

Product and Process Evaluation of Handwriting Difficulties

Sara Rosenblum,^{1,2,3} Patrice L. Weiss,² and Shula Parush¹

Handwriting is a complex human activity that entails an intricate blend of cognitive, kinesthetic, and perceptual-motor components. Children are expected to acquire a level of handwriting proficiency that enables them to make skillful use of handwriting as a tool to carry out their work at school. Poor handwriters have difficulty developing their writing skills and, as a result, often suffer in their educational and emotional development. This article highlights the importance of handwriting and reviews the development of methods used to evaluate handwriting difficulties. Included also is a discussion of methodological aspects of current handwriting evaluations and a presentation of research on the use of a computerized system that may be helpful in better understanding the handwriting process of poor writers. The article concludes by outlining future directions in handwriting evaluation that combine the assessment of the handwriting product with computerized analysis of the handwriting process.

KEY WORDS: handwriting assessment; legibility; handwriting speed; computerized analysis.

INTRODUCTION

Handwriting is a complex human activity that entails an intricate blend of cognitive, kinesthetic, and perceptual-motor components (Bonny, 1992; Reisman, 1993). To produce written text a student must initiate and execute simultaneously a number of motor and cognitive tasks including

¹Faculty of Medicine, School of Occupational Therapy, The Hebrew University, Jerusalem, Israel.

²Department of Occupational Therapy, Faculty of Social Welfare & Health Studies, University of Haifa, Haifa, Israel.

³To whom correspondence should be addressed at Department of Occupational Therapy, Faculty of Social Welfare & Health studies, University of Haifa, Mount Carmel, 31905 Haifa, Israel. E-mail: rosens@construct.haifa.ac.il

ideation, planning, text production, spelling, punctuation, grammar, self-monitoring, evaluation, and orthographic-motor integration (Berninger, 1994; Hooper et al., 1993; Jones and Christensen, 1999). Handwriting skills, particularly handwriting fluency, improve with age and schooling (Graham et al., 1998; Hamstra-Bletz and Blote, 1990). During their first 3 years of school, children are expected to acquire a level of handwriting proficiency that enables them to make skillful use of handwriting as a tool to carry out their work at school (Laszlo and Broderick, 1991; Maeland and Karlsdottir, 1991). As of the fourth grade, writing assignments become longer and more frequent. Children are required to hand in papers, write essays, and give longer responses to test-questions (Cornhill and Case-Smith, 1996; Reisman, 1993). Most children find that they are ready to handle these demands, and the proficiency of their handwriting is reflected by their ability to produce legible text with minimum effort. Furthermore, for typical children, handwriting becomes automatic so that text generation does not interfere with their creative thinking process (Scardamalia et al., 1982).

Those children who do not succeed in developing proficient handwriting are defined by some authors as “poor writers” and by others as “dysgraphic” (Marr and Cermak, 2001). Hamstra-Bletz and Blote (1993) defined “dysgraphia” as a disturbance or difficulty in the production of written language that has to do with the mechanics of writing. The difficulty is manifested in the inadequate performance of handwriting among children who are of at least average intelligence level and who have not been identified as having any obvious neurological or perceptual-motor problems. It is reported that the prevalence of handwriting difficulties among school-aged children varies between 10% and 34% (Rubin and Henderson, 1982; Smits-Engelsman et al., 1995, 2001). Handwriting difficulties are especially prevalent among children diagnosed with Developmental Coordination Disorder [DSM-4, American Psychological Association (APA), 1994] and learning disabilities (Waber and Bernstein, 1994) or who are defined as clumsy by their teachers (Laszlo, 1990; Laszlo et al., 1988).

Because 30%–60% of a child’s school day is spent in the performance of fine motor tasks, consisting primarily of handwriting tasks (McHale and Cermak, 1992), it is likely that the quality of one’s handwriting skill affects academic performance. Several authors have suggested that difficulty in the mastery of the mechanical aspects of handwriting may interfere with higher order processes required for the composition of text (Berninger and Graham, 1998). Graham (1990) found that handwriting mechanics influence the quality and quantity of the written product. This finding is supported by that of Berninger et al. (1997), who reported that handwriting performance was significantly related to fluency and quality of composition in elementary school students. Graham et al. (2000) summarized views on the negative

implications of handwriting difficulties in a recent article. They, together with others (Briggs, 1980; Chase, 1986; Hughes, Keeling, and Tuck, 1983), have suggested that teachers tend to give higher marks for neatly written papers than those for messy ones. It thus 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 et al., 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 (Berninger and Graham, 1998; Jones and Christensen, 1999). 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 and Weintraub, 1996).

Unfortunately, some children who have difficulty mastering handwriting skills may avoid writing altogether, resulting in arrested writing development (Berninger et al., 1991), or become less willing to devote the extra effort needed for planning composition or revising their work (McCutchen, 1996). These avoidances affect children's performance in a circular fashion because increased writing may help improve handwriting quality (Graham, 1992). Some of these children may respond by simply giving up, having developed a mind-set that they cannot write (Berninger et al., 1991). In fact, the results of a series of studies carried out in Canada and in the United Kingdom have indicated that difficulties with handwriting in the early years might be used as a predictor of more general learning difficulties later on (Harvey and Henderson, 1997; Simner, 1982, 1985, 1986, 1990).

Problems that stem from difficulties in handwriting do not disappear when a student graduates from school. In fact, in many cases, the problems become more complicated and difficult to resolve. The demands made in the course of pursuing a higher education or of advancing in the workplace can often exacerbate an already difficult situation. In some cases, the individual's problems become masked and further complicated by increasing stress. It appears that inadequate handwriting can affect many areas of life, resulting in a loss of self-confidence, and may have serious consequences for career prospects and even personal relationships (Sassoon, 1997).

In summary, it appears that the quality of handwriting has a marked effect on the writing and academic performance of school-aged children (Berninger et al., 1997; Graham et al., 2000; Jones and Christensen, 1999)—a finding that reinforces the importance of identifying handwriting difficulties as early as possible. School psychologists often play a role in the overall management of handwriting difficulties of school children, whether as consultants to teachers as to how to assess handwriting, as planners and evaluators of

programs that include goals for improving handwriting performance, or as evaluators of children's progress. Therefore, it seems advisable that school psychologists be familiar with the issues and procedures that are related to handwriting assessment (Graham, 1986a,b).

Over the years, many methods have been developed for the evaluation of handwriting difficulties. Most are based on analyzing the handwritten product and speed. These evaluations formed the basis for research into the developmental sequence of writing and in the clinical identification of children with handwriting problems. Comparative studies of the handwritten output of children with and without handwriting difficulties reveal differences in the accuracy and readability of letters, words, and sentences. The handwriting quality of children with difficulties has been described in studies as "poor" and can be characterized by inappropriate spacing between letters or words, incorrect or inconsistent shaping of letters, poorly graded pencil pressure, letter inversions, and mixing of different letter forms (i.e., script and square) (Hamstra-Bletz and Blote, 1993; Kaminsky and Powers, 1981; Maeland and Karlsdottir, 1991; Rubin and Henderson, 1982; Sovik et al., 1987a,b). The process of describing the features that characterize the written output of children with handwriting difficulties has formed the basis for the development of scales for handwriting evaluation.

The purpose of this article is to critically review the various methods used to evaluate handwriting difficulties. To date, the most commonly used methods have been (1) global-holistic evaluations of legibility and (2) analytic evaluations that assess readability in relation to predetermined criteria. This article's first section provides a historical review of these evaluations and describes the process by which they were developed and the results of studies carried out to determine their psychometric properties. This article's second section addresses issues related to methodological aspects of these evaluations. The article's third section is devoted to a description of recently published studies on the use of computerized methods for understanding the handwriting process. The methods use real-time measures of various performance criteria during the actual performance of handwriting. The article concludes with a brief discussion of future directions in the evaluation of children with poor handwriting that focuses on the combination of handwriting product evaluation with computerized analysis of the handwriting process.

A HISTORICAL VIEW OF HANDWRITING EVALUATIONS

Assessment of Handwriting

The primary aim of researchers who composed the various handwriting evaluation scales was to develop standardized evaluations capable of

producing quantitative scores for handwriting quality (Chu, 1997; Reisman, 1991; Rubin and Henderson, 1982). Their dilemma was how to define the “quality of handwriting” or “readability” (Ayres, 1912) in specific, measurable terms. The handwriting evaluations that were developed over the years can be categorized as either global-holistic evaluations of handwriting “readability” or analytic evaluations that rated the readability of a handwritten product in relation to predetermined criteria. The global evaluation scales are used to form an overall judgment of a written product in terms of how readable it is in comparison to a group of standard handwriting samples previously graded from “readable” to “unreadable.” In contrast, the analytically based evaluations are based on the assumption that a relationship exists between the general look (i.e., the readability), and certain criteria of performance, such as the shaping of the letters, the spaces between the letters and the words, etc. The handwriting sample is judged by grading each criterion individually for the passage and then calculating an overall score.

The Global Holistic Evaluations in the Early Years

Research concerning the development of handwriting evaluation scales was conducted as early as the second decade of the twentieth century. One of the earliest reported scales developed for the evaluation of handwriting was a global-holistic scale for students from fifth to eighth grade (Thorndike, 1910). A handwriting product was first evaluated for “general merit.” Then it was graded according to the average value awarded it by a group of judges who compared it to the graded handwriting samples that were provided (Tseng and Cermak, 1991). Ayres (1912) also developed a global handwriting evaluation scale that was based on the rating of the “general merit” of handwritten products. In contrast to the previous scale, the grading criterion used by Ayres (1912) was the median time taken by 10 judges to read the passage. However, researchers remained unsatisfied with either evaluation, deeming them as impractical for regular use in the schools and far too subjective and unreliable (Starch, 1919). In response to dissatisfaction with the existing scales, Freeman (1959) proposed that the reliability of handwriting scales be improved by using clearly defined criteria by which to grade handwriting samples. For this purpose, he developed a scale for assessing handwriting samples that included a system for grading handwriting quality according to the following five criteria: tilt, height, shaping of letters, line quality, and general merit (Freeman, 1959). Though more comprehensive, these scales still provided only a crude grading system. Therefore, Freeman (1959) later revised this scale and substituted what he considered

to be general excellence for the use of specific criteria in evaluation of handwriting (Tseng and Cermak, 1991).

