Consistency of Handwriting in Early Elementary Students

Deborah Marr,
Sharon Cermak

KEY WORDS
• handwriting
• kindergarten
• longitudinal

OBJECTIVE. The purpose of this longitudinal study was to examine the consistency of handwriting in children from the beginning of kindergarten to the middle of the first-grade year. Consistency was defined as retaining the same qualitative performance and relative ranking over time.

METHOD. Ninety-three children were tested at the beginning of the kindergarten year and again in the middle of the first-grade year on the Scale of Children's Readiness In PrinTing (SCRIPT). Data were analyzed with a repeated measures analysis of variance (ANOVA) (3 performance groups [lower, middle, upper] x 2 grades) followed by a Tukey's HSD post hoc analysis. An analysis of the change in performance of individual children also was conducted.

RESULTS. The main effects and the interaction in the ANOVA were significant. Post hoc analysis for grade demonstrated that both the middle and the lower performing groups showed a significant improvement from kindergarten to first grade, whereas the upper performing group remained unchanged. Post hoc analysis found significant differences among all three performance groups in kindergarten. In first grade, the lower performing group continued to score significantly lower than the middle and upper performing groups, whereas the middle and upper performing groups were not significantly different. Analysis of individual data revealed that 60% of the students were consistent in their performance from kindergarten to first grade.

CONCLUSION. The correlation between kindergarten SCRIPT scores and first-grade SCRIPT scores provides initial evidence that a moderately consistent pattern of handwriting performance exists. Students with low- and middle-ranked handwriting skills show relatively similar improvement over time, whereas change for students ranked high is limited.


Handwriting is an important functional task used frequently in every grade beginning in kindergarten (Marr, Windsor, & Cermak, 2001). Children are expected to gain skill gradually in handwriting legibility as formal instruction is introduced in the kindergarten and first-grade curriculum (Vreeland, 1999). When improvements in legibility are not seen or when an individual student’s progress does not match others in the class, educators develop strategies to assist the student. One strategy frequently considered is therapeutic intervention.

Kindergarten children often are referred to occupational therapy for poor fine motor performance, including poor letter formation and handwriting. Therapy referrals made early in a student’s academic career are seen as beneficial to the student so that a deficit can be addressed, and hopefully corrected, before the student experiences some type of academic failure. Therapeutic intervention also can be seen as negative: a process of labeling and stigmatizing a student when no significant deficit exists. Because kindergarten often is a child’s first exposure to formal handwriting instruction, a deficiency in handwriting performance may be temporary. Developmental maturation, academic instruction, and practice may be sufficient to counteract any initial lack of skill. Educators and occupational therapists

Deborah Marr, MS, OTR/L, is Doctoral Student, Boston University, and Associate Professor, Utica College of Syracuse University, 1630 Burrstone Road, Utica, New York 13502; dmarr@utica.ucsu.edu.

Sharon Cermak, EdD, OTR/L, FAOTA, is Professor, Department of Occupational Therapy, Boston University, Sargent College of Health and Rehabilitation Sciences, Boston, Massachusetts.
need to know whether poor handwriting ability is a consistent trait in early elementary students so that appropriate classroom strategies and therapeutic interventions can be developed.

Consistency of Handwriting

We, the authors, define consistency of handwriting as a level of qualitative handwriting performance that persists over time. This definition of handwriting consistency does not mean that the skill does not improve but that the quality in relation to a peer group does not vary. A child with poor skill in kindergarten would continue to have poor skill over time relative to his or her peers despite improvements due to maturational change, experience, and academic instruction. If early elementary students demonstrate consistency in handwriting skill (i.e., a performance level relative to their peers), then therapists and educators can be confident that children identified early as poor performers will continue to be poor performers in future grades.

A variety of measures have been used to assess the consistency of handwriting performance over time. These measures were reviewed by Marr and Cermak (2001) and include teacher ratings, computerized assessment of specific handwriting characteristics (i.e., speed, trajectories), and assessment of overall legibility. Standardized tests, such as the Concise Evaluation Scale for Children's Handwriting (abbreviated BHK [Netherlands]; Hamstra-Bletz, De Bie, & Den Brinker, 1987), also were used in cross-sectional and longitudinal research examining consistency in handwriting performance.

