.41 and 4.12, respectively, and there were no influential cases (the maximum Cook’s distance value was 0.07; referring to an F_{0.290} distribution, this value corresponds to percentile 0.13).

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>SE B</th>
<th>Std. B</th>
<th>T</th>
<th>R^2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>32.52</td>
<td>0.55</td>
<td></td>
<td>58.55*</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>2.92</td>
<td>0.21</td>
<td>0.79</td>
<td>14.21*</td>
<td></td>
</tr>
<tr>
<td>Age^2</td>
<td>0.25</td>
<td>0.07</td>
<td>-0.14</td>
<td>-3.47*</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>-3.01</td>
<td>0.57</td>
<td>-0.19</td>
<td>-5.31*</td>
<td></td>
</tr>
<tr>
<td>MLPE</td>
<td>1.24</td>
<td>0.57</td>
<td>0.08</td>
<td>2.19*</td>
<td></td>
</tr>
<tr>
<td>Age × Gender</td>
<td>-0.62</td>
<td>0.27</td>
<td>-0.12</td>
<td>-2.34*</td>
<td>.63</td>
</tr>
</tbody>
</table>

Notes: Encoding of the predictors: Age = calendar age – 12; Age^2 = (calendar age – 12)^2. Gender: Women = 0, men = 1. Mean Level of Parental Education: Low = 0, high = 1. R^2 is adjusted for the number of predictors in the regression equation.

*p < .05.
As shown in Table 2, there was a significant interaction between age and gender on the LDST score. Fig. 1 shows the effects of age and gender on the expected LDST score graphically (for children who have parents with a low MLPE; the shape of these curves is identical for children who have parents with a high MLPE, up to a constant; see Table 2). As shown, the gender differences in the LDST scores were small for young children but increased as a function of age. For example, the test scores of 8-year-old boys and girls were approximately equal, but by the age of 15.5, girls scored on average 5.2 points higher (i.e., 15% higher) than boys.

In addition to age and sex, MLPE was found to significantly affect the LDST performance. Children who had parents with a high MLPE scored on average 1.24 points higher than children who had parents with a low MLPE. The regression model explained 63% of the variance in the LDST scores.

The Four-Step Normative Procedure

The LDST score of a tested child is normed by means of the four-step procedure described above. For example, suppose that an 8-year-old boy who had parents with a high MLPE made 15 correct substitutions in 60 s. The predictors to be used in the regression model are thus coded as follows: Age = −4 (=8 − 12), Age^2 = 16 (=([8 − 12]^2)), Gender = 1, MLPE = 1, and Age × Gender = −4 (=([8 − 12] × 1)). In the first step of the normative procedure, the expected LDST score for the tested child is computed, that is, expected test score = 32.52 + (−4 × 2.92) + (16 × −0.25) + (1 × −3.01) + (1 × 1.24) + (−4 × −0.62) = 17.55. Secondly, the residual is calculated, that is, −2.55 (=15 − 17.55). Thirdly, the residual is standardized. As the predicted LDST score for this child equaled 17.55, the SD(residual) value to be used equals 3.29 (=√(−6.09 + (0.962 × 17.55))). Thus, the standardized residual is −0.78 (=−2.55/3.29). Fourthly, the standardized residual is converted into a percentile value by means of the standard normal distribution. A standardized residual that equals −0.78 corresponds to a percentile value of 22. Thus, 22% of the “normal” population of 8-year-old boys who have parents with a high MLPE obtain an LDST score that is lower than 15. It can, thus, be concluded that the information processing ability of this child (as assessed with the LDST) is within normal limits.
User-Friendly Normative Data

The four-step normative procedure that was described in the previous paragraph lacks user-friendliness, because the user of the norms has to actively make the required computations. To increase the user-friendliness of the normative data, we provided a normative table that requires no computations at all (see Supplementary material online, Table S1). For example, Supplementary material online, Table S1 immediately shows that the LDST score equal to 15 that was obtained by the 8-year-old boy from the previous example corresponds to a percentile value between 20 and 30.

The simplified normative table is easy to use but lacks some accuracy because the child’s age has to be rounded off (if he or she is not exactly 8, 8.5, . . . , or 15.5 years old), and because only a limited number of percentile values can be presented in this table (because of space limitations). To maximize both the user-friendliness and the accuracy of the norms, the normative conversion procedure was implemented into an Excel worksheet (which can be downloaded at http://home.deds.nl/~wimvde/). The use of this worksheet is straightforward: The user simply types in the age, gender, and MLPE of the tested child together with his or her obtained LDST score, and the algorithm automatically computes the corresponding percentile value.