During the following years, additional attempts were made to produce an improved handwriting scale with more accurate scoring criteria. In 1962, Bezzi presented a scale similar to the one developed by Freeman in 1959, but it did not really represent a significant improvement. The Wisconsin scale, developed by Herrick and Elebacher (1963) in the following year, represented an attempt at improving accuracy by providing, for each of three school grades, 200 samples of handwriting graded according to letter size, tilt, and readability. Not surprisingly, this scale was considered inefficient for use in schools because of the difficulty and time required to distinguish between the samples. It was subsequently used for research purposes only (Herrick and Elebacher, 1963; Tseng and Cermak, 1991).

In summary, the examination of the early literature on handwriting evaluation development demonstrates a distinct trend in which global scales (Ayres, 1912; Thorndike, 1910) paved the way for the development of the analytic approach (Bezzi, 1962; Freeman, 1959; Herrick and Elebacher, 1963).

Global Readability Assessment in Recent Years

During the past 20 years, some researchers developed a renewed interest in attempting to develop global readability assessment methods. The developers of these tests stressed the need to use experienced evaluators who have practiced scoring a minimum of writing samples before scoring research samples.

The *Test of Legible Handwriting* (TOLH) (Larsen and Hammill, 1989) is designed to evaluate the overall readability of manuscript (print) and cursive writing of children from the 2nd to 12th grade. Originally called TOWL (Test of Written Language), the authors of the TOLH constructed a scale of writing samples graded from 1 to 9 (from “least” to “most readable”). Writing samples consisted of written stories based upon pictures or passages written by the students during school. The samples were made up of three writing types: print (i.e., manuscript), script writing (i.e., cursive) that was tilted vertically or to the right, and script writing tilted to the left. The objective of the evaluator is to match the written passage, as closely as possible, to one of the given samples. The written product’s readability is given standard and percentile scores, and an informal protocol is prepared to summarize the analysis of the child’s mistakes. The early version of the TOLH was used by Graham, Boyer-Schick, and Tippets (1989), who showed that evaluators who only complete the minimal practice requirements of the manual did not obtain reasonable interrater reliability when evaluating the writing of children

with learning difficulties. The researcher's team scored 70 writing samples for which a high interrater reliability ($r = 0.95$) was calculated. However, because reliability is most properly determined by outside researchers, this result is not considered definitive (Graham and Weintraub, 1996). Graham and Weintraub (1996) criticized the validation process of the scale. Although the authors conducted several studies that support the validity of the TOLH, these studies typically involved small numbers of handwriting samples. Of equal importance, the data presented provide little insight into what the test actually measures (Graham and Weintraub, 1996). Thus, although this scale is unique in its capacity to evaluate three types of writing, further research is necessary to determine its psychometric properties (Graham and Weintraub, 1996).

Despite such criticism, the TOLH scale was used to study the relationship among writing styles (manuscript, cursive, mixed-mostly manuscript, and mixed-mostly cursive), writing speed, and readability (Graham et al., 1998). To improve the reliability of the study results, the evaluators participated in a training workshop. The results indicated that students who combined slanted manuscript with cursive handwriting had the most fluent writing. However, no differences were found between the readability of passages written in either of the styles. In another study based on the TOLH, Simner (1996) examined its ability to identify children at risk during the first years of school. The assumption was that illegible handwriting is related to the general level of learning performance of the child. Results of handwriting samples evaluated by the TOLH were correlated with CIBS (Comprehensive Inventory of Basic Skills; Brigance, 1983) achievement scores and their subsequent reading, spelling, and arithmetic performance in second grade. In light of these results, Simner (1996) concluded that it is possible to use the TOLH to identify children at risk for future learning disabilities. In a recent study, the TOLH was used by classroom teachers to select experimental groups of poor and proficient handwriters for research (Graham et al., 2001).

The Evaluation Tool of Children's Handwriting (ETCH) (Amundson, 1995) was developed by an occupational therapist for the purpose of evaluating the readability and handwriting speed generated on written tasks that are similar to those expected in the classroom. One part of the tool tests manuscript (print) writing (ETCH-M) and the other tests cursive handwriting (ETCH-C). The time needed to administer each part of the ETCH is 20–30 min (Shneck, 1998). The writing tasks include writing uppercase and lowercase letters from memory, writing numbers from memory, copying a near-point text, copying a text from a distance, dictation, and composing a sentence.

Scoring focuses on overall readability, writing speed, component features of readability, and biomechanical aspects of writing. The evaluator

counts and scores occurrences of various readability components (such as shape, size, and spacing). The mechanical aspects of the child's writing, such as pencil grasp, pencil pressure, and in-hand manipulation, are observed during task performance and noted on the evaluation sheet (Diekma et al., 1997). Evaluators, professionals in health and education, are expected to practice the scoring procedures described in the instruction manual. They are required to achieve scoring competency on the trial tests in the manual before attempting to score children's performances in the practice setting (Amundson, 1995).

The interrater reliability studies for the ETCH completed by the test developer showed moderate-to-high results for different parts of the ECTH-M and the ECTH-C (see Table I) (Amundson, 1995). Test retest reliability for readability, according to studies of the ETCH-M that were conducted on first- and second-grade children, was moderate (see Table 1) (Diekma et al., 1997). These results did not demonstrate that the ETCH scales had better reliability than previous scales, a disappointing finding for its authors (Alston, 1983; Phelps et al., 1985; Stott et al., 1984; Ziviani and Elkins, 1984). However, Shneck (1998) points out that in contrast to the reliability studies done for prior assessment scales, the ETCH-M was researched among children who have handwriting difficulties, which would tend to reduce its reliability (Diekma et al., 1997). On the other hand, Shneck asserts that because it is based on global readability, the ETCH method is, in fact, a subjective evaluation. Diekma et al. (1997) suggest that therapists take into account the limited reliability of a writing assessment tool (i.e., its subjectivity and absence of studies applicable to children with handwriting difficulties) when planning to use it for assessing the efficacy of treatment. In fact, no significant relationship was found between the ETCH scores and teacher questionnaire scores in either general legibility or task-specific legibility (Sudsawad et al., 2001). Thus, it has been suggested that further changes for scoring criteria are warranted before the ETCH scores are considered related to actual performance in the classroom as determined by teachers (Sudsawad et al., 2001).

The writing speed and readability of 372 typical children aged 7–14 years in Australia was measure by Ziviani and Watson-Will's scale (Ziviani and Watson-Will, 1998). Unlike the methods used previously (see above; Ziviani and Elkins, 1984), this scale evaluates the global readability of handwriting, measuring the written product on a 7-point scale. No significant differences were found between boys and girls in mean writing speed. However, the readability of the girls' handwriting was significantly better than that of the boys. A low correlation was found between writing speed and readability (Ziviani and Watson-Will, 1998). Reliability studies were not found in the literature.

ANALYTIC EVALUATIONS OF READABILITY

Most of the handwriting evaluation scales developed during the last 25 years belong to those of the analytic category; that is, evaluation according to specific criteria of “readability” that can be defined objectively. Although most researchers of analytic scales agreed as to what those criteria are—letterform, size, slant, spacing, and line straightness (Bruinsma and Nieuwenhuis, 1991), the approaches used to actually measure the criteria have varied. In essence, this was a period of trial and error in which researchers attempted to find the best combination of objectivity and utility for a handwriting product evaluation. In this section, the major analytic handwriting evaluation methods are described in a chronological sequence. To facilitate comparisons between the scales, psychometric data and handwriting speed values are presented in Table I and Table II, respectively.

The first analytic approach to evaluating a handwriting sample was referred to as the “transparent overlays” method (Collins et al., 1980; Helwing et al., 1976; Jones et al., 1977). Scoring is based on the use of transparent overlays to determine if specified standards of performance have been met, and to assess topographic features such as shape, size, and other descriptive criteria. The writing sample is printed on top of a transparency and the examiner compares each letter of the written passage to what is printed on the transparency. Letters that protrude out from the boundaries of the letters on the transparency are considered mistakes taking into consideration criteria such as stability of the pen’s stroke and consistency of the letter size, among others. For example, to assess letter shape, the examiner typically determines if the target letter fits within a 1- to 3-mm wide outline of the letter (Graham, 1982). If the writer adds a little flourish at the end of the letter or writes it a little larger than the standard, it is scored as incorrect. Graham (1982) noted that this technique was highly reliable, with inter-rater reliability coefficients ranging from 0.86 to 0.97. However, evidence on the validity and utility of these instruments is virtually nonexistent. Collins et al. (1980), Sims and Weisberg (1984), and more recently Graham and Weintraub (1996) concluded that despite their reliability, “transparent overlay” evaluations lack the sensitivity needed to monitor gradual improvement adequately, yet are overly sensitive to variations in personal style. In addition, this method’s construct and concurrent validity have not been adequately addressed.