Longitudinal studies of children in later elementary grades who are typically developing demonstrate inconsistency in handwriting performance when specific handwriting characteristics are evaluated. Blöte and Hamstra-Bletz (1991) followed 63 students with typical development from second grade to sixth grade. The students' handwriting characteristics (i.e., spacing, letter alignment, use of joins) were measured once yearly with the BHK. The authors found that from the second to fourth grade, some characteristics improved while others were constant. Following fourth grade, deterioration in some handwriting characteristics was noted. Sovik and Amtnsen (1991) examined a group of 12 children with typical development, using a kinematic evaluation of handwriting characteristics in third grade and again in sixth grade. They found significant improvements in some characteristics (i.e., line accuracy, movement rhythm) but not in others (i.e., speed). These studies suggest that performance of specific handwriting characteristics in groups of children who are typically developing may not be consistent across time in the later elementary years.

Several studies examined the change of handwriting skill over time with both a group of children with typical development and a group of children with dysgraphia and identified different patterns of skill acquisition between the two groups. Mæland and Karlsdottir (1991) evaluated 12 children with typical development and 12 children with dysgraphia in third grade and again in sixth grade. Two independent judges assessed handwriting with an overall measure of legibility using a 7-point scale. The authors did not state the inclusion criterion for the dysgraphic group. Handwriting was significantly better for the typical group than the dysgraphic group in both the third grade and the sixth grade, indicating consistent relative performance between the two groups. Change in handwriting for the typical group from third grade to sixth grade was not significant. Improvement in handwriting of children in the dysgraphic group was significant despite the fact that the dysgraphic group continued to perform more poorly than the typical group. Analysis of the ability to predict sixth-grade writing scores from scores in third grade indicated that for the dysgraphic group, 59% of the variance was explained by the third-grade scores compared with 36% of the variance in the typical group. The authors characterized the children in the dysgraphic group as more stable in their writing because their deficits tended to persist with age, making their performance more consistent relative to the typical group.

In another longitudinal study, Hamstra-Bletz and Blöte (1993) followed 121 children from second through sixth grade. Handwriting characteristics were measured once a year in the spring semester with the BHK. Dysgraphic writers were defined as 10% of the second graders scoring most poorly on the BHK, a criterion chosen because 10% of the students in the Netherlands are estimated to have writing problems (Hamstra-Bletz et al., 1987) and included 12 children from the sample of 121. The children with dysgraphia scored significantly lower on 7 of the 11 specific handwriting characteristics than the children without dysgraphia. For each subsequent grade, the number of characteristics found to be significantly different between the two groups was progressively fewer. No significant between-group differences were noted by the sixth grade. This decrease in the difference between groups was due in part to a decline in some skills of the nondysgraphic writers and improvements in some skills in the dysgraphic group.

Smits-Engelsman and van Galen (1997) examined 16 students from 7.6 years to 11 years of age. Eight had been identified by teachers as being poor handwriters, whereas 8 matched students were chosen from a pool of proficient
handwriting. Each matched pair was evaluated twice kinematically, with a 1-year interval, on the specific handwriting characteristics of overshoots, undershoots, movement time, writing dysfluencies, stroke curvature, and “neuromotor noise” (lack of movement precision and consistency). Results at the second test session indicated significant improvement for the sample as a whole on four of the six measures. However, the degree of improvement for each group in overshoots, undershoots, movement time, and neuromotor noise were significantly different. The proficient handwriters produced significantly fewer overshoots and undershoots and decreased movement time than the poor handwriters. These changes signify improved handwriting. The poor handwriters demonstrated significantly more neuromotor noise than the proficient handwriters, contributing to ongoing poor performance. The authors concluded that although both poor and proficient handwriters improved over time, poor handwriting persists at least for the 1-year period studied.

In summary, research on children with typical development suggests inconsistency in specific handwriting characteristics from early elementary years to later elementary years but suggests that children with typical handwriting perform consistently better over time than a lower performing group. Research on children with poor handwriting performance has provided some evidence that although handwriting improves with age, poor handwriting relative to peers with typical development is a consistent, stable characteristic (Mæland & Karlsdottir, 1991; Smits-Engelsman & van Galen, 1997). These studies include handwriting assessed by overall legibility or by specific handwriting characteristics. Studies to date have only included children who have had several years of formal handwriting instruction; no studies have examined the consistency of poor handwriters in the early elementary grades. Children in the early elementary grades are more likely to be referred for therapeutic evaluation because of their poor performance. Educators and therapists need to know whether poor performance in the early grades will improve on its own with continued instruction and practice or whether this poor performance is a consistent trait. This information is needed so that appropriate strategies for improvement can be explored.

