Handwriting evaluation for developmental dysgraphia:
Process versus product

Sara Rosenblum, Ph.D¹, Patrice L. Weiss, Ph.D¹ and Shula Parush, Ph.D²

1. Department of Occupational Therapy, Faculty of Social Welfare & Health Studies,
   University of Haifa, Haifa, Israel

2. School of Occupational Therapy, The Hebrew University Faculty of Medicine, Jerusalem,
   Israel

Abstract

The act of writing presents difficulties for 10-30% of elementary school children. This study’s objectives were to compare the abilities of digitizer-based evaluation of the handwriting process and conventional evaluation of the handwriting product to discriminate between children with proficient and dysgraphic handwriting. Copied and dictated writing samples were collected from 3rd grade students, 50 with proficient and 50 with dysgraphic handwriting. Results indicated that both digitizer-based and conventional evaluations differentiated between children with proficient and dysgraphic handwriting, and that together they provided an improved understanding of writing difficulties. Moreover, copying and dictated writing task results significantly differed. The results demonstrate the advantages of combining both handwriting process and product testing, and utilizing both copying and dictation tasks, in order to achieve a more comprehensive understanding and superior evaluation of developmental dysgraphia.

Key words: digitizer, dysgraphia, evaluation, handwriting, school-aged children
Introduction

Handwriting is an important means of communication that enables the expression, recording, and transmission of ideas of students throughout their educational careers (Dennis & Swinth, 2001; Hamstra-Bletz & Blote, 1993; Tseng & Cermak, 1993). Elementary school children typically spend up to 50% of the school day engaged in writing tasks, some of which are performed under time constraints (Amundson & Weil, 1996; McHale & Cermak, 1992; Tseng & Chow, 2000). A child’s ability to write legibly, as well as quickly and efficiently, enables him or her to achieve both functional written communication and academic advancement (Amundson & Weil, 1996; Phelps, Stempel & Speck, 1985; Tseng & Cermak, 1993; Tseng & Hsueh, 1997).

Handwriting difficulty or dysgraphia was defined by Hamstra-Bletz & Blote (1993) as a disturbance or difficulty in the production of written language that is related to the mechanics of writing. It has also been referred to as a specific learning disability (Brown, 1981). The problem is manifested in the inadequate performance of handwriting among children who are of at least average intelligence and who have not been identified as having any obvious neurological problems. Teachers have estimated that 11-12% of female students and 21-32% of male students have handwriting difficulties (Rubin & Henderson, 1982; Smits-Engelsman, Van Galen & Michels, 1995). Dysgraphia or poor handwriting is a common complaint among children and adults with learning disabilities, appearing with or without other academic difficulties (Cratty, 1994; Johnson, 1995; Waber & Bernstein, 1994), as well as in children diagnosed with Developmental Coordination Disorder (DSM-4, American Psychiatric Association (APA), 1994), children defined by their teachers as clumsy (Laszlo, 1990; Laszlo, Bairstow, & Bartip, 1988) and children diagnosed with developmental right-hemisphere syndrome (Gross-Tsur, Shalev, Manor & Amir, 1995). In fact, Cratty (1994) found that 30-40% of the surveyed children with learning disabilities had handwriting
Several authors have suggested that difficulty in the mastery of the mechanical aspects of handwriting may interfere with the higher order processes required for the composition of text (Berninger & Graham, 1998). Graham (1990) found that handwriting mechanics influence the quality and quantity of the written product. This finding is supported by Berninger et al. (1997), who reported that handwriting performance was significantly related to fluency and quality of composition in elementary school students. Recently, Graham, Harris, & Fink (2000) summarized existing views on the negative consequences of handwriting difficulties. They, along with others (Briggs, 1980; Chase, 1986; Hughes, Keeling & Tuck, 1983), suggested that teachers tend to give higher marks for neatly written papers than for those in which legibility is poorer. Thus, it appears that poor penmanship may influence perceptions about children’s competence as writers. Other authors have proposed that the act of handwriting among children with difficulties can interfere with the simultaneous execution of composition (Graham, 1990; Scardamalia, Bereiter, & Goleman 1982). It may be that when letter production is not fully automatic, the act of handwriting makes increased demands on memory and attentional resources, which, in turn, constrain the higher level cognitive processes required for composition (Jones & Christensen, 1999; Berninger & Graham, 1998). Additionally, some suggest that if handwriting is very slow, children may forget the ideas and plans held in memory before they succeed in transferring them to paper (Graham & Weintraub, 1996).

The most frequent complaint made regarding children with poor/dysgraphic handwriting is in the quality of their script (Smits-Engelsman & Van Galen, 1997). Compositions and essay exams are a frequent classroom requirement and are a great source of frustration for those struggling with the skill of handwriting. It is no wonder, therefore, that researchers have suggested that handwriting difficulties may have serious
consequences for the student’s academic progress, emotional well-being and social functioning (Cornhill & Case-Smith, 1996; Kaminsky & Powers, 1981; Martlew, 1992; Modlinger, 1983). Hence, it reinforces the importance of identifying handwriting difficulties as early as possible both as a preventive and as a corrective aid (Phelps & Stempel, 1988). The process of describing the characteristics of the written output of children with handwriting difficulties began early in the twentieth century and formed the basis for the development of handwriting evaluation scales (Ayres, 1912; Bezzi, 1962; Feldt, 1962; Freeman, 1959; Thorndike, 1910). The objective of these scales was to determine how best to define the “quality”, “readability” (Ayres, 1912) or "legibility" of handwriting in specific, measurable terms. Over the years, two main approaches were utilized for evaluating the “readability” or "legibility" of handwriting: a global-holistic evaluation of readability (Ayres, 1912; Freeman, 1959), and an analytic evaluation of readability utilizing predetermined criteria (Alston, 1983; Amundson, 1995; Hamstra-Bletz, De Bie & De Binker, 1987; Larsen & Hammil, 1989; Reisman, 1993).