The scale of Rubin and Henderson (1982) was developed to enable teachers to identify children with handwriting difficulties. Following a few trials, six assessment criteria were chosen: readability, accuracy of letter formation, unity of letters size and letters tilt, spaces between letters and words, and straightness of the written line. A 4-point scale was developed for each

Table 1. Reliability and Validity Values of Analytic and Global Handwriting Scales from Recent Years

Handwriting evaluation scale	Reliability		Validity
	Intrater	Test-retest	
Alston's evaluation scale (1983)	0.64-0.68	0.63-0.82	—
BHK (Hamstra-Bletz et al., 1987)	0.76-0.89		0.78-Teachers scores
CHES (Phelps et al., 1985)	0.64-0.82		
DRHP (Stott, Moyes, and Henderson, 1984)	0.56-0.66		
ETCH (Amundson, 1995)	0.85-0.92	0.63	
ETCH-M	0.53-Total numbers		
ETCH-C	0.30-Total words		
	0.97-Total letters		
HHE (Erez et al., 1999)	0.75-0.79	0.95	
Minnesota (Reisman, 1993)	0.99-Experienced		
	0.77-0.83- Unexperienced		
	0.95-Researcher and her assistants		
Rubin and Henderson (1982)	0.95	0.97	
TOLH (Larsen and Hammill, 1989)	0.95		
"Transparent overlays" (Collins, Baer, Walls, and Jackson, 1980)	0.86-0.97		
Ziviani and Elkins (1984)	0.76-0.97	0.44-0.93	0.52-0.76
Ziviani and Watson-Will (1998)	—	—	

Table II. Results of Research into the Average Writing Speed of Children (Characters Per Minute) (Erez et al., 1999; Graham, 1986a,b; Tseng and Cermak, 1991; Ziviani and Watson-Will, 1998)

Country Group Grade	Ayres, Freeman 1912 1959		Groff Ziviani and 1961 Elkins 1984		Phelps et al., 1985		Sasson et al., 1986		Phelps et al., 1988		Hamstra-Bletz and Blote, 1990 ^a			Wallen et al., 1996		Erez et al., 1999	
	U.S. 1578	U.S. 810	U.S. 575	Australia 1365	U.S. 1365	England 294	U.S. 1365	U.S. 1365	Germany 127	Australia 1292	Israel 900	Cop ^b	Dic ^b				
1						U	R		1	2	3						
2		30				46	55	32	24	25	25						
3			33		25	64	82	35	34	39		38	49				
4	55		34		37			46	42	49	54	46	57				
5	64		38		47			54	59		64	81					
6	71		46		57			66			94						
7			52		62						100						
8		80			72						115						
9											116						
10						117	140				124						
11											133						
12																	

^a"1 2 3" denotes that there were three different groups.
^b"Cop" denotes copying tasks and "Dic" denotes dictated tasks.

of the criteria. The children were asked to copy a paragraph of 57 words on unlined paper within 5 min. Writing speed was calculated as the number of letters written per minute. Both test-retest reliability and interrater reliability of the scale were extremely high (see Table I).

In developing the Alston Evaluation Scale (1983), the authors used a novel approach to handwriting assessment. A 20-item questionnaire was constructed to gather information from teachers regarding features defined by the researcher as influencing readability (e.g., letter formation, letter size, spacing, and straightness of the written line). The questionnaire includes questions such as “Are the letters that are supposed to be rounded indeed rounded?” The assessment was designed to measure children’s performance on a freestyle writing task: children were asked to write an essay about their favorite person on lined paper within 20 min, using any writing tool. Both interrater reliability and construct validity of the scale were moderate to high (see Table I). A further study conducted on this scale found that only 15% of the 23 items had a significant relationship with readability, leading to the recommendation that the scale be redefined (Graham and Weintraub, 1996; Tseng and Cermak, 1991). However, no further study concerning the scale appears in the literature.

Ziviani and Elkins (1984) developed an evaluation scale for manuscript (printing) handwriting ability of children aged 7–14 years. Handwriting ability was judged on the basis of the evaluation of readability components (letter formation, size, spacing, and straightness) and on handwriting speed. Exact specifications for forming individual letters and symbols were defined to enable the most objective measurements possible. Children were asked to copy shapes, letters, and words. Their speed was determined by the number of times that they succeeded in writing the phrase “cat and dog” within 2 min. A transparent overlay with straight lines drawn on it was used together with a ruler to measure the spaces between the words or the deviations of words from the horizontal alignment.

The authors conducted studies on the reliability and validity of the scale. Interrater reliability and test-retest reliability were found to be moderate to high. Content validity of Ziviani and Elkins’ scale was investigated by use of a table of specifications examining whether legibility components as found in the research literature were represented and by examining the scale’s internal consistency (Ziviani and Elkins, 1984). Results showed a moderately high agreement between items measuring the same legibility components (see Table I). Item analysis was performed to determine which criteria had the greatest influence on readability: shape, spacing, size, or line straightness. Criterion validity for legibility, determined by comparing the handwriting evaluation results to the teachers rating on students handwriting samples, yielded a moderate validity coefficient (see Table I). Norms for speed were

compared to those found in studies by Ayres (1912) and Groff (1961) and were consistent with Groff's data (Groff, 1961). Graphs indicating normative performance on the handwriting tasks were prepared.

The advantage of Ziviani and Elkins' tool, as compared with its predecessors, is in the resolution with which criteria were defined (i.e., spacing and size were measured to the nearest millimeter through the transparent overlays). Unfortunately, the authors do not clarify how they determined these criteria as critical ones that support readability (Stott et al., 1987). In addition, the scale makes no attempt to measure inconsistency of size and slant or distortion of letter forms due to poor perceptual-motor control, despite the fact that these are the components of legibility that pose the greatest challenge for measurement (Stott et al., 1987).

The Children Handwriting Evaluation Scale (CHES) was developed by Phelps et al. (1985) to enable teachers and clinicians to measure quality and fluency of cursive handwriting among third to eighth graders. The scale, part of an evaluation battery to identify children suspected of having learning disabilities, is comprehensive and contains specifically defined criteria (Phelps et al., 1985). Children are first asked to read a story (containing 197 letters) and then to copy it onto a blank paper "in the same manner in which they usually write." To evaluate handwriting speed, the examiner marked the point the child reached after 2 min of copying. The number of letters copied was compared with a table defining writing speed norms for children of different ages. On that basis, a score was assigned according to a 5-point scale, ranging from "worst" to "best." The quality was also judged on a scale of 5 points according to a number of criteria: letter shapes, tilt, rhythm, spacing, and general look. Of the total 1352 writing samples collected, 150 were evaluated by each of two researchers, a speech therapist, and a teacher to determine the scale's reliability. Interrater reliability was moderate to high (see Table I). Handwriting speed norms that were provided (Phelps et al., 1985) for Grades 3–8 are appreciably lower than those observed in other investigations (see Table II). In light of the success of the CHES, a similar scale, the CHES-M, was developed to evaluate manuscript (print) writing for children in first and second grades (Phelps and Stempel, 1988).

There are various critics of the CHES scale. According to Graham (1986a,b), the CHES developers do not specify if it is meant as a screening or an evaluation tool. Graham (1986a,b) contends that a tool constructed according to a 5-point scale is not sensitive enough to pick up slight changes that might result from maturation, age, or treatment. Daniel and Froude (1998) claim that asking children to write on a blank paper probably affects their performance. Similarly, Burnhill et al. (1983) found that the quality of children's handwriting on blank paper suffers as compared with handwriting

on lined paper. Contradicting results were found among various researchers on this issue (Krzesni, 1971; Lindsay and McLennan, 1983).

The Concise Evaluation Scale for Children's Handwriting—BHK (translation from German) (Hamstra-Bletz et al., 1987) was developed as a screening tool to examine the readability and speed of writing performance in young dysgraphic children. The authors of the BHK chose a writing task that resembles school-type assignments. Specifically, children are asked to copy a standard text that is presented to them on a card for 5 min. The text is graded according to the complexity of its contents. The first five sentences are composed of one-syllable words at first-grade level and the following sentences are progressively more complex.

The writing passage is evaluated by judging deviations of the child's writing from the standard handwriting text according to 13 criteria. A total score on all 13 criteria items is calculated to determine writing quality which is subsequently used to categorize the child as a poor or proficient writer. Writing speed is calculated according to the number of letters written in 5 min.

The BHK is distinguished by the amount of research devoted to investigating its psychometric properties, by its development of norms for second and third graders, and by its use among children in various populations (Blote and Hamstra-Bletz, 1991; Hamstra-Bletz and Blote, 1990, 1993; Reinders-Messelink et al., 1996; Smits-Engelsman et al., 1996). Interrater reliability for the total score is high (see Table I). The percentage agreement for the single items is 80% (Hamstra-Bletz et al., 1987; Hamstra-Bletz and Blote, 1990, 1993). The BHK scores correlate well with teachers' evaluations of writing quality ($r = 0.78$). Longitudinal studies conducted in Germany with 127 schoolchildren from the second to seventh grade found that the test is sensitive to developmental changes during the elementary school years (Blote and Hamstra-Bletz, 1991; Hamstra-Bletz and Blote, 1990). The BHK scale has also been found to discriminate between children with and without dysgraphia (Hamstra-Bletz and Blote, 1993), a finding recently confirmed by Smits-Engelsman et al. (2001). As a result, its authors suggest that the BHK can be used in the early identification of children with handwriting difficulties (Hamstra-Bletz and Blote, 1993).

The BHK's diagnostic sensitivity is further illustrated by the results of two other studies. Reinders-Messelink et al. (1996) used the tool to evaluate the performance of children with Acute Lymphoblastic Leukemia who had received chemotherapy. They found that the tool is sensitive enough to identify late long-term effects of the therapy: 2 years after 17 children participating in the study received chemotherapy, difficulties were found in the writing and fine-motor activities. Smits-Engelsman et al. (1996) used the BHK to test the efficacy of physiotherapy given to children aged 7–11 years

who had writing difficulties. The children were evaluated by the BHK before and after therapy. At the same time, the evaluation was also used with a control group of children who were identified as needing therapy but were still waiting to receive it. It was found that the physiotherapy improved the quality and writing speed of the children who received it, whereas no such change occurred among children in the control group.

An evaluation scale developed by Stott et al. (1984) is part of a system designed to evaluate and treat handwriting difficulties. It is called the *Diagnosis and Remediation of Handwriting Problems (DRHP)*. The goal of the three-part evaluation is to identify one of three possible *causes* for handwriting difficulty:

1. Features such as improper spacing and letter shaping, which are related to flawed teaching or learning of writing rules.
2. Performance mistakes, such as inconsistency in letter size, tilt, and the tilt of words on the line that can occur from a lack of perceptual-motor control.
3. Inefficient writing manner and position. This part of the assessment involves the direct observation of the writer, but it is not formally encoded.

The DRHP evaluation scale includes the measurement of quantitative and qualitative mistakes made during performance relational to handwriting samples provided to help the evaluator score each of the features. The overall score is based on whether the mistakes in writing affect its readability. The novelty of this evaluation is in the combination of an observational evaluation together with an analysis of the written product. However, clear instructions are lacking to guide the process of test scoring and interpretation. In addition, the research into the scales' psychometric properties is weak: norms were not provided for the scores and test-retest reliability and validity studies were not reported (Stott et al., 1984, 1987; Tseng and Cermak, 1991). Interrater reliability was studied and found to be moderate (see Table I).