Discussion

The main aim of the present study was to establish the normal range of performance on the LDST for school-aged children. As a first step in the normative analyses, the influence of demographic variables on the LDST performance was evaluated. The results showed that the relationship between age and LDST performance was curvilinear (i.e., the relative increase in test performance was more pronounced for younger children than for older children; see Fig. 1). Moreover, the relation between age and LDST performance was moderated by gender (i.e., the age-related increase in test performance was more pronounced for girls than for boys). For example, the average relative increase in the LDST score of a 9-year-old child when compared with the test score of an 8-year-old child was 27.5% for girls and 24.5% for boys, while the average relative increase in the LDST score of a 15.5-year-old child when compared with the test score of a 14.5-year-old child was only 3.6% for girls and 2.3% for boys. These results are in line with previous studies (Camarata & Woodcock, 2006). The developmental increase in processing speed has been attributed to a variety of neuropsychological processes, such as an increase in the speed of stimulus identification and encoding, faster response selection, and a reduction in decision-making time (Miller & Vernon, 1997). At the neurological level, the age-related increase in information processing speed has been attributed to myelination processes and increases in the number of connections in the nervous system during childhood and early adolescence (Bjorklund & Harnishfeger, 1990; Kail, 2000; Kail & Salthouse, 1994). The advantage of girls in substitution test performance has been attributed to their superior verbal abilities (Fry & Hale, 2000; Joy, Kaplan, & Fein, 2004), which is at the neurological level reflected by a more bilateral pattern of activation of the frontal and temporal regions in girls (Burman, Bitan, & Booth, 2008). In addition to age and gender, MLPE was found to affect the LDST performance significantly. As expected, children who had parents with a high MLPE had higher LDST scores than children who had parents with a low MLPE. The total variance in the LDST score that was explained by all predictors together (i.e., age, age², gender, MLPE, and the age × gender interaction) equaled 63% and was thus high.

The LDST and its normative data provide a useful tool in applied and research settings when information processing speed needs to be assessed, but there are some limitations that warrant further discussion. First, all children who participated in the study were native Dutch speakers. The question arises whether the established normative data can also be applied to children who have a different native language or a different cultural background. An earlier study has shown that the LDST performance of adults who lived in eight European countries and the USA was comparable (Møller et al., 1998). This finding suggests that the LDST performance of children from different Western countries may also be similar, but empirical research is needed to substantiate this claim. In non-Western countries, the use of the normative LDST data is not recommended because the letters and/or digits may not be familiar symbols for these children and because of general cross-cultural differences which are known to affect performance on cognitive tests (such as degree of familiarity with formal testing; Ardila, 1995).

Second, the normative data that were established in the present study were based on a sample of children who were aged between 8.03 and 15.87. The question rises whether the regression model and its SD(residual) values that were established in the present study can also be used to norm the LDST scores of children whose age falls outside this age range. Extrapolating a regression model requires several assumptions. For example, if we want to estimate the expected test score of a 17-year-old child (as based on the regression model shown in Table 2), it has to be assumed that the relation between age and LDST performance in the age range between 15.87 and 17 is identical (or at least very similar) to the relation between age and LDST performance in the age range between 8.03 and 15.87. If this assumption is not fulfilled, the extrapolation of the regression model may result in biased conclusions (e.g., a child with a “normal” LDST score may be classified as being “impaired,” or vice versa). Since this assumption is not testable (as based on the present data set), it is not recommended to extrapolate...
the model (except perhaps when the value to be extrapolated is very near to the age limits that were considered in the normative sample, e.g., for a person aged 16 years).

On a related note, it is not recommended to administer the LDST to children who are younger than 8 years. Apart from the extrapolation problem that was described in the previous paragraph, a prerequisite of valid LDST administration is that the letters and digits are well-known symbols for the testee. By the age of 8 years, most children are sufficiently familiar with letters and numbers (as they had at least 1 year of practice in reading and writing at school), but this may not be the case for younger children (see also the SDMT manual, which also uses 8 years as the lower age limit for test administration; Smith, 1982).

Third, the LDST responses can be given both in writing and verbally. In the present study, the written LDST version was administered to all children. Previous research in adults has shown that the test scores that were obtained with the oral LDST version were on average higher than the scores that were obtained with the written LDST version (Van der Elst et al., 2006a). It is conceivable that similar effects of response modality also occur in children, and consequently, the normative data that were established in the present study should not be used when the oral version of the LDST is administered. Future studies should establish the normal range of performance for the oral LDST version in children, which may be especially useful in clinical settings (e.g., to assess the information processing speed of children with paresis or other motor problems that hamper writing).

Fourth, the sample was not balanced with respect to gender and MLPE in each of the age groups (e.g., there were \( n = 45 \) girls and \( n = 25 \) boys in the 10–12 years age group; see Table 1). This is, however, not a problem when a regression-based normative approach is used, because regression weights are estimated in an unbiased way no matter what the distribution of the predictors in the sample is (as long as the regression model is correctly specified, i.e., as long as no predictors or interaction terms are overlooked; Fox, 1997). Moreover, regression models use the entire sample for estimating and testing effects of the predictors (i.e., the analyses are not based on age subgroups, but treat age as a continuous variable). There is only some loss of statistical power when an unbalanced sample is analyzed, because the standard error of a regression weight is proportional to the \( \sqrt{\text{VIF}} \) of the predictor at hand (Kleinbaum, Kupper, Muller, & Nizam, 1998)—but this was not a problem in the present study because both gender and MLPE were significant.

**Supplementary Material**

Supplementary material is available at *Archives of Clinical Neuropsychology* online.

**Conflict of Interest**

None declared.

**References**