The global evaluation scales are based on techniques used to form an overall judgment of a written passage, that is, how readable it is as compared to a group of standard writing samples that have been previously graded from readable to unreadable. Dissatisfaction with these types of scales stems from their inherent subjectivity, which places their overall reliability into question, and from the fact that scoring the readability of a writing sample requires an exhaustive and time consuming comparison to numerous pre-graded samples. This approach has only limited practicality (Feldt, 1962; Freeman, 1959; Graham, 1982).

In contrast, the assumption underlying the analytic method is that a relationship exists between the general look, the “readability” or legibility of the writing, and the quality of specific features that characterize it. Many of these specific features are related to spatial characteristics (e.g., the spacing of the letters and the words), while others are related to letter
formation (e.g., letter shape and size). The writing sample is judged by first grading each
feature individually for the entire passage and then calculating an overall score (Alston, 1983;
Amundson, 1995; Erez & Parush, 1999; Hamstra-Bletz, De Bie and De Binker, 1987; Larsen
& Hammil, 1989; Phelps et al., 1985; Reisman, 1991; Stott, Moyes & Henderson, 1984;

It had been hoped that the analytically based scales would provide a satisfactory solution
to handwriting evaluation in that they are less subjective than the global-holistic scales.
However, the analytic scales are also limited in ways that hinder the development of
optimally effective handwriting assessment. Analytic scales require prolonged processing
time, lack the overall holistic perspective found in global scales, and are plagued by
methodological variations that prevent direct comparisons between them (Bonny, 1992;
Graham, 1986a, 1986b; Graham & Weintraub, 1996; Rosenblum, Weiss, & Parush, 2003;
Rubin & Henderson, 1982).

Hence, clinicians, educators, and scale development researchers continue to have
reservations regarding each of these approaches to handwriting assessment. Moreover,
neither the global nor the analytic evaluation scales are able to go beyond an analysis of the
written product in order to provide substantive information about the writing process. This
poses a significant limitation since it is believed that only a comprehensive description of the
real-time, dynamic characteristics of a child’s handwriting can provide insight into the motor
control mechanisms of normal handwriting and an understanding of the underlying pathology
of handwriting difficulties (Dobbie & Askov, 1995; Graham & Weintraub, 1996; Longstaff &

In recent years, computer, digitizer-based technology has led to the development of an
innovative approach to handwriting evaluation that is capable of analyzing the writing
process during the actual performance of a written task. Digitizer-based technology enables
not only the attainment of measures of the handwriting process, but also provides access to
information that the human eye is incapable of discerning. Initial results indicate that the use
of this technology reveals the presence of objective, measurable features that distinguish
between the writing of children who have and those who do not have handwriting difficulties
(Rosenblum, Parush, & Weiss, 2001; Smits-Engelsman, 1995; Smits-Engelsman, Van Galen
& Portier, 1994a, 1994b; Smits-Engelsman, Van Galen & Shoemaker, 1998; Sovik et al.,
1987a, 1987b). The key instrument in this technology is a digitizer, an electronic surface
capable of recording x and y coordinates when used in tandem with a wireless pen and
personal computer. The digitizer/computer system is able to record kinematic and kinetic
variables as the writer engages in various writing tasks. The result is a comprehensive
collection of dynamic features that generates a more objective and precise set of handwriting
outcome measures than had been previously attainable.

Recently, several research groups have examined the writing performance of children
with handwriting difficulties via computerized systems (e.g., Rosenblum et al., 2001; Smits-
Engelsman & Van Galen, 1997; Sovik et al., 1987b). Their results have shown that poor
writers are characterized by having greater variability in temporal and spatial variables, lack
writing continuity and fluency, and pause for longer periods between strokes, letters, and
words than their peers (Sovik et al., 1987b; Wann & Jones, 1986; Wann & Kardirkamanathn,
1991; Shoemaker, Shellekens, Kalverboer, & Kooistra, 1994; Shoemaker & Smits-
Engelsman, 1997; Wann, 1987; Smits-Engelsman & Van Galen, 1997). It was also found
that children with poor/dysgraphic handwriting achieved higher absolute scores of “neuro-
motor noise” (variability of the motor output system) in comparison with more proficient
writers, and demonstrated less ability to regulate muscle force (Van Galen, Portier, Smits-
Engelsman, & Shoemaker, 1993).

Smits-Engelsman (1995) ascribed their deficits to a poor muscular initiation process
Mojet (1991) and Wann & Kardirkamanathan (1991) found that the writing of poor handwriters tends to lack continuity within letter sequences and shows a wide variability in the orientation of individually written letter segments. Further, the results of a study carried out by Smits-Engelsman et al. (1994a) indicated that children with poor handwriting made more spatial errors than proficient writers, leading them to conclude that poor handwriting may stem from a problem in spatial control. Longitudinal research performed on children with handwriting difficulties supports the view that children with dysgraphic handwriting fail to obey spatial constraints, and that their handwriting lacks consistency. For example, the size of their letters was inconsistent and they lacked adequate curvature in their strokes (Smits-Engelsman, & Van Galen, 1997).

In previous research (Rosenblum, Parush & Weiss, 2003b), we used a digitizer to compare the writing process of children with poor and proficient handwriting, through the analysis of writing tasks that replicated the types of assignments commonly required in the classroom (i.e., letters, words, sentences and paragraphs in both copying and dictated writing modes). These studies introduced variables designed to illuminate both the temporal and spatial characteristics of the handwriting process. Both the time and path lengths during which the subjects’ pens were in contact with the writing surface (On Paper variables), as well as the time and path lengths in which their pens were not in contact with the writing surface (In Air variables) were recorded. Results of these studies indicated that when writing sentences and paragraphs, the In Air time of children with poor handwriting was significantly longer than that of proficient handwriters, demonstrating that they maintain contact with the writing surface for less time than do their peers. Further analysis revealed that all the subjects tended to keep their pens in motion during the In Air time, performing what was described as a “motion tour”. Moreover, the motion tour produced by the children with poor handwriting
was longer and more intricate than that of those with proficient handwriting, reflected by the significantly longer In Air time and path lengths that were measured for this group. Thus, it can be seen that research in the 20th century has led to the advancement of knowledge about the nature of deficient handwriting in several important areas. There now exist a number of evaluation scales that can be used to manually assess the legibility of various handwriting products through global-holistic and analytic means. Moreover, technology has given us the means to perform computerized online monitoring of the handwriting process. However, when each of these methodologies is considered by itself, specific limitations become apparent. Although the use of conventional writing assessments by educators and clinicians have enabled the evaluation of handwriting legibility, the reliability and validity of such tools are, to some extent, limited due to their dependence on subjective judgment for scoring and interpretation. On the other hand, although the development of computerized analysis has advanced the understanding of the spatial and temporal characteristics of children with poor and proficient handwriting, further study is still needed to determine how these data are related to the actual functional deficits in handwriting.