Occupational therapists working within School Health Support Services are receiving increasing numbers of referrals of children who have handwriting difficulties (Miller et al., 2001). Therefore, the *Minnesota Handwriting Test* (Reisman, 1991, 1993) was developed to assist school-based occupational therapists in the identification of children with handwriting difficulties and to assess treatment efficacy. The assessment was normed on first and second graders. It evaluates manuscript (print) writing and is supposed to be sensitive to small changes in performance. The children are asked to copy a typed sentence on the lines beneath it. Because the sentence presented to them is a common one ("The quick brown fox jumped over the lazy dog"),

the words are printed in jumbled order to eliminate the advantage of children who read better or have better memories. After writing for 2.5 min, the children are asked to stop and circle the last letter written and then to continue writing.

Handwriting quality is evaluated according to the same five criteria mentioned for previous assessments: readability, shape, line straightness, size, and spacing (Freeman, 1959; Graham, 1982; Kaminsky and Powers, 1981; Rubin and Henderson, 1982; Ziviani and Elkins, 1984). The scoring process begins with letter readability, followed by the other four criteria mentioned above, because the researchers reasoned that if a letter is not legible it is not possible to measure the remaining criteria reliably. Writing speed is calculated according to how many letters are written in 2.5 min. The author mentions other important elements of writing that can be observed during task performance as well such as pencil holding, attention to the assignment, and posture; however, these elements are not included in the scoring manual. The latest version of the manual provides three sets of 10 writing samples on which evaluators can practice scoring. Reisman (1993) includes scored writing samples for instruction and comparison.

Interrater reliability of the Minnesota Handwriting Test was studied by the author (Reisman, 1993) and other researchers (Krupa, 1991; Lilly, 1987) and was found to be high for both experienced and nonexperienced evaluators (see Table I). Moreover, the correlation between experienced and nonexperienced evaluators was also high (Reisman, 1993). Because all evaluators in the reliability studies learned the test by referring to the instruction manual, this correlation between experienced and nonexperienced evaluators indicates the clarity of the manual's instructions.

All of the above-mentioned assessments were developed for languages using a Latin-based character set. The Hebrew written language contains unique features that would make it impossible to evaluate by use of these evaluation tools. For example, text is written from right to left, spaces occur not only between words and letters but within the letters themselves (for example: פ * ם), letters do not connect (a fact that results in many breaks while writing) and, furthermore, some letters change their form when they terminate a word (Modlinger, 1983).

The *Hebrew Handwriting Evaluation* (HHE) (Erez et al., 1996, 1999) was developed to assess the handwriting of children suspected of having difficulty writing in Hebrew. Children are asked to perform three assignments:

1. Copying the letters of the Hebrew alphabet (in atypical order to avoid a possible influence of the familiar order on the copying).
2. Copying a short story (of 30 words) onto lined paper.

3. Writing a short story from dictation (also containing 30 words) onto lined paper.

The tool enables the assessment of four factors: writing speed, writing quality, ergonomic factors, and writing mistakes. Writing speed is measured by the number of letters written in 1 min. Writing quality is tested along two dimensions, letter shaping and spatial organization, each of which is then subdivided into a number of items. All items from both dimensions are scored according to a Likert scale (1–4) using detailed, accurate criteria (spatial organization is measured in millimeters) ranging from “very good” to “very poor.” An overall score for each of the two dimensions is the summation of the respective individual item scores. Ergonomic factors—pressure, pencil grasp, grip consistency, body posture, paper position, and stabilization—are scored according to defined criteria, again on a scale of 1–4, with 1 indicating “good performance” and 4 indicating “poor performance.” Writing mistakes are counted in each of the passages written by the child.

Interrater reliability for the HHE scale was moderate to high (see Table I). Test-retest reliability was not reported. Norms for handwriting speed of second and third graders were reported in the test manual (Erez et al., 1999). Construct validity was indicated by the significant differences found to exist between children who write well and those who have difficulties (Dvash et al., 1995; Lifshietz and Parush, 1996).

The internal reliability of the tool was calculated for all of the items within each of the two dimensions rated. Reliability, calculated separately for copying and dictation, was high (see Table I). This indicates that the HHE’s test items assess the same skill content areas or different aspects of the same skill (Erez et al., 1999). Finally, the scale distinguished the performance of children with and without handwriting difficulties with regard to the ergonomic factors measured (i.e., pencil position, paper position, body posture, body stabilization, and the affect of fatigue) (Parush et al., 1998a,b).

In summary, during the past 25 years a wide variety of global and analytic handwriting evaluation scales have been developed in an attempt to find an optimally reliable and practical method of assessing handwriting. The next section presents a discussion of the relative merits and limitations of the different methodological approaches that were used.

METHODOLOGICAL ISSUES RELATED TO ANALYTIC AND GLOBAL HANDWRITING EVALUATIONS

From the historical review presented in the first section of this article, the consensus of opinion among most handwriting evaluation developers is

that general readability is an important factor in judging the quality of the written product. Similarly, there is agreement among the authors of analytic handwriting scales regarding the main qualitative criteria by which writing readability should be judged (size, slant, spacing, shape, general merit) (Bruinsma and Nieuwenhuis, 1991). However, there exist many methodological variations among the scales in terms of factors that may affect students' outcome scores. These factors include the nature of the handwriting assignments, instructions given to the examinees, writing accessories, specific assessment criteria, and methods for measuring handwriting speed. Handwriting scales also differ as to the extent of the investigation into their psychometric properties and the applicability of the scales to different populations. Finally, scales often differ with regard to the type of evaluator, sensitivity to variability in personal writing style, practicality of the evaluation's administration, the nature of the examinees' involvement in the process, and other varied performance factors that may influence test outcome. This section includes a discussion of these factors and describes how they may impede progress in the development of a maximally reliable and effective handwriting assessment tool (Bonny, 1992; Graham, 1986a,b; Rubin and Henderson, 1982).

The Evaluator

Most of the assessments do not specify who is certified to administer them, whether it be a teacher, a therapist, or a student's self-assessment, nor do they specify what preparation is required before performing the evaluation (Daniel and Froude, 1998). Moreover, precise instructions concerning the administration of the assessment are sometimes missing (e.g., Stott et al., 1984). In addition, it is not always clear whether the evaluator is expected to practice using the tool before its administration. These factors may significantly affect a student's score. The ETCH-M is an example of an evaluation for which developers required high standards of evaluator preparation before actual administration (Diekma et al., 1997). Only in two of the tests that were described in this article was it specified that evaluators should practice on writing samples (Amundson, 1995; Reisman, 1991, 1993). Different studies have found that teachers who have experience in administering a particular test tend to have a more reliable judgment of writing and that the evaluator may respond, consciously or unconsciously, to different criteria (e.g., letter shape, size) (Feldt, 1962; Otto and Askov, 1962, Otto, Askov, and Cooper, 1967). According to Graham (1986a,b), the evaluator's familiarity with the purpose of the tool may affect the severity of the assessment. The combination of a lack of precise instructions together with a lack of practice

in using the tool raises doubts regarding the reliability of evaluators who are not part of the research team (Graham et al., 1989).

The Grading of Assessment Criteria

Discussions regarding which criteria constitute the critical components of handwriting readability, as well as how to measure them, are still ongoing (Bonny, 1992; Daniel and Froude, 1998; Graham, 1986a,b; Phelps et al., 1985; Reisman, 1993; Rubin and Henderson, 1982). Although there is great variability in the definition of “readability,” most researchers accept the criteria of size (height, width), slant, spacing (spaces between letters/words), the degree of line straightness, shape (letter form and shape), and the general merit of the writing (Bruinsma and Nieuwenhuis, 1991; Mojet, 1989). However, the grading scales for each criterion are different from one assessment to another. Ambiguous grading of criteria may result in a lack of scale reliability. Furthermore, the importance of each criterion and which combinations of criteria best produce readable handwriting remain unclear (Graham, 1986a,b). For example, letter formation, defined by Formsma (1988) as the way in which letters are made, is one of the criteria considered in most evaluations (e.g., Alston, 1983; Rubin and Henderson, 1982; Ziviani and Elkins, 1984). Yet, different scales evaluate this criterion differently. In the BHK scale, the evaluator needs only to check “yes” or “no” in response to questions relating to general letter forms (i.e., collisions of letters, inconsistent letter size, or correction of letter forms). In Rubin and Henderson’s scale, letter formation is judged on a 4-point scale (Rubin and Henderson, 1982). In comparison, Ziviani and Elkins (1984) give exact specifications to determine the accuracy of each letter. Graham, Weintraub, and Berninger (2001) reinforced the importance of measuring letter legibility. They found that letter legibility made a significant contribution to the prediction of text legibility after all other predictor variables were controlled for. They also found that a small number of letters accounted for a large proportion of the overall legibility. Therefore, it seems that although authors unanimously agree that letter formation is a legitimate criterion relating to overall legibility, it is difficult to compare legibility data from different evaluations because most authors describe different ways of measuring it. Moreover, no data are available on the age-related performance of children regarding legibility criteria features (Hamstra-Bletz and Blote, 1990).

The Assignment

The existing scales do not consider the effect of the complexity of the handwriting assignments on test outcomes, nor do they specify the rationale

underlying their choice of handwriting assignment (Ziviani and Watson-Will, 1998). Various levels of task complexity are seen in the following examples of tasks that are used in different evaluations. Some tools give a variety of assignments, such as copying shapes, letters, and words (Ziviani and Elkins, 1984). Others ask the child to copy a paragraph (Erez et al., 1999; Hamstra-Bletz et al., 1987; Phelps et al., 1985; Rubin and Henderson, 1982), to write a paragraph according to dictation, or to write letters and/or numbers from memory (Amundson, 1995; Erez et al., 1999). In yet another approach, children are asked to write a 20-min essay about their favorite person (Alston, 1983). What was the rationale of the various test developers for making their choice of assignment task? Was the tool specifically constructed to test writing acquisition according to the developmental sequence, to fit the assignment to ones required in school, or conform to some other rationale? For example, test developers who specified the assignment of copying a passage containing 57 words (Rubin and Henderson, 1982) did not provide an explanation as to the choice of assigning exactly 57 words.