The present study is designed to examine the consistency of handwriting in children from kindergarten to the first-grade year. The study addresses the following questions:

- What is the consistency of handwriting performance in children from kindergarten to the first-grade year?
- Will children with lower handwriting performance in kindergarten perform more consistently over time than children with moderate or high skill?

Method

Design

A longitudinal, repeated measures, descriptive design was used. Early elementary students were assessed with a handwriting measure twice over a 15-month period.

Participants

The participants comprised a convenience sample of 93 children with typical development from 10 kindergarten classes who remained in the same school district from kindergarten to the first-grade year and were in class for both test sessions. They included 42 boys and 51 girls with a mean age at the kindergarten test session of 5.6 years ($SD = 0.3$, range $= 4.9–6.2$ years). At the first-grade test session, the mean age was 6.8 years ($SD = 0.3$, range $= 6.2–7.4$ years). Eighty children were right-hand dominant, and 13 were left-hand dominant. This sample is representative of the majority of kindergarten students in a middle-income, suburban community in Upstate New York. None of the participants had an individualized education program, and all spoke English as their primary language. Socioeconomic and ethnic information was not collected; however, 6% of this district’s students are in families below the poverty level, with 10% using the free or reduced lunch program. White students comprise 94.8% of the district; Black students comprise 0.9%; Hispanics comprise 0.6%; and 3.7% are listed as “other” (State Education Department, 1998). The parents of all participants signed consent forms, and assent of the children was obtained at the time of testing.

Measure

The Scale of Children’s Readiness In PrinTing (SCRIPT) is a letter form copying test developed for kindergarten children (Weil & Amundson, 1994). The test booklet consists of five pages with a maximum of 8 letters per page using the Zaner-Bloser manuscript alphabet. All 26 lowercase letters are included, as are the following 8 uppercase letters: A, K, M, N, V, W, Y, and Z. The student sees the stimulus letter printed in the center of a square and is asked to copy the letter in the blank square located directly below the stimulus letter. The test developer provided scoring criteria. Point-by-point reliability was reported at 90% to 100% (Weil & Amundson, 1994). This level of reliability was not achieved during initial scoring in an earlier study conducted by Marr et al. (2001). Reliability of scoring for this test was investigated at that time, and more refined parameters, which included criterion from the Test of Copied and Dictated Writing, were developed (Windsor, 1995) (see Appendix). Using these refined parameters, a letter was scored 1 if all the criteria were met and 0 if any part of the criteria was not
met. A maximum total score of 34 was possible. An inter-rater reliability of .95 was established between the first and second author. Test–retest reliability was examined on two randomly chosen classrooms of first-grade participants (n = 37) from the present study tested 1 week apart and found to be .77. Validity of the SCRIPT has not been established.

Procedure

Children were tested twice with the SCRIPT: in the beginning of the kindergarten year and in the middle of the first-grade year (approximately 15 months later). Test sessions were scheduled in collaboration with the teachers. The first author administered all tests. The SCRIPT was administered to the participants grouped by classroom with the following directions:

Please copy the letters you see in the box below each letter. When you are done with one page, go on to the next page until you have done all the pages. When you are done, please put your pencil down in front of you on the desk.

Data Analysis

The participants were divided into three groups on the basis of z scores calculated from the kindergarten SCRIPT scores: the lower group scoring less than 1 standard deviation (n = 19), the middle group scoring between −1 and +1 standard deviations (n = 61), and the upper group scoring greater than 1 standard deviation (n = 13). A repeated measures analysis of variance (ANOVA) with one between-subjects factor and one within-subjects factor [Group (3) x Grade (2)] was calculated, using raw scores to evaluate main effects and interactions (Portney & Watkins, 1993). Visual inspection of descriptive data was used to describe change in individual performance over the period studied.

Results

To initially examine the consistency of handwriting performance in children from kindergarten to the first-grade year, Pearson product-moment correlations were run between the kindergarten and first-grade scores for the SCRIPT (r = .42, p < .01) with the entire sample. This correlation value represents a moderate relationship between the SCRIPT scores in kindergarten and the SCRIPT scores in first grade.

As described previously, the participants were divided into three performance-based groups. Table 1 displays the SCRIPT mean scores, standard deviations, and effect sizes of the total sample as well as for each group at the kindergarten and first-grade test sessions. Because the groups were formed on the basis of the kindergarten z scores, participants were kept in their designated group for the first-grade analysis. Effect sizes, which indicate the degree to which change occurred, demonstrated that children in the lower and middle groups made very large changes in handwriting ability from kindergarten to first grade, whereas the upper group made no change (Cohen, 1988).