The purpose of this study was to compare the relative abilities of digitizer-based evaluation and conventional evaluation to discriminate between the writing of children with proficient and dysgraphic handwriting. A further purpose was to determine whether the digitizer, through its unique ability to measure spatio-temporal handwriting variables that cannot be discerned with the human eye, provides additional information important to the characterization of handwriting difficulties.

Finally, a third objective was to address an issue that has, to date, received insufficient attention, namely the effect of different writing modes (copying versus dictation). In most studies, students’ handwriting has been evaluated via copying assignments (Ziviani & Elkins,
1984; Cornhill & Case-Smith, 1996), although in some cases students were required to perform free style writing (Alston, 1983) or to write a sentence from memory (Sovik, Arntzen, Samuelstuen, & Heggberget, 1994). This poses a problem, since children in their everyday life are confronted with the need to write under a variety of modes, including copying, writing in response to dictation, free style writing, or writing from memory (such as while taking notes). This becomes more significant if one considers the fact that each of these modes relies on different underlying sensory modalities. For example, copying requires a student to rely on visual stimuli, while dictation requires a reliance on auditory stimuli (Fryburg, 1997). Moreover, it was found that students in grades 1 through 6 wrote more legibly while copying than while writing free style (Graham, Berninger, & Weintraub, 1998). Further, studies concerned with agraphia among patients with brain damage emphasized the importance of utilizing different modes of functional writing (i.e., copying and dictation) for a better understanding of the phenomenon (Hua, Chen, & Chu, 2001).

Method

Participants

Two groups of handwriters (proficient and dysgraphic), each consisting of 50 third grade pupils, aged 8 and 9 years old, were recruited from eight regular public schools located in four different types of municipalities in northern Israel (large town, small town, kibbutz and community settlement). All participants were Caucasian Jews from either Ashkenazi or Sephardic origin, who were born in Israel, used the Hebrew language as their primary means of verbal and written communication, and were right hand dominant. The parents of all the children who participated in the study gave their informed consent.

Third grade pupils were selected as the target population since the handwriting development literature has indicated that by the time a child reaches third grade, his or her
handwriting has become more automatic, organized, and readily available as a tool to facilitate the development of ideas (Berninger, Mizokawa, & Bragg, 1991; Levine, 1987; O’Hare, 1999; Ziviani & Elkins, 1984). Thus, a deficiency in these qualities at this age may be a sign of a problem.

The 100 participants were identified as having proficient or dysgraphic handwriting with the aid of the standardized and validated Teachers’ Questionnaire for Handwriting Proficiency (Rosenblum, Jessel, Adi-Japha, Parush & Weiss, 1997), completed by their classroom teachers. The questionnaire was constructed from criteria selected from the literature and handwriting assessments, including handwriting legibility, speed, fatigue, and complaints of pain or discomfort while writing (Alston, 1983; Cornhill & Case-Smith, 1996; Rubin & Henderson, 1982). Children with documented developmental delay, neurological deficits, or physical impairment were excluded from the study. After the children were classified into groups of proficient versus poor handwriters according to the Teachers’ Questionnaire for Handwriting Proficiency, one of the authors (SR) administered the Hebrew Handwriting Evaluation (HHE) (Erez & Parush, 1999), a standardized, reliable and valid handwriting assessment for the Hebrew language, to all of the children. One hundred percent agreement between these two measures was found (i.e., the same 50 children who were categorized as poor, and the fifty categorized as proficient writers according to the Teachers’ Questionnaire for Handwriting Proficiency were also categorized as poor - or proficient with the HHE). Seventy percent of these children were boys, a gender bias that concurs with clinical reports and with the literature (Rubin and Henderson, 1982; Smits -Engelsman, Schomaker, Van Galen, & Michels, 1996).

The children with proficient handwriting were matched to the participants in the dysgraphic handwriting group, on the basis of gender, age, school, and grade. For each child in the dysgraphic handwriting group, a matched control participant was chosen from his or
her classroom peers, and was taught by the same classroom teacher. Thus, there were no differences between the two groups with respect to their age (8.68 ± 0.27 years for the children with proficient handwriting and 8.61 ± 0.35 years for the children with dysgraphic handwriting) and gender ratio (30% girls versus 70% boys in both groups).

Instruments

Digitizing Tablet and On-line Data Collection and Analysis Software

A unique data collection and analysis software program, developed specifically for the current study, including a suite of on-line, computerized tasks programmed in C++ and implemented via Matlab software toolkits, was used to administer the stimuli and to collect and analyze the data. The writing tasks were performed on A4 size lined paper affixed to the surface of a WACOM (407 X 417 X 36.3 mm) x-y digitizing tablet using a wireless electronic pen with a pressure sensitive tip (model Up 401). Pen size and weight were similar to those of normal pens (length=150mm, circumference=35mm, weight=11gm). The paper was lined (spacing=0.5 cm). Displacement, pressure, and pen tip angle were sampled at 150 Hz via a 90 MHz Pentium laptop computer.