These issues are important in light of research indicating that the type of assignment affects the performance outcome. Researchers have found that people write differently when asked to copy than when asked to write creatively (Lewis, 1964). Moreover, writing can be affected by the individual meaning that the assignment holds for the writer (Graham, 1986a,b). Task parameters are important in clinical and educational settings where children observed to succeed in short-writing assignments (words, sentences) fail to complete longer assignments (paragraphs) or succeed in copying tasks but not in dictation tasks (Levine, 1993).

The Instructions

Different handwriting scales specify the use of different types of instructions for completing the writing assignments. A child may perform differently when asked to “write as quickly as possible without stopping for corrections” (Ziviani, 1984; Ziviani and Watson-Will, 1998), “write as you usually do when you try to write well” (Reisman, 1993), or “write as you are used to” (Erez et al., 1999; Phelps et al., 1985). It is possible that the nature of the instructions and even the way writers perceive them in view of past experience with writing (e.g., fear or frustration) may affect an individual’s handwriting performance.

The Writing Accessories/Format

In some assessments, the child is asked to write on unlined paper (Phelps et al., 1985; Rubin and Henderson, 1982) and in others, on lined paper

(Alston, 1983; Erez et al., 1999). According to Phelps and Stempel (1988) the rationale for using unlined paper in the CHES-M is that this format enables younger students to arrange the letters and words in an unstructured manner in the space. The decision to use unlined paper was based on previous studies in which it had been found that younger children write more legibly on a page with no lines (Hackney et al., 1973; Lindsay and McLennan, 1983). However, contradictory results can also be found. Burnhill et al. (1983) and Krzesni (1971) showed that the quality of writing on unlined paper was inferior to that done on lined paper. Moreover, young schoolchildren are used to writing on lined paper.

The effect that writing on different types of paper has on handwriting was also the subject of a study by Trap-Porter et al. (1983), who found that not only does the presence or absence of lines affect handwriting quality but the width of the line affects quality as well. These researchers found that letters written with a wider line (1.11 cm) were more accurate than those written on paper with ordinary line width (0.5 cm). It is possible that the different findings regarding paper format are the result of the children's ages or the specific assignments given to them. Regardless of the reason, most of the assessments lack a description of the rationale for giving a page with or without lines, and no consideration is given to the way in which the child is used to writing at school.

Another factor that is not specified in many handwriting assessments is which writing tool (pen or pencil; the one the writer is used to or another one) should be used for the handwriting assignment. Most assessments that specify the writing tool request that pencils be used. In some assessments, sharpened pencils are required (Diekma et al., 1997; Phelps et al., 1985), whereas in others children are permitted to use their personal pencils (Alston, 1983). Most of the assessments do not specify whether the child is allowed to use an eraser. Diekma et al. (1997) indicated that the pencils given to the children should be equipped with erasers. However, this approach is not universal. Erez et al. (1999) emphasized that the children should not have an eraser during the assessment. Common sense dictates that the presence or absence of an eraser during the assessment may affect the amount of time taken to perform handwriting tasks as well as the degree of readability of the written product.

Variability in Personal Writing Style

Individuals tend to develop their own personal writing styles. In fact, a person's handwriting may change from one day to the next or even within the same written passage (Herrick, 1960). Assessment scales that are not

sensitive enough to personal or developmental changes in the individual's handwriting may result in children being falsely judged as having handwriting difficulties. Yet, none of the authors of the scales described in this article discussed whether they had considered this issue or whether they had found a resolution to it in the scale they developed.

Consideration of Performance Factors

Most of the assessment scales measure handwriting quality according to set criteria that were defined by their developers. Unfortunately, only two scales (Erez et al., 1999; Stott et al., 1984) are designed to alert examiners to behaviors commonly observed among poor writers, such as stress, fatigue, or the tendency to take frequent breaks while writing. The observation and subsequent documentation of such behaviors occurring as the child is writing may give the evaluator additional important information regarding the child's handwriting process.

Practicality of Administration

In order for the scale to be useful either in the educational system or for research purposes, it should require as little time as possible for administration and scoring. However, researchers' desire to produce a scale that was both brief enough to be efficient and objective enough to satisfy reliability requirements posed a continuing dilemma. The first approach, attempted by Ayres (1912) and Thorndike (1910), was to produce a handwriting scale that could be quickly scored by obtaining an overall impression of handwriting quality. This approach proved subjective and lacked sensitivity (e.g., Ayres, 1912; Thorndike, 1910). A more refined approach to scoring was that used in a handwriting tool in which graded writing samples were provided and compared with written samples submitted by students (Herrick and Elebacher, 1963). However, this method proved impractical for use in schools (Herrick and Elebacher, 1963). Recently, Erez et al. (1999) developed an evaluation in which a calibrated measuring device is used to obtain precise measurements of spatially related data—a relatively time-consuming process. In fact, no satisfactory solution has been found as yet to satisfy the demand for both efficiency and accuracy in the collection of quantitative handwriting data. As a result, the evaluation of handwriting in schools is most commonly accomplished through teacher observations (Stowitschek et al., 1987). Perhaps, the lack of a reliable and practical handwriting assessment largely explains the prevalence of students with handwriting difficulties and why schools have made little progress evaluating handwriting.

Reliability and Validity

Many of the existing handwriting tools are limited in research on psychometric properties. Some of the scales that compare of a sample of a student's handwriting to previously graded samples do not provide samples representative of the entire range of handwriting performance, nor do they include samples representative of the worst and of the best writers. Moreover, most tests make no reference to different subgroups of writers (e.g., boys vs. girls), even though this may affect the assessment.

For those scales whose interrater reliability or test-retest reliability have been checked, results indicate a wide range of values (0.44–0.98) (Alston, 1983; Hamstra-Bletz and Blote, 1990; Phelps et al., 1985; Ziviani and Elkins, 1984) indicating that not all scales are sufficiently reliable. Moreover, scales with reliability coefficients around 0.44 and lower should not be used for decisions regarding a student's handwriting status (i.e., deficient or adequate).

The situation regarding the proven validity of scales is even less satisfactory (see Table I). For many scales, no validity studies were conducted at all. Some of the scales were used only with typical children and not with those with difficulties—for whom the tools were actually intended. Research studies raise questions regarding the validity of the writing assessment tools and stress the need for development of a more reliable and valid assessment tools (Daniel and Froude, 1998).

Applicability of the Scales to Different Populations

A review of the literature demonstrates that a number of handwriting scales were developed but not further researched in terms of applicability (e.g., Alston, 1983; Helwing et al., 1976; Phelps et al., 1985; Rubin and Henderson, 1982; Stott et al., 1984; Ziviani and Elkins, 1984). The two scales receiving the most thorough investigation are the TOLH (Larsen and Hammill, 1989) and the BHK (Hamstra-Bletz et al., 1987). Results of studies performed using the TOLH show that the tool is sensitive enough to assess the handwriting of children with difficulties, test the relatedness between writing style and readability, and identify children at risk (Graham and Weintraub, 1996; Simner, 1996). Extensive research into the psychometric properties of the BHK (Hamstra-Bletz et al., 1987; Hamstra-Bletz and Blote, 1990, 1993) indicates that it is sensitive enough for assessing developmental changes, identifying dysgraphic children, identifying the secondary effects of chemotherapy, and determining the efficiency of physiotherapy treatment (Blote and Hamstra-Bletz, 1991; Hamstra-Bletz and Blote, 1990, 1993; Reinders-Messelink et al., 1996; Smits-Engelsman et al., 1996).

Student Involvement

An underemphasized element in many methods of handwriting assessment is the writer's own handwriting evaluation. Bruinsma and Nieuwenhuis (1991) wrote that the writer's responsibility for his/her own written product must be stressed. The researchers recommend that students be taught to evaluate their own handwriting even at the initial stages of writing acquisition. They suggest that self-assessment would encourage students to improve their handwriting and to become aware of changes. Studies indicate that students can be effective evaluators of their own handwriting (Stowitschek et al., 1987). Children can be taught how to successfully document, evaluate, and correct their own handwriting. However, in order for students to assess their own handwriting, defined objective criteria and scales are needed (Moxley et al., 1990). Unfortunately, the possibility of self-assessment is not mentioned in any of the writing assessment scales described in this review. It is not clear if this is due to the researchers' being unaware of the educational value afforded by self-assessment or to their belief that evaluation requires professional experience. Bruinsma and Nieuwenhuis (1991) performed a study examining self-assessment. They asked students to evaluate their handwriting according to five defined criteria: slant, size, space, shape, and general look. Many of the students (51%) were not satisfied with different aspects of their handwriting, although it was globally legible according to the handwriting evaluation score. The researchers emphasized the importance of the writer's self-awareness of his/her handwriting as a basis for quality improvement. Self-evaluation is an issue that warrants further development and research.

Writing Speed Measurement

Functional writing should be both legible and performed in a reasonable amount of time. Therefore, many handwriting evaluations include a segment that determines writing speed. However, different writing tools vary widely in how writing speed is measured. Typically, speed is calculated either by recording the amount of time required to write a specific text or the amount of text reproduced within a specific time period. Some evaluations test speed on the basis of the number of letters written in a minute (Rubin and Henderson, 1982), others in 2 min (Phelps et al., 1985; Ziviani and Elkins, 1984), and still others in 5 min (Hamstra-Bletz et al., 1987). Furthermore, handwriting speed studies done in different countries yielded inconsistent results (Table II). This disparity may be attributed to the different approaches used to teach writing internationally and differences in letter forms.

Alternatively, the inconsistency in writing speed measurements may also reflect variable in methodological factors such as those detailed in this article (i.e., differences in the type and duration of writing assignments, the time needed to perform the writing task, the writing accessories used, the evaluator or the instructions given to children on how to perform). What follows are examples that demonstrate different approaches used to measure handwriting speed in different scales. Ayres (1912) asked children to copy a passage until it became familiar to them and then to write it from memory. Groff (1961), in an attempt to have the assessment procedure resemble handwriting demands in real life, had children read a passage until they were familiar with all the words and then write it. In contrast, Ziviani and Elkins (1984) asked children to copy the phrase “cats and dogs” as quickly as possible on lined paper for 2 min. In contrast, Phelps et al. (1985) asked children to copy a passage on unlined paper, at their own usual pace for 2 min. In a longitudinal study performed with children in Germany, children were requested to copy a sentence at their usual pace for 5 min (Hamstra-Bletz and Blote, 1990). In contrast, Wallen et al. (1996) asked children to copy a sentence (The quick brown fox jumps over the lazy dog) “as quickly as you can, but as organized as you can” on a lined page for 3 min. In an assessment tool developed for the Hebrew language (Erez et al., 1999), children were asked to copy a 30-word passage and to write a passage dictated to them. Finally, Sasson et al. (1986) gave children two different sets of instructions: first, to write at their usual writing pace (“U” in Table II) and then, to write as fast as they can (“R” in Table II).