Results of the repeated measures ANOVA showed that Group was significant, $F(2, 90) = 51.4, p < .001$; Grade was significant, $F(1, 90) = 60.3, p < .001$; and the Group-by-Grade interaction was significant, $F(2, 90) = 11.6, p < .001$ (see Figure 1). Tukey’s HSD post hoc analysis demonstrated significant differences on all pairwise comparisons for group in kindergarten, indicating that the lower group scored significantly lower than the middle and upper groups ($p < .001$). The middle group scored in the mid range and significantly lower than the upper group, which scored highest ($p < .001$). In first grade, the lower group scored lowest and was significantly lower than the middle and upper groups ($p < .005$), whereas the middle and upper groups were not significantly different ($p = .33$). Post hoc analysis for grade demonstrated that both the middle and lower groups showed a significant improvement from kindergarten to first grade ($p < .001$), whereas the upper group remained unchanged ($p = .97$).

The preceding analysis addressed change in group performance over time. Occupational therapists and educators need to consider whether individual children identified as having lower skill in the kindergarten year are the same chil-

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>SCRIPT Score (SD)</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower</td>
<td>19</td>
<td>8.1 (2.0)</td>
<td>1.1</td>
</tr>
<tr>
<td>Middle</td>
<td>61</td>
<td>16.4 (3.0)</td>
<td>1.9</td>
</tr>
<tr>
<td>Upper</td>
<td>13</td>
<td>25.2 (2.6)</td>
<td>0.0</td>
</tr>
<tr>
<td>Total</td>
<td>93</td>
<td>15.9 (5.7)</td>
<td>2.4</td>
</tr>
</tbody>
</table>

Note. SCRIPT = Scale of Children’s Readiness In PrinTing.
children with lower performance in subsequent years. Therefore, data on individual participants were analyzed. With the original kindergarten groupings serving as a baseline and using the same definitions to describe lower, middle, and upper performing groups, an analysis of change in an individual child's SCRIPT $z$ score from kindergarten to first grade was performed. First-grade SCRIPT $z$ scores were compared with kindergarten SCRIPT $z$ scores to determine whether a change in performance group had occurred. Sixty percent of all participants had first-grade SCRIPT $z$ scores that placed them in the same performance group as their kindergarten SCRIPT $z$ scores. An examination of change in each of the three performance levels separately revealed that 42% of the kindergarten low group remained in the low group in first grade; 70% of the kindergarten middle group remained in the middle group; and 38% of the kindergarten upper group remained in the upper group. Figure 2 displays the full range of change patterns demonstrated by the participants.

**Discussion**

The correlation between kindergarten SCRIPT scores and first-grade SCRIPT scores provides initial evidence that a moderately consistent pattern of handwriting performance exists in young children. Examination of the descriptive data in Table 1 and the results of the repeated measures ANOVA provide further evidence of the consistency of handwriting performance. The mean for each group remains consistent in relation to the other groups over time, with the lowest group remaining the lowest; however, the change from kindergarten to first grade is not similar across groups (see Figure 1). The lower and middle groups show relatively similar improvement, whereas change in the upper group is limited. This interaction is, to some extent, consistent with previous research that found different patterns of change for children with and children without dysgraphia. Mæland and Karlsdottir (1991) found that students with typical development did not show significant improvement in handwriting legibility from third to sixth grade, whereas improvements in children with dysgraphia were significant. Hamstra-Bletz and Blöte (1993) also found that some handwriting characteristics of children with typical development from second to sixth grade deteriorated, whereas children with dysgraphia improved across time. For both studies, the typical group remained significantly better in most aspects of handwriting performance. Although the present study did not include children labeled dysgraphic, similar patterns between lower and upper performers were identified. An alternative explanation is regression to the mean for the upper and lower groups. Examination of individual, rather than group, performance on the SCRIPT also substantiates consistency, as 60% of the participants remained in their same performance group from kindergarten to first grade. When performance groups are examined individually, less consistency is found in lower and upper performers than in middle performers. The middle performers had the highest level of consistency over time, with 70% in the middle group at kindergarten remaining in the same performance group in the first grade. Lower and upper performers had moderate levels of performance consistency: 42% of the lower group and 38% of the upper group were performing at the same levels in first grade. These results differ from Mæland and Karlsdottir (1991) and Smits-Engelsman and van Galen (1997), which suggest more consistency in dysgraphic performance.