The Hebrew Handwriting Evaluation (HHE) (Erez & Parush, 1999)

The HHE was used to examine the handwriting product, assessing legibility through both global and analytic measures. It contains one text for assessing writing performed through a copying mode and a separate text for assessing writing performed through dictation. Both texts are of the same level of difficulty and appropriate for third graders. In addition, each text contains all the letters in the Hebrew alphabet, and included the same number of words (30) and letters (107) (Erez & Parush, 1999). The inter-rater reliability of the HHE is \( r = .75-.79; p < .001 \). Construct validity of the HHE has been established by demonstrating statistical
significant differences ($t = -2.34; p = .027$) between the performance of children with proficient and poor handwriting (Dvash, Levi, Traub, & Shapiro, 1995). The standardized texts from the Hebrew Handwriting Evaluation (Erez & Parush, 1999) were also chosen in the current study as the paragraph writing tasks to be analyzed by the digitizer, to enhance the fidelity between conventional (HHE) and the digitizer-based approaches.

All the 100 handwriting products were analyzed by the same evaluator. The evaluator was blind as to the group membership of each child. The evaluator was certified in HHE administration having completed a course for this purpose that was conducted by the tool developers. As part of the certification process, inter-rater reliability of the course participants was demonstrated ($r = .75 - .79, p < .01$).

**Outcome Measures**

The primary outcome measures included temporal and spatial measures of handwriting kinematics. The temporal measures included the total time taken to complete the writing tasks, On Paper time, and In Air time (i.e., the time during the writing of the paragraph that the pen was not in contact with the writing surface). The spatial measures included the total path length of all the characters written in the paragraph, the On Paper path length, and the In Air path length (i.e., the excursion of the pen when it was not in contact with the writing surface while writing the paragraph).

The outcome measures of the HHE assessment of the written product included global legibility (scored on a 4 point Likert scale, from the most legible [1] to the least legible [4]), which refers to the overall clarity of the handwriting. The analytic measurement of legibility used in the HHE examined the following three variables:

1. Letters erased and/or overwritten – the number of letters that were erased and/or written over.
2. Unrecognizable letters – the total number of letters that could not be recognized due to the quality of letter closure, rounding of letters, or letter reversals.

3. Spatial arrangement of the written text, as determined according to detailed and precise criteria, utilizing a caliper that is calibrated to the millimeter. Specifically, these criteria included vertical alignment of letters (including the extensions of letters above and below the lines), the spacing of words and letters (whether too wide or overlapping), and letter size. The minimum score for spatial arrangement is 9, and the maximum is 24. For all of the four outcome measures of the HHE, a low score indicates good performance and a high score indicates poor performance.

Procedure

All participants were tested individually under similar environmental conditions. Specifically, testing occurred in a quiet classroom in their school during the morning hours. In addition, in an effort to achieve writing samples that would resemble those typically produced by the subjects, all environmental factors were kept as similar as possible to writing conditions that the child would normally experience. Thus, the participant was seated on a standard school chair and in front of a school desk, appropriate to his or her height.

Paragraph copying and dictation tasks were selected to represent handwriting tasks in which a child would typically engage. The copying task was presented visually on the screen in Hebrew font type Gutman Yad-Brush size 20 point (shown in Figure 1). The content of the dictation task was read to the child by the test administrator, who paced her reading in accordance with the participant's writing speed. As shown in Table 2, the mean dictation time for the proficient writers was $124.45 \pm 29.49$ s and for poor writers was $208.95 \pm 78.35$ s. The tasks were written on normal writing paper with printed lineature, which was affixed to the digitizing tablet. Each subject was instructed in the same fashion about what he or she would be required to do. The testing for this part of the study took approximately 15
minutes; the subjects had previously engaged in other handwriting tasks for an additional 25 minutes (see Rosenblum et al. (2001) for a full description). The same tester carried out all computerized data collection sessions.

With respect to the data analysis, MANOVA procedures were used to examine group differences across the dependent conventional variables of the handwriting product (e.g., subjective legibility, unrecognizable letters), and the digitizer-based variables of the handwriting process (e.g., total time/length, In Air time/length), for each writing task (e.g., paragraph copying versus dictation). To examine the source of significance, the data from each task were subjected to univariate ANOVAs. Discriminant analysis was conducted in order to determine which of both the conventional and digitizer-based variables were the best predictors of group membership (i.e., proficient and dysgraphic) in each of the two kinds of tasks (copying or dictation). Pearson correlations were calculated in order to check for associations between the conventional and digitizer-based measures and between copying and dictation performance of all the 100 children. An effect size (Eta squared) was provided for the cases where p-values were reported.

Results

Prior to a detailed presentation of the statistical analysis, we have included some samples of paragraphs written by typical participants from each group. In Figure 1, two representative samples of the copying paragraph task are shown. Figure 1A shows how this paragraph appeared on the computer screen to participants. The lower two panels of the figure show two samples of the same paragraph as it was written by a child with proficient (1B) and a child with dysgraphic (1C) handwriting. Heavy lines represent the actual trajectory of the child’s pen when in contact with the writing surface, and thin lines show the In Air trajectory, i.e., when the pen was above the writing surface. A visual examination of the extent of these
thin line trajectories (indicating the In Air path length) reveals that the trajectory, and hence the time spent In Air by the child with dysgraphic handwriting, was considerably greater than that of the child with proficient handwriting.

[Fig. 1 here]

The first group of variables to be analyzed were the conventional measures obtained via the Hebrew Handwriting Evaluation (HHE) (see Table 1). The MANOVA applied to these four variables (global legibility, letters erased and/or overwritten, unrecognizable letters and spatial arrangement) yielded statistically significant differences between the children with proficient and the dysgraphic handwriting for both the copying task ($F(4, 95) = 13.57, p < .001$) and for the dictation task ($F(4, 95) = 57.67, p < .001$). Univariate ANOVAs, used to examine the source of the significance, showed that children with proficient handwriting received significantly lower scores (i.e. performed better) in all the four variables than did the children with dysgraphic handwriting, for both the copying and dictation tasks.