In light of the variable methods, it is not possible to judge how handwriting speed relates to age. In fact, the only clear tendency observed from the results is that performance speed increases with age.

Another weakness of the rate studies is that subjects’ gender is not reported even though writing speed between boys and girls differs (Groff, 1963; Ziviani and Elkins, 1984). Some studies (Cohen, 1997; Ziviani, 1996) indicate that boys write faster, whereas others (Dutton, 1990; Ziviani, 1984) indicate that girls write faster. Results of yet another study found that between the ages of 7 and 10 years, girls write faster and that at age 11, boys write faster (Ziviani and Watson-Will, 1998). Despite these inconsistencies, gender appears to impact writing speed.

Summary of Methodological Issues

The methodological variability of the studies reviewed limits comparisons of results and precludes the development of an information database about handwriting characteristics, such as legibility and speed of children at

various ages. Such a database would help researchers and educators assess and remediate handwriting difficulties.

It is important to note that the scales developed thus far relate to the written *output* and not to the *process* of handwriting performance, although the latter may yield valuable information about the characteristics of the writer's handwriting. For the purpose of this article, writing output refers to the handwriting product that can be analyzed by analytic and global-type evaluation scales. In contrast, the examination of the handwriting process refers to the computerized measurement and analysis of a different set of variables, such as time, space, and pressure while the child is actually performing a writing task. Research concerning the features that distinguish the handwriting processes of children in general, and of children with difficulties in particular, is still in the early stages. In the following section, a brief review of studies designed to analyze the handwriting processes of children with handwriting difficulties is presented along with a discussion of new avenues of research initiated as a result of technological developments.

FUTURE DIRECTIONS: INVESTIGATING THE HANDWRITING PROCESS

Handwriting is a multilevel process consisting of numerous, concerted, cognitive, and motor actions of which the final, static outcome, cursive script, is the result (Van Galen, 1991; Van Galen and Morasso, 1998). In general, studies of dysgraphia have been conducted from a descriptive, product-oriented approach, and the application of the results of handwriting evaluation, notably in educational settings, is often based on the static final product without reference to the underlying source of the difficulty (Bruinsma and Nieuwenhuis, 1991; Hamstra-Bletz and Blote, 1993; Longstaff and Heath, 1997). The development of computerized technology over the last 20 years has made it possible for researchers to examine handwriting in a whole new light, enabling the quantitative measurement of the handwriting process instead of relying solely on the assessment of the written product. This switch in orientation is appropriate because handwriting is a highly dynamic process (Longstaff and Heath, 1997).

Analysis of the handwriting process is accomplished through the use of a digitizing tablet, an electronic surface which, when used in tandem with a special pen and computer, allows for the recording of the "x" and "y" coordinates of the pen on the paper. Such recordings reveal the spatial and temporal features of handwriting in real time. Sensors located in the pen record the pressure used by the writer while writing. As with any skilled performance, the production of legible handwriting requires movement patterns that can

be reproduced with little variability in time and space (Longstaff and Heath, 1987). Because a skilled movement is characterized by precise organization in time and space and by appropriate force regulation, documentation of the spatial, temporal, and pressure measures while writing supplies important information about the degree of handwriting proficiency.

Applications of technologically driven research into handwriting began with Teulings and Thomassen's study outlining advanced techniques for recording handwriting (Teulings and Thomassen, 1979). Recent research has emphasized the handwriting process of adults (e.g., Alimi and Plamondon, 1996; Rogers and Found, 1996; Teulings, 2001; Van Galen et al., 2001). In contrast, less research exists in relation to the handwriting process of children in general and of children with difficulties in particular (Wann and Kadirkamanathan, 1991). Until recently, insufficient attention was given to the diagnostic potential of computer analysis in identifying the handwriting characteristics of children with handwriting difficulties.

Computerized research on the analysis of the handwriting of children with difficulties first arose from the investigation of handwriting as a protocol for research on basic aspects of motor control. The researchers using this approach reasoned that skilled handwriting requires a high level of control in both time and space (Graham and Weintraub, 1996; Wann, 1987). As a result of the work of these motor control researchers, it became apparent that the handwriting process of children with difficulties differed with respect to the process of typical children in terms of specific spatial and temporal characteristics. For example, Wann (1987), who based his study on Hay's three movement categories—ballistic, step, and ramp—indicating levels of movement control/maturation (Hay, 1979), found that children with handwriting difficulties tend to use less mature movement patterns and rely less on visual feedback. Furthermore, their ability to regulate force is decreased.

Van Galen et al. (1993) also found evidence of immature movement control among poor handwriters, which the authors expressed in terms of "movement noise" or "neuromotor noise" in referring to the children's lack of movement precision and consistency. These researchers suggested that this so-called neuromotor noise is a dynamic influence on the spatial variability of movement (Smits-Engelsman, Van Galen, and Shoemaker, 1998). Results showed that poor writers got higher absolute scores of "neuromotor noise" than did typical writers. Poor writers used faster movements and larger movement beats (Van Galen et al., 1993). Moreover, poor handwriters were less successful in adapting the level of noise to the variable spatial accuracy demands of the tasks. The findings of Van Galen et al. (1993) are consistent with those of Smits-Engelsman et al. (1995), who found that poor handwriters fail to make fine adaptations to the spatial demands in motor tasks. This may be due, in part, to deficiencies of the muscular initiation

process or to difficulties in muscular initiation. Poor handwriters are less effective in the management of the natural neuromotor noise in their motor control system (Smits-Engelsman et al., 1998). Regardless of the underlying cause, the research of Van Galen et al. (1993) and that of Smits-Engelsman et al. (1995) provided the scientific community with objective evidence that the handwriting process of children with poor handwriting differs from that of typical children and can be characterized on the basis of its lack of precision in time and space.

The spatial characteristics distinguishing the handwriting of children with handwriting problems from that of typical children were investigated in a study performed by Smits-Engelsman et al. (1994a). This study's aim was to relate handwriting problems to one of the three psychomotor processing stages of the Van Galen writing model (for more details, see Smits-Engelsman et al., 1998; Van Galen, 1991; Van Galen et al., 1993). Results of this study were unexpected; there was no evidence that poor handwriters had specific problems in either letter form retrieval or size control. Further, there were no signs that handwriting performance problems were related to cognitive ability. Instead, findings indicated that the primary distinguishing factor among poor writers was the increased number of spatial errors that they made and their apparent failure to make accommodations for the spatial accuracy constraints of the experimental handwriting tasks. The contribution of this study cannot be overstated; it implies that poor movement control among this group of poor handwriters was largely the result of deficiencies in spatial accuracy.

To further investigate performance proficiency for different task parameters, Smits-Engelsman et al. (1994b) conducted another study 1 year later on eight poor and eight proficient handwriters from the previous study's sample groups. The participants viewed a series of letters on a computer monitor and copied them individually onto a digitizing tablet. This stimulus was varied according to the type of letters (i.e., garlands or arcades), letter sequence, and the size of the space in which the children wrote. The results showed that improvements in the handwriting of poor handwriters lagged behind that of the proficient handwriters. Specifically, poor writers made relatively more overshoots (lines extending beyond available space) in the larger letter condition and more undershoots (lines not reaching the limit of available space) in the smaller letter condition. The researchers found that the proficient writers showed fairly consistent improvement, whereas the poor writers did not catch up, nor did their performance improve much in that 1-year period (Smits-Engelsman et al., 1998). These results confirm that the one of the most important characteristics of poor handwriting is the difficulty that children have in controlling spatial accuracy. However, this study also showed that the spatial inaccuracy

problems that contribute to the poor handwriting of children persisted over time.

The researchers next conducted a longitudinal study to confirm their results regarding the persistence of such problems over time (Smits-Engelsman and Van Galen, 1997). They selected 48 pupils on the basis of their handwriting proficiency (24 poor writers and 24 proficient writers) from Grades 2 to 4 (mean age = 9.1 years) in 10 different elementary schools spread throughout the Netherlands. The children were asked to write letter strings of varying complexity related to motor control demands (shape, size, and accuracy). A representative sample of 16 participants (8 poor writers and 8 proficient writers) was retested again after a 1-year period. Outcome measures included the number of overshoots and undershoots, total movement/writing time, writing dysfluencies, stroke curvature, and “neuromotor noise.” Results indicated that poor writers had less stroke curvature and significantly more overshoots, undershoots, and neuromotor noise than did proficient writers. Handwriting size was also significantly different, with the proficient group writing 10% smaller than the poor writers. In summary, the findings of this study provide support for the view that poor psychomotor skill, including handwriting performance, persists in children over time.

Smits-Engelsman and Van Galen’s finding that movement time and writing dysfluencies do not distinguish poor and proficient writers (Smits-Engelsman, 1997) contradicts the results of other studies that examined the temporal aspects of the handwriting process (Shoemaker et al., 1994; Shoemaker and Smits-Engelsman, 1997). For example, in one study, the spatial accuracy of clumsy children who had poor handwriting was compared with that of a control group on a task requiring them to copy figures of different levels of complexity (Shoemaker et al., 1994). Although the researchers were not surprised to find that clumsy children showed a higher incidence of spatial undershoots and overshoots, no differential speed/accuracy trade-off was found. Despite the fact that no differences were found relating to handwriting speed, the movement patterns of the clumsy children were characterized by greater number of dysfluencies and longer pause durations. The researchers concluded that the clumsy children used a different movement strategy from that of the other children and suggested that these pauses were used for additional programming of movements (Smits-Engelsman and Van Galen, 1997). In other words, perhaps clumsy children only manage to process the global aspects of a task during the reaction time interval and postpone further programming until the execution phase. It is this serial processing strategy that appears to disrupt the fluency of movement.