**Table 2. Characteristics of Children in the Lower Group**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>$n$</th>
<th>Male</th>
<th>Female</th>
<th>$M$</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kindergarten lower group</td>
<td>19</td>
<td>11</td>
<td>8</td>
<td>5.5</td>
<td>5.1–5.9</td>
</tr>
<tr>
<td>Children who started in lower group and remained in lower group</td>
<td>8</td>
<td>5</td>
<td>3</td>
<td>5.6</td>
<td>5.3–5.8</td>
</tr>
<tr>
<td>Children who started in lower group and moved to middle and upper groups</td>
<td>11</td>
<td>6</td>
<td>5</td>
<td>5.4</td>
<td>5.0–5.8</td>
</tr>
</tbody>
</table>
groups than in typical groups. However, none of the previous studies looked at individual changes. Additionally, the present study has examined only 15 months of change, whereas earlier studies have examined longer periods. It is possible that following young children over the entire course of their elementary career would reveal further variations in the course of their handwriting development.

Age and gender are thought to influence handwriting performance, with either younger children of either gender or male children having lower performance (Blöte & Hamstra-Bletz, 1991; Graham, Berninger, Weintraub, & Schafer, 1998; Sovik & Arntzen, 1991). Table 2 provides a closer examination of the age and gender characteristics of the children in the lower group beginning in kindergarten and their characteristics based on first-grade group designation. The children that began in the lower group in kindergarten and stayed in the lower group in first grade did not appear to differ in age or gender from those who started in the lower group but moved to a higher group in first grade. This finding runs counter to the widely held view that younger children that began in the lower group in kindergarten and their male children having lower performance (Blöte & Hamstra-Bletz, 1991; Graham, Berninger, Weintraub, & Schafer, 1998; Sovik & Arntzen, 1991).

The results of this study suggest implications for occupational therapy consideration. When kindergarten children are identified as having lower performance in handwriting, should their needs be addressed? Although slightly less than half of the lower performing children remained low in first grade, could some type of educational or therapeutic strategies be targeted to the whole low performance group to alter the consistent low performance of a portion of that group? In this participating school district during the kindergarten year, all students received weekly group activities provided in the classroom by an occupational therapist. The activities were designed to improve fine motor, gross motor, and postural stability skills. Kindergarten teachers also received consultation in these areas. The school board initiated these prereferral occupational therapy activities outside the scope of the present study in an attempt to improve teachers’ understanding of fine motor development. The participating school district also believed that this program could possibly improve the skills of all children. A model of prereferral services may be effective and offers an arena for future research.

Acknowledgments

We thank Erin Cain, occupational therapy student, who assisted with the data and contributed to the development of ideas found in this article.

This study was supported in part by a leadership training grant to Boston University from the U.S. Department of Health and Human Services, Health Resources and Services Administration, Maternal and Child Health Bureau (MCJ-000-901), and by a grant from the American Occupational Therapy Foundation.

This study was completed in partial fulfillment of the first author’s requirements for the degree of Doctor of Science in Therapeutic Studies at Boston University.

Appendix

Scoring the Scale of Children’s Readiness In PrinTing

1. The letter is quickly and easily recognized as itself and no other symbol using the “peek hole” method; no gross errors in proportion are present. Case (upper or lower) is correct.
2. The letter has no missing parts and no extra parts. This includes the need to have the “stick” on a lowercase n.
3. No lines extend beyond the intersection by more than 2 mm.
4. Baselines and toplines must be parallel to the horizontal boundary lines of the blank stimulus box within 3 mm. Toplines and bottom boundary lines are not used for the letters a, b, d, q, g, r, and p, and the bottom of u.
5. Upstrokes and downstrokes must be parallel to the vertical boundaries within 3 mm. The capital letter M and the dots on i and j are not included in this criterion. The side points of z, s, x, k, e, and c must fall within a 3-mm space of each other, which is perpendicular to the horizontal boundaries.
6. Letter forms must be closed correctly, with no more than a 2-mm gap. For k, this means that the intersection of the two angled lines can be no more than 2 mm apart.
7. Curved lines must be curved, and straight lines must be able to fit within a 2-mm space. These criteria include any extension lines that may be present.
8. Angles must be present.
9. There is no rotation of more than 45° in any part of the letter. No reversals are present.
10. Each side of the horizontal line in t and f must be within 2-mm length of the other. The bottom portion of the vertical line of the t must be at least 2 mm longer than the top side.
11. Oblique lines cannot be perpendicular to the outer boundary lines (e.g., v, w, y).

Note. Each letter must pass each criterion to be awarded 1 point. Failure on any one criterion results in a score of 0 for that letter (Marr et al., 2001).

References