[Table 1 here]

The second group of measures included the spatial and temporal variables obtained from the computerized digitizer system (see Table 2). As shown previously, the six variables obtained via the digitizer (total length, On Paper length, In Air length, total time, On Paper time, and In Air time) yielded statistically significant differences between the children with proficient and the dysgraphic handwriting for both the copying ($F(5, 91) = 9.82, p < .0001$) (Rsoenblum, Parush, & Weiss, 2003a, 2003b). These same variables also differentiated between the two groups for the dictation tasks ($F(5, 93) = 11.96, p < .0001$). When the data from each variable were subjected to univariate ANOVAs, results showed that children with proficient handwriting required significantly less total time, On Paper time, and In Air time for both the copying and dictation tasks, than did the children with dysgraphic handwriting. Moreover, they had shorter total, On Paper and In Air path lengths during both the copying and dictation
tasks than did the children with poor handwriting.

Discriminant analysis was conducted to determine whether the eight variables (four conventional and four digitizer based measures) could predict group membership (i.e., children with proficient versus children with dysgraphic handwriting). Results indicated that the discriminant function was found to be significant for both the copying (F(8, 87) = 9.95, p < .0001) and the dictation tasks (F(8, 89) = 25.96, p = .0001) (see Table 3). Amongst the conventional variables, the highest predictor was the subjective legibility variable during the dictation task. Amongst the digitizer-based measures, the highest predictor was In Air time during the copying task. In contrast, On Paper length was the lowest predictor.

When the data from all the participants were analyzed, significant correlations, ranging from .31 to .65, were found between all four conventional measures and three of the digitizer based measures (In Air time, In Air length and On Paper time), for both copying and dictation tasks (see Table 4.1). However, when correlations between the conventional and digitizer-based measures were calculated separately for each study group (Table 4.2 versus Table 4.3), results indicated that the majority of significant correlations were found to exist amongst the dysgraphic writers.

Significant and high correlations between copying and dictation writing tasks, ranging from .54 to .86, were found for all of the digitizer-based measures (see Table 4). In general, the values of the correlations were higher for the digitizer-based measures than for the conventional measures.

The results of copying and dictation performance of the proficient and dysgraphic
handwriters were previously presented in tables 1 and 2. As a second phase, a MANOVA was used to compare the results of the conventional and digitizer-based measures between the copying and dictation for each of the study groups. Results indicated that no significance differences (F(4, 96) = 0.54, p < .71) were found in the conventional measures between copying and dictation tasks within the proficient group, but significant differences were found in the dysgraphic group (F(4, 96) = 24.61 p < .0001). To examine the source of these significant differences in the dysgraphic group, the data from each variable were subjected to univariate ANOVAs. The results indicated that the children with dysgraphic handwriting performed significantly better (i.e., achieved lower scores) for subjective legibility and spatial arrangement during the copying task (p < .01, partial eta squared is .37 for subjective legibility and .58 for spatial arrangement). In contrast, both children with proficient and dysgraphic handwriting performed significantly differently when copying and when writing from dictation, for all the temporal and spatial digitizer based measures (p < .01, Partial eta squared ranging between .29 - .63).

Discussion

The current study represents one of the first attempts to utilize digitizer-based data as an adjunct to conventional handwriting assessment in order to examine the contribution of each method to the identification and characterization of dysgraphic handwriting. This study also examines the importance of evaluating handwriting in both copying and dictation writing modes. In order to address these goals we began by examining whether both conventional and digitizer-based handwriting measures used in this study differentiated significantly between dysgraphic and proficient handwriting, and to what extent they complement each other in the evaluation process.

The fact that the conventional measures used in the current study differentiated significantly between dysgraphic and proficient handwriting is not trivial. Researchers over
the years have not been satisfied that the criteria that have been relied on thus far truly represent the critical components of handwriting legibility nor comprise the optimal way to measure them (Bonny, 1992; Daniel & Froude, 1998; Graham, 1986b; Phelps et al., 1985; Reisman, 1993; Rubin & Henderson, 1982). This sentiment is echoed by Dennis & Swinth’s (2001) recent statement that although legibility is certainly a crucial component of written communication it remains quite difficult to define. In addition, a review of handwriting evaluation scales (Rosenblum et al., 2003) reveals that although many handwriting evaluation developers have attempted to define the characteristics of handwriting that determine legibility, they have not yet been shown to discriminate sufficiently well between dysgraphic and proficient handwriting. For example, the BHK scale (Hamstra Bletz, De Bie, & De Binker, 1987) is an evaluation tool that has been found to discriminate between children with dysgraphic and non-dysgraphic handwriting. Yet, further research indicated that 25% of the items used in the original BHK to measure different aspects of handwriting legibility were eliminated in a revised edition, since they were found not to discriminate sufficiently well between dysgraphic and non-dysgraphic handwriting (Hamstra-Belt and Blote, 1993). The items that were found to discriminate well between these groups were those that examined various aspects of letter formation (poor letter alignment, acute turns in connecting joins to letters, irregularities in joins, absence of joins, collisions of letters and inconsistent letter size), word alignment and word spacing.

Our results show that the conventional criteria used in this study (i.e., global legibility, letter erased or overwritten, unrecognizable letters, and spatial arrangement) did successfully discriminate between dysgraphic and proficient handwriting. These criteria are congruent with those used in previous studies. In the Hebrew Handwriting Evaluation, Erez and Parush (1999) demonstrated significant differences between children with poor and proficient handwriting using these same criteria (Sary, Parush, & Jarus, 1996; Lifshietz &
Parush, 1993; Pindack & Parush, 1996). Graham, Boyer-Schick, & Tippets, (1989) found that almost two-thirds of the legibility scores of students with writing difficulties on the Test of Legible Handwriting criteria (TOLH; Larsen & Hammil, 1989), were due to problems in letter shaping, spacing and organization. According to the BHK researchers (Hamstra-Belt and Blote, 1993), dysgraphic writing can be characterized by a consistently lower quality of individual spatial writing features. These include inconsistent letter size, acute turns in the letters and joints, relatively uneven and unsteady writing, and sudden changes in size and direction of letter writing. Graham, Weintraub, & Berninger (2001) found that letter legibility made a significant contribution to the prediction of text legibility.