Shoemaker and Smits-Engelsman (1997) compared dysgraphic children who had no gross motor problems to dysgraphic children who had generalized motor problems. They used a handwriting evaluation (BHK—Concise

Evaluation Scale for Children's Handwriting) to measure the quality and speed of handwriting and used a drawing task performed on a digitizer to measure handwriting kinematics. The researchers found that both groups of dysgraphic children drew more slowly and with elongated pause intervals between strokes in comparison with the control group.

As a result of their investigations of the handwriting process in poor vs. proficient handwriters, Smits-Engelsman et al. (2001) hypothesized that poor handwriting is part of a wider neuromotor condition characterized by faster and cruder movements, lack of inhibition of comovements, and poor coordination of fine motor skills. To test their theory, they collected digitizer-based kinematics measures of drawing movements of poor writers in the fourth and fifth grades using the flower-trail drawing item of the M-ABC test as the research task (Henderson and Sugden, 1992). They found effects for the groups regarding their movement time (time needed to complete the figure) and movement velocity while drawing. The poor writers finished the drawing task in less time and also used a higher movement velocity than did proficient writers. Proficient writers tend to spend, on average, more time pausing above the paper with the digitizer pen. Differences were not found for pen pressure nor for the number of times the pen was raised.

The Use of Computerized Temporal and Spatial Measurements in Evaluating Handwriting Performance

The possibility of reliably evaluating handwriting through the use of spatial and temporal measurement of the handwriting process was examined in a study performed on adults who were poor handwriters (Longstaff and Heath, 1997). These investigators employed the temporally sensitive techniques of the digitizer to investigate the relationship between spatial (i.e., legibility) and kinematics (i.e., dynamic) aspects of handwriting production of poor adult handwriters. Participants were asked to write a pseudo-word 10 times on the digitizer. The spatiotemporal variables were analyzed both between trials and within subjects using coherence analysis. Participants previously rated as proficient handwriters by three independent judges displayed a greater degree of temporal consistency than did the less proficient writers. Thus, these results, indicating that spatial inconsistencies are related to dynamic variability, also suggest the possibility that the methodology employed is useful as a tool for the quantitative assessment of handwriting quality.

The diagnostic potential of computer analysis in identifying children with handwriting difficulties has also been studied. Wann and Kadiramanathan (1991) selected 16 children with handwriting difficulties

and 16 children who write well according to the scores that they received in the writing assessment developed by Rubin and Henderson (1982). The study protocol required that the children write separate letters (w, a) and a sequence of letters on a digitizer. Wann and Kadiramanathan found that the handwriting of the children with difficulties was characterized by a lack of continuity when writing a sequence of letters and by variability in the orientation of the main characters of the letters when writing each letter separately. Similar results were also found in a study conducted by Mojet (1991) on children in Grades 3–6 in Germany.

A comparison of the writing process parameters of children with and without handwriting difficulties and children diagnosed as having dyslexia was performed by Sovik, Arntzen, and Thygesen (1987a,b). The children were chosen on the basis of their reading and spelling abilities. Seventy percent of them were boys. Different kinds of graded tasks were given to the children to gain a deeper understanding of the cognitive and psychomotor aspects of children's spelling and writing performance. The tasks included were brief and presented only different words (Sovik et al., 1987a) or letters (Sovik et al., 1987b). In both studies, the measures of handwriting performance were the total time spent on handwriting, the number of spelling errors, and accuracy. The accuracy variable was measured according to three parameters: the overall time of stops during the performance, the maximal absolute writing speed (millimeter per second) of each of the items, and the average height of the assignment's items. Results showed that when comparing "accuracy" scores, the time of performance, and the spelling errors of the three groups, the children without handwriting difficulties had the best scores and the dysgraphic children performed worse than the dyslexic children on all variables except for that of writing duration (Sovik et al., 1987a).

In the second study, Sovik et al. (1987b) found that the dyslexic children wrote more slowly than the other groups of children and their average number of writing mistakes was the highest. The dysgraphic children had the lowest accuracy score in writing and rhythm. The main characteristic of the third-grade dysgraphic children found in Sovik et al.'s studies is that their handwriting was less "smooth" than the writing of their peers (Sovik et al., 1987a,b). Consistent differences in the amount of time poor and proficient writers spent pausing as they write, though, was not obtained.

Wann and Jones (1986) compared the writing performance of Australian children who write well with that of children who have difficulties. They focused on aspects of time and space during movement. They asked children to copy a letter/word that was written at the beginning of every line. The children were given the opportunity to practice until they felt confident about the shape/size required in each assignment. Wann and Jones found that children with difficulties took intermissions (e.g., pauses) at greater frequencies

and for longer periods of time in comparison to their counterparts when writing *letters*. In contrast, they did not find that poor writers paused more often and for longer periods of time than did good writers when writing *letters* (Graham and Weintraub, 1996), as was the case in the second study by Sovik et al. (1987a,b).

Another interesting finding concerns the variability of handwriting speed during the performance of a writing task (Wann and Jones, 1986). Although the data suggest that the overall performance speed of children with handwriting difficulties does not differ significantly from that of proficient handwriters, Wann and Jones (1986) noticed that for individual children with handwriting difficulties, writing speed showed significantly more variability during writing performance than that for the controls. The researchers suggested that the degree of variability in handwriting speed and in the duration of intermissions (e.g., pauses) during handwriting performance are the best indicators of writing difficulties, even more so than the overall time of performance or the number of intermissions taken during it.

METHODOLOGICAL ISSUES RELATED TO COMPUTERIZED STUDIES OF HANDWRITING DIFFICULTIES

The computerized studies described earlier offer different ways of defining and measuring digitizer data. Instructions and assignments vary widely as do the definitions of criteria for measuring handwriting proficiency. In most studies, the researchers focused mainly on the writing of letters and words (Smits-Engelsman et al., 1994a; Smits-Engelsman and Van Galen, 1997; Sovik et al., 1987b; Van Galen et al., 1993; Wann, 1987; Wann and Jones, 1986; Wann and Kadiramanathan, 1991) sometimes at different levels of complexity. What is conspicuously absent is the use of a computerized system to investigate the handwriting process used for text lengths that approximate those typically required of children at school and at home. Sovik, Arntzen, Samuelstuen, and Heggberget (1994) note that in functional writing (text production), the properties of the writing tasks (the words) can be expected to affect the process as well as the product of the handwriting performance.

Another limitation of the computerized studies is the small sample sizes [for example, the 24 poor and 24 proficient writers studied by Smits-Engelsman and Van Galen (1997)] especially in comparison to the large samples used for studies done on handwriting product scales. Hence, generalizing results of process studies to the entire population of poor writers is problematic. Moreover, the contrasting results found between studies (e.g., Sovik et al., 1987a, in contrast to Sovik et al., 1987b, and Wann and Jones, 1986) may be due to the small sample sizes used.

In summary, the common aim of the computerized studies was to show that the differences between children with and without handwriting difficulties lie not only in their written products but also in the dynamics of their handwriting performance. According to these studies, the main temporal and spatial features that distinguish the handwriting process of poor writers from proficient writers include less mature movement patterns with “neuromotor noise,” various irregularities in movement control (Smits-Engelsman et al., 1994a, 1995; Van Galen et al., 1993; Wann, 1987), variability in writing time, pauses at greater frequencies and for longer periods of time, lack of continuity and fluency (Wann and Jones, 1986; Wann and Kadirkamanathan, 1991), failure to obey spatial constraints, and lack of consistency (Smits-Engelsman and Van Galen, 1997).

One conclusion ensuing from these studies is that variables such as speed or total handwriting performance time do not differentiate between poor and proficient writers. This conclusion conflicts with anecdotal evidence reported by most clinicians and teachers. In contrast, pauses and the temporal variability while writing are meaningful variables for differentiating between poor and good writers. These results lead to a number of questions: What is the influence of task length on the finding that total time or speed does not differentiate between handwriting groups? If children would be asked to perform longer tasks (such as they are requested to do in school), would the results be different?

From this brief review of the computerized studies on the handwriting process, it is apparent that the goal of describing the features of the writing process of children who have difficulties has been accomplished. Unfortunately, digitizer studies to date have neglected to take the research one logical and significant step further: relating the writing process features of children with handwriting difficulties, to the writing *products* that characterize these children.

CONCLUSION

Handwriting researchers over the years were faced with significant problems in attempting to identify poor writers in an objective and standardized manner and in differentiating between poor and proficient writers on the basis of distinguishing writing characteristics. Research in the twentieth century led to several important advancements in the area of evaluation of handwriting difficulties. As described in this article, two main directions of research predominated. The first direction includes the development and testing of evaluation scales dealing with global-holistic evaluations of readability and analytic evaluations that assess readability in relation to predetermined criteria. The second direction consists of computerized “on-line”

investigations of the handwriting process. Yet, despite these advances, educators, clinicians, and researchers continue to search for tools that provide greater insight into the motor, perceptual, and cognitive components underlying poor handwriting.

Each approach has its advantages and limitations. Subjective analyses of the handwriting product, via both global and analytic methods, are readily available, inexpensive, and technically simple to implement in environmentally friendly settings such as in the child's classroom. Moreover, the human mind, because of its ability to detect the "gestalt" of complex images, enables human evaluators to attain a global impression of a writing sample's readability. This ability is used routinely by teachers and handwriting evaluators, but has not become successfully automated.

Nevertheless, subjective handwriting evaluations suffer from limited accuracy, sensitivity, and reliability. In contrast, objective digitizer-based analyses enable the documentation of handwriting dynamics, providing data beyond that which is observable to the human eye. Computer-based analyses are more accurate, sensitive, and reliable than the subjective analyses and much of the procedure is rapid and automated. However, the equipment and software is considerably more expensive than are traditional handwriting evaluation scales. Moreover, the tester must make a serious effort to ensure that the instruments are organized and presented in a way that does not encumber the child and disturb his or her ability to write in a natural manner. Finally, as indicated earlier, practical applications are still limited. For example, a computer is not yet capable of making a global decision about the legibility of the written product and even simple operations such as the identification of the start and stop of successive characters are not fully automated and thus still require human intervention.

A combined approach to handwriting evaluation, one that takes advantage of the strengths of both human and digitizer-based evaluation procedures, is possible with greater communication between their respective developers. Such a partnership could not only lead to a richer and broader pool of information than has been available in the past—but could conceivably stimulate the development of new ideas and improved approaches for both researchers and individuals committed to helping children with handwriting difficulties.