Having confirmed that all of the conventional and the digitizer based measures were significantly different for dysgraphic versus proficient handwriting, discriminant analysis was then used to determine to what extent the conventional and digitizer based variables discriminated between them. The results indicated that global legibility discriminated best. The fact that conventional measures discriminated so well between proficient and dysgraphic handwriting supports the importance of continuing to include conventional handwriting assessments in the process of evaluating children’s handwriting. Furthermore, it is interesting to observe that among the conventional measures, including analytical handwriting legibility criteria, global legibility discriminated best. The fact that some handwriting researchers over the past five years have returned to collecting data on handwriting quality through the use of global legibility ratings by classroom teachers seems to reinforce the importance of our finding that conventional measures of global legibility are highly effective. Dennis and Swinthe's (2001) suggestion that it is easier to achieve agreement between testers when using overall legibility scores than when measuring handwriting quality through analytic legibility criteria provides further support for our results.
Results of the analysis of the digitizer based data in the current study revealed that the child’s pencil ‘traveled’ above the writing surface between the writing of successive character segments, letters and words (the In Air phenomena), for a significantly longer period of time and trajectory among the non-proficient handwriters than among the proficient writers (Rosenblum et al., 2003b). These results may provide clues as to the underlying difficulties that limit the performance of children with dysgraphic handwriting during the process of writing and that are manifested in the quality of their written products. This contention is supported by the fact that all four conventional measures were significantly correlated with the digitizer based spatio-temporal (In Air length and In Air time) measures, mainly among the children with dysgraphic handwriting. That is, the poorer the score a child received for global legibility, the longer the child’s measured In Air time and length were found to be. This relationship was considerably less among the children with proficient handwriting. Therefore, one may speculate that these phenomena are an expression of the reduced continuity, consistency, and efficiency of writing movements that characterize the writing performance of children with dysgraphia, as described in earlier clinical (Ziviani & Elkins, 1984) and digitizer studies (Mojet, 1991; Smits-Engelsman & Van Galen, 1997; Van Galen et al., 1993; Wann, 1987; Wann & Kardirkamanathan, 1991). These results are quite intriguing since they demonstrate that digitizer technology has the unique capability to reveal some of the perceptual-motor processes underlying handwriting that would not be apparent from classical handwriting evaluations.

The literature provides support for our suggestion that the In Air phenomenon may be an expression of the underlying processing difficulties experienced by children with dysgraphic handwriting. Stott, Henderson & Moyes (1987) wrote that the graphic production of a letter involves two processes. The first process is the formation of a cognitive (or visual) schema of the letter, including knowledge of its defining features and recognition of its
variant forms. The second process involves the development of a motor schema of the letter by which the visual picture is translated into a sequence of movements. The failure to form letters may result from a deficiency in either of these processes. A different explanation offered by Eidlitz & Simner (1999), is that the form errors characterizing the writing of children with dysgraphic handwriting might result from a poorly developed or unstable memory image which gradually fades and then disappears as children print.

Our data, provided by the digitizer, is not sufficient to determine which of the above explanations is correct; rather it allowed us to document the In Air phenomena occurring between letters and letter strokes, and revealed its consistent appearance in both copying and dictation modes of handwriting. Clearly, further research is needed to determine the actual relationship between the In Air phenomena and the underlying mechanisms that contribute to dysgraphia. Indeed such research may contribute to the formulation of new and improved educational strategies to help dysgraphic children, as well as determine the effectiveness of such strategies. For example, consider a child whose performance on a conventional assessment indicates a tendency to have difficulty producing certain letterforms, through repeated erasures or overwriting of the letters. If conventional evaluation was followed-up by a digitizer-based evaluation that revealed prolonged ‘motion tours’, similar to those observed in this study, it would provide the clinician with clues as to the nature of the child's hesitation, and lack of sufficient internalization of these letterforms. Such detailed data, provided by the digitizer, would demonstrate clearly the child's need to spend extra time, energy, and attentional reserves in contemplating how to write and erase that which has not been formed correctly. This type of pictorial representation of the child's difficulties in writing could do more than help focus remediation efforts; it could also be used to illustrate the existing problem to the child, his parents and teachers, as well as provide effective documentation of improvement in handwriting.
The process of using digitizer-based data to discover clues regarding the underlying mechanisms responsible for clinically manifested behaviors has begun to be appreciated in other areas of research as well. For example, in the field of gait analysis, the physical therapist typically uses conventional visual analysis to assess the quality of gait performance in clinical practice (Halligan & Wade, 1998). In contrast, computerized gait analysis is occasionally available to provide clinicians with a quantitative analysis of temporal-spatial parameters (e.g. gait kinematics and kinetics). Such computerized analysis systems enable the objective measurement of phenomena that the human eye cannot discern, but is less suitable for providing an overall judgment of the quality of gait (Halligan & Wade, 1998). As in the current study, researchers have discovered correlations between the objective computerized and the conventional visual gait analysis measures (Damiano & Abel, 1996; O’Sullivan, Said, Dillon, Hofman, & Hughes, 1998; Schwartz, Novacheck & Trost, 2000). The computerized assessment procedures aim to provide a deeper understanding of the basic causes and mechanics of abnormal gait, whereas the clinically oriented assessments are more suited to the resolution of issues concerned with the global characteristics of gait. Such considerations prompted Mulder, Nienhuis & Pauwels (1998) to argue that both conventional observation as well as computerized gait analysis are necessary for a comprehensive characterization and understanding of human gait. Thus, based on findings that seem to parallel those derived from our study, the position of gait analysis scientists reinforces the conclusions we have stated in terms of the evaluation of handwriting. That is, the use of a combination of conventional and computerized objective methods may lead to a better understanding of the handwriting process and product of children with proficient and dysgraphic handwriting.

Our study also served the purpose of examining the importance of including both copying and dictation tasks in the evaluation of handwriting. This information would be quite useful to educators since most handwriting tests targeting children with developmental
dysgraphia evaluate writing in only one mode or the other, despite the fact that school children are required to perform handwriting tasks in both modes.

This question was examined in two ways. First, a correlational matrix was used to assess the extent of convergence between the variables tested. This was done to determine the degree to which the same measure, under two different conditions (copying and dictation) were correlated. As was observed, the digitizer based measures in the copying mode correlated with high magnitude with the corresponding measures in the dictation mode. This appears to indicate that quantitative handwriting measures used in this study test the construct no matter which mode is used, suggesting that testing in one mode would be sufficient. However, the correlations between the corresponding conventional measures in copying versus dictation modes were only moderate, suggesting that the variables being measured are related but are expressions of different underlying mechanisms depending on the mode being used. The conclusion that ensues from this analysis supports the importance of looking at both modes of writing.

The second way in which the need for testing in both writing modes was investigated was through the use of MANOVA and Univariate analyses. Significant differences were found between copying and dictation performance for two of the conventional measures in the dysgraphic group and for all of the digitizer based measures in both groups. No significant differences were found between the conventional measures of the proficient writer’s group. These results clearly support the fact that handwriting should be tested in both modes, especially when using quantitative measures.

One can speculate that the statistical differences in results between copying and dictation tasks resulted from the fact that each writing mode may entail the use of different underlying perceptual-motor processes. Melvin and Levine (1993) provided support for this contention when they stated that in a copying task, the child only needs the ability to imitate,
while in a dictation task the child needs the ability to access the graphic representation of letters and words. This may provide an explanation as to why the scores achieved on the conventional measures in the current study were better in the dictation task, and why the digitizer based In Air time and length measurements were longer in the copying task. Although significant differences were found between writing modes for the children in the dysgraphic writer’s group, it was not found for the conventional measures among the children in the proficient writer’s group, suggesting that proficient writers do not need to be examined in both modes. In light of the results of our study, further research is recommended to expand on the applications and possibilities offered by using both dictation and copying modes in handwriting evaluation and remediation.

Conclusions

The findings of this study serve to expand on the information provided by previous digitizer handwriting research and indicate that a relationship exists between the conventional appearance of the written product of children with dysgraphia and quantitative digitizer measures (In Air time and length). Pooled results from both conventional and digitizer based methods may provide information related to the problematic underlying processing mechanisms of these children. In addition, our findings provide concrete data indicating the importance of utilizing both copying and dictation tasks in the evaluation of children’s handwriting. These results have broad implications for professionals in the educational and clinical fields who are responsible for the identification and remediation of handwriting problems in children. More research is needed aimed at the development of techniques incorporating the use of computerized data in order to establish a comprehensive and systematic evaluation system for children with dysgraphic handwriting of all ages, based on a combination process of examining both handwriting process and product.
Acknowledgements

We would like to thank Dr. Esther Adi-Japha for her creative work in designing and programming the data collection and analysis program. We also thank Yaron Segal for development of the Matlab analysis modules, Gil Rosenzweig for assistance with the research design and statistical analysis, and Sarina Goldstand for editorial assistance. Financial support from the Israeli Ministry of Education is gratefully acknowledged.
References


Modlinger, I. (1983). Disorders of Written Language: Diagnosis and Remedial Teaching. Tel Aviv; Sifriyat Hapoalim.


Pindak, V., & Parush, S. (1996). The Connection between Motor Programming and
the ability of Finger Identify - to Handwriting Quality of Third Grade children.

Unpublished Thesis. Israel, Jerusalem: Faculty of Medicine, Hebrew University, Jerusalem.


Rubin, N., & Henderson, S. E. (1982). Two sides of the same coin: Variation in
teaching methods and failure to learn to write. *Special Education: Forward Trends* 9, 17-24.


Figures

Figure 1: Handwriting samples of a paragraph copying of a typical proficient and a typical
dysgraphic handwriting.

(1A) Stimulus on the computer monitor:

(1B) Sample of proficient handwriting
(1C) Sample of dysgraphic handwriting
Tables

Table 1

Comparison of mean performances of participants in the proficient and the dysgraphic groups, in conventional measures, for both copying and dictation tasks.

<table>
<thead>
<tr>
<th></th>
<th>Conventional (HHE) measures – Copying task</th>
<th>Conventional (HHE) measures – Dictation task</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Proficient</td>
<td>Dysgraphic</td>
</tr>
<tr>
<td>Global legibility</td>
<td>1.26</td>
<td>2.42</td>
</tr>
<tr>
<td>Letters erased or</td>
<td>1.98</td>
<td>4.14</td>
</tr>
<tr>
<td>overwritten</td>
<td>2.62</td>
<td>6.12</td>
</tr>
<tr>
<td>Unrecognizable letters</td>
<td>6.58</td>
<td>8.28</td>
</tr>
<tr>
<td>Spatial arrangement</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: **p<.01
Table 2

Comparison of mean performances of participants in the proficient and the dysgraphic groups, in digitizer-based measures for both copying and dictation tasks.

<table>
<thead>
<tr>
<th>Digitizer-based measures – Copying task</th>
<th>Proficient</th>
<th>Dysgraphic</th>
<th>Sig</th>
<th>Partial eta squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Means</td>
<td>SD</td>
<td>Means</td>
<td>SD</td>
<td></td>
</tr>
<tr>
<td>Total length (mm)</td>
<td>3029.4</td>
<td>569.9</td>
<td>4191.9</td>
<td>1331.0</td>
</tr>
<tr>
<td>On paper length (mm)</td>
<td>978.4</td>
<td>178.5</td>
<td>1079.3</td>
<td>292.5</td>
</tr>
<tr>
<td>In air length (mm)</td>
<td>2051.0</td>
<td>467.8</td>
<td>3112.6</td>
<td>1191.6</td>
</tr>
<tr>
<td>Total time (s)</td>
<td>171.3</td>
<td>55.8</td>
<td>313.1</td>
<td>131.1</td>
</tr>
<tr>
<td>On paper time (s)</td>
<td>47.5</td>
<td>11.6</td>
<td>66.9</td>
<td>25.1</td>
</tr>
<tr>
<td>In air time (s)</td>
<td>123.8</td>
<td>847.</td>
<td>246.2</td>
<td>117.8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Digitizer-based measures – Dictation task</th>
<th>Proficient</th>
<th>Dysgraphic</th>
<th>Sig</th>
<th>Partial eta squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Means</td>
<td>SD</td>
<td>Means</td>
<td>SD</td>
<td></td>
</tr>
<tr>
<td>Total length (mm)</td>
<td>2569.2</td>
<td>434.6</td>
<td>3516.2</td>
<td>1041.0</td>
</tr>
<tr>
<td>On paper length (mm)</td>
<td>862.7</td>
<td>169.4</td>
<td>1000.9</td>
<td>391.5</td>
</tr>
<tr>
<td>In air length (mm)</td>
<td>1706.5</td>
<td>329.2</td>
<td>2515.3</td>
<td>862.3</td>
</tr>
<tr>
<td>Total time (s)</td>
<td>124.4</td>
<td>29.5</td>
<td>208.9</td>
<td>78.3</td>
</tr>
<tr>
<td>On paper time (s)</td>
<td>42.8</td>
<td>10.0</td>
<td>58.8</td>
<td>20.6</td>
</tr>
<tr>
<td>In air time (s)</td>
<td>81.7</td>
<td>24.9</td>
<td>150.2</td>
<td>68.4</td>
</tr>
</tbody>
</table>

Note: *p<.05  **p<.01
Table 3

The degree of discrimination between study groups, for all the conventional measures (HHE) and the digitizer-based measures, for copying and dictation tasks separately.

<table>
<thead>
<tr>
<th>Measures</th>
<th>Copying</th>
<th>Dictation</th>
</tr>
</thead>
<tbody>
<tr>
<td>All eight conventional and Digitizer-based measures</td>
<td>$F(8,87) = 9.95$ P&lt;.0001</td>
<td>$F(8,89) = 25.96$ p&lt;.0001</td>
</tr>
<tr>
<td>Conventional</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global legibility</td>
<td>0.81</td>
<td>0.95</td>
</tr>
<tr>
<td>Letters erased and/or overwritten</td>
<td>0.64</td>
<td>0.59</td>
</tr>
<tr>
<td>Unrecognizable letters</td>
<td>0.77</td>
<td>0.81</td>
</tr>
<tr>
<td>Spatial arrangement</td>
<td>0.57</td>
<td>0.86</td>
</tr>
<tr>
<td>Digitizer-based</td>
<td></td>
<td></td>
</tr>
<tr>
<td>In air time</td>
<td>0.79</td>
<td>0.65</td>
</tr>
<tr>
<td>On paper time</td>
<td>0.63</td>
<td>0.52</td>
</tr>
<tr>
<td>In air length</td>
<td>0.71</td>
<td>0.62</td>
</tr>
<tr>
<td>On paper length</td>
<td>0.29</td>
<td>0.26</td>
</tr>
</tbody>
</table>
Table 4

Correlation matrix between the Hebrew Handwriting evaluation (HHE) measures and spatial and temporal digitizer measures of paragraph copying

4.1 Entire sample

<table>
<thead>
<tr>
<th>HHE variables</th>
<th>COPY (mm)</th>
<th>DICT (mm)</th>
<th>COPY (s)</th>
<th>DICT (s)</th>
<th>COPY (s)</th>
<th>DICT (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global legibility</td>
<td>.42**</td>
<td>.52**</td>
<td>.31**</td>
<td>.41**</td>
<td>.51**</td>
<td>.65**</td>
</tr>
<tr>
<td>Letters erased and/or overwritten</td>
<td>.47**</td>
<td>.50**</td>
<td>.38**</td>
<td>.31*</td>
<td>.54**</td>
<td>.54**</td>
</tr>
<tr>
<td>Unrecognizable letters</td>
<td>.41**</td>
<td>.52**</td>
<td>.36**</td>
<td>.50**</td>
<td>.52**</td>
<td></td>
</tr>
<tr>
<td>Spatial arrangement</td>
<td>.58**</td>
<td>.54**</td>
<td>.31**</td>
<td>.36**</td>
<td>.62**</td>
<td>.57**</td>
</tr>
</tbody>
</table>

4.2 Children with proficient handwriting

<table>
<thead>
<tr>
<th>HHE variables</th>
<th>COPY (mm)</th>
<th>DICT (mm)</th>
<th>COPY (s)</th>
<th>DICT (s)</th>
<th>COPY (s)</th>
<th>DICT (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global legibility</td>
<td>.41**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Letters erased and/or overwritten</td>
<td>.33*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unrecognizable letters</td>
<td>.31**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spatial arrangement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4.3 Children with dysgraphic handwriting

<table>
<thead>
<tr>
<th>HHE variables</th>
<th>COPY (mm)</th>
<th>DICT (mm)</th>
<th>COPY (s)</th>
<th>DICT (s)</th>
<th>COPY (s)</th>
<th>DICT (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global legibility</td>
<td>.35**</td>
<td></td>
<td>.33*</td>
<td></td>
<td>.39**</td>
<td></td>
</tr>
<tr>
<td>Letters erased and/or overwritten</td>
<td>.35**</td>
<td>.24*</td>
<td>.44**</td>
<td>.27**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unrecognizable letters</td>
<td>.30*</td>
<td></td>
<td>.36*</td>
<td>.34**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spatial arrangement</td>
<td>.50**</td>
<td>.31*</td>
<td>.59**</td>
<td>.35**</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

COPY = copying task  DICT = Dictation task  *p<.05  **p<.01
Table 5

Correlations between copying and dictation (HHE and Digitizer variables) of entire sample

<table>
<thead>
<tr>
<th>COPY/DICT Correlation</th>
<th>0.84</th>
<th>0.86</th>
<th>0.80</th>
<th>0.81</th>
<th>0.70</th>
<th>0.54</th>
<th>0.58</th>
<th>0.64</th>
</tr>
</thead>
<tbody>
<tr>
<td>N=100</td>
<td>DICT</td>
<td>COP</td>
<td>**p&lt;.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>