ACKNOWLEDGMENTS

We thank Sarina Goldstand, OT, for her dedicated help and professional editorial support. Financial support from the Ministry of Education is gratefully acknowledged.

REFERENCES

- Alimi, A. M., and Plamondon, R. (1996). A comparative study of speed-accuracy trade-off formulations. The case of spatially constrained movements where both distance and spatial precision are specified. In: Simner, N. L., Leedham, C. G., and Thomassen, W. M. (eds.), *Handwriting and Drawing Research*. IOS Press, Burke, pp. 127-142.
- Alston, J. (1983). A legibility index: Can handwriting be measured? *Educ. Rev.* **35**: 237-242.
- American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders* (4th Ed.), American Psychiatric Association, Washington, DC.
- Amundson, S. J. (1995). *Evaluation Tool of Children's Handwriting (ETCH) Examiner's Manual*, O.T. Kids, Homer, AK.
- Amundson, S. J. (1995). *Evaluation Tool of Children's Handwriting (ETCH)*, O.T. KIDS, Homer, AK. See www.alaska.net/~otkids/about_children_legibility.htm
- Anderson, P. L. (1983). *Denver Handwriting Analysis*, Academic Therapy, Novato, CA.
- Askov, E., Otto, W., and Askov, W. (1970). A decade of research in handwriting: Progress and prospect. *J. Educ. Res.* **64**: 100-111.
- Ayres, L. (1912). *A Scale for Measuring the Quality of Handwriting of School Children*, Russel Sage Foundation, New York.
- Benbow, M. (1995). Principles and practices of teaching handwriting. In: Henderson, A., and Pehoski, C. (eds.), *Hand Function in the Child*, Mosby, St. Louis, MN.
- Berninger, V. W. (1994). *Reading and Writing Acquisition: A Developmental Neuropsychological Perspective*, Brown & Benchmark, Dubuque, IA.
- Berninger, V., and Graham, S. (1998). Language by hand: A synthesis of a decade of research on handwriting. *Handwriting Rev.* **12**: 11-25.
- Berninger, V., Mizokawa, D., and Bragg, R. (1991). Theory-based diagnosis and remediation of writing. *J. Sch. Psychol.* **29**: 57-59.
- Berninger, V. W., Vaughan, K. B., Abbott, R. D., Abbott, S. P., Woodruff Rogan, Brooks, A., Reed, E., and Graham, S. (1997). Treatment of handwriting problems in beginning writers: Transfer from handwriting to composition. *J. Educ. Psychol.* **89**: 652-666.
- Bezzi, R. (1962). A standardized manuscript scale for grades 1, 2, and 3. *J. Educ. Res.* **55**: 339-340.
- Blote, A., and Hamstra-Bletz, L. (1991). A longitudinal study on the structure of handwriting. *Percept. Mot. Skills* **72**: 983-994.
- Bonny, A. M. (1992). Understanding and assessing handwriting difficulties: Perspective from the literature. *Aust. Occup. Ther. J.* **39**: 7-15.
- Brigance, A. H. (1983). *Comprehensive Inventory of Basic Skills*, Curriculum Press, North Billarica, MA.
- Briggs, D. (1980). A study of the influence of handwriting upon grades using examination scripts. *Educ. Rev.* **32**: 185-193.
- Bruinsma, C., and Nieuwenhuis, C. (1991). A new method for the evaluation of handwriting material. In: Wann, J., Wing, A. M., and Sovik, N. (eds.), *Development of Graphic Skills*, Academic Press, New York, pp. 41-51.
- Burnhill, P., Hartly, J., and Lindsay, D. (1983). Line paper illegibility and creativity. In: Hartley, J. (ed.), *The Psychology of Written Communication*, Kogan, London, England, pp. 82-91.
- Chase, C. (1986). Essay test scoring: Interaction of relevant variables. *J. Educ. Meas.* **23**: 33-41.
- Chu, S. (1997). Occupational therapy for children with handwriting difficulties: A framework for evaluation and treatment. *Br. J. Occup. Ther.* **60**: 514-520.
- Cohen, M. R. (1997). Individual and sex differences in speed of handwriting among high school students. *Percept. Mot. Skills* **84**: 1428-1430.
- Collins, F., Baer, G., Walls, N., and Jackson, M. S. (1980). The development of a behavioral assessment technique for evaluation gradual change in handwriting performance. *Behav. Assess.* **2**: 369-387.

- Cornhill, H., and Case-Smith, J. (1996). Factors that relate to good and poor handwriting. *Am. J. Occup. Ther.* **50**: 732–739.
- Daniel, M. E., and Froude, E. (1998). Reliability of occupational therapist and teacher evaluations of handwriting quality of grade 5 and 6 primary school children. *Aust. Occup. Ther. J.* **45**: 48–58.
- Diekma, S. M., Deitz, J., and Amundson, S. J. (1997). Test-retest reliability of the evaluation tool of children's handwriting manuscript. *Am. J. Occup. Ther.* **52**: 248–258.
- Dutton, K. P. (1990). Writing under exam conditions: Establishing a baseline. *Profession. Dev. Initiativ.* 1989–90: 189–210.
- Dvash, L. E., Levi, M., Traub, R., and Shapiro, M. (1995). *Reliability and Validity of the Hebrew Handwriting Evaluation*, Unpublished manuscript, Faculty of Medicine, School of Occupational Therapy, Hebrew University, Jerusalem, Israel.
- Erez, N., Yochman, A., and Parush, S. (1996). *The Hebrew Handwriting Evaluation*, Faculty of Medicine, School of Occupational Therapy, Hebrew University of Jerusalem, Israel.
- Erez, N., Yochman, A., and Parush, S. (1999). *The Hebrew Handwriting Evaluation* (2nd Ed.), Faculty of Medicine, School of Occupational Therapy, Hebrew University of Jerusalem, Israel.
- Feldt, L. S. (1962). The reliability of measures of handwriting quality. *J. Educ. Psychol.* **53**: 288–299.
- Formis, R. (1988). *Literature on Evaluation of Handwriting*, Indiana University at South Bend, Indiana.
- Freeman, F. N. (1959). A new handwriting scale. *Elem. Sch. J.* **59**: 218–221.
- Graham, S. (1986a). A review of handwriting scales and factors that contribute to variability in handwriting scores. *J. Sch. Psychol.* **24**: 63–72.
- Graham, S. (1986b). The reliability, validity, and utility of three handwriting measurement procedures. *J. Educ. Res.* **79**: 373–380.
- Graham, S. (1990). The role of production factors in learning disabled students' compositions. *J. Educ. Psychol.* **82**: 781–791.
- Graham, S. (1992). Issues in handwriting instruction. *Focus Except. Child.* **25**: 1–14.
- Graham, S., and Weintraub, N. (1996). A review of handwriting research: Progress and prospects from 1980–1994. *Educ. Psychol. Rev.* **8**: 7–86.
- Graham, S., Boyer-Schick, K., and Tippetts, E. (1989). The validity of handwriting scale from the test of written language. *J. Educ. Res.* **82**: 166–171.
- Graham, S., Berninger, V. W., and Weintraub, N. (1998). The relationship between handwriting style and speed and legibility. *J. Educ. Res.* **5**: 290–296.
- Graham, S., Harris, K. R., and Fink, B. (2000). Is handwriting causally related to learning to write? Treatment of handwriting problems in beginning writers. *J. Educ. Psychol.* **4**: 620–633.
- Graham, S., Weintraub, N., and Berninger, V. (2001). Which manuscript letters do primary grade children write legibly. *J. Educ. Psychol.* **3**: 488–497.
- Groff, P. G. (1961). New spectrum of handwriting. *Elem. Engl.* **38**: 564–565.
- Groff, P. G. (1963). Who writes faster? *Education* **83**: 367–369.
- Hackney, C. S., Myers, E. M., and Zaner-Bloser, P. (1973). *"Sound of!" Teacher's Edition. Expressional Growth Through Handwriting*, Zaner-Bloser, Ohio.
- Hamstra-Bletz, L., and Blote, A. (1990). Development of handwriting in primary school: A longitudinal study. *Percept. Mot. Skills* **70**: 759–770.
- Hamstra-Bletz, L., and Blote, A. (1993). A longitudinal study on dysgraphic handwriting in primary school. *J. Learn. Disabil.* **26**: 689–699.
- Hamstra-Bletz, L., DeBie, J., and Den Brinker, B. (1987). *Concise Evaluation Scale for children's handwriting*, Lisse, Swets & Zeitlinger, Germany.
- Harvey, C., and Henderson, S. (1997). Children's handwriting in the first three years of school: Consistency over time and its relationship to academic achievement. *Handwriting Rev.* **11**: 8–25.
- Hay, L. (1979). Spatial-temporal analysis of movements in children: Motor programs versus feedback in the development of reaching. *J. Mot. Behav.* **11**: 189–200.

- Helwing, J. J., Johns, J. C., Norman, J. E., and Cooper, J. O. (1976). The measurement of manuscript letter strokes. *J. Appl. Behav. Anal.* **9**: 231–236.
- Henderson, S. E., and Sugden, D. A. (1992). *Movement Assessment Battery for Children: A Manual*, The Psychological Corporation, Harcourt Brace Jovanovic, Sidcup, Kent, OH.
- Henderson, S. E., and Sugden, D. A. (1992). *Movement Assessment Battery for Children; Manual*, The Psychological Corporation, A Harcourt Assessment Company, London.
- Herrick, V. E. (1960). Handwriting and children's writing. *Elem. Engl.* **37**: 248–258.
- Herrick, V. E., and Elebacher, A. (1963). The evaluation of legibility in handwriting. In: Herrick, V. E. (ed.), *New Horizons of Research in Handwriting*, University of Wisconsin Press, Madison.
- Hooper, S., Montgomery, J., Swarts, C., Reed, M., Brown, T., Levine, M., and Wasileski, T. (1993). Prevalence of writing problems across three middle school samples. *School Psychol. Rev.* **2**: 610–621.
- Hughes, D. C., Keeling, B., and Tuck, B. F. (1983). Effects of achievement expectations and handwriting quality on scoring essays. *J. Educ. Meas.* **20**: 65–70.
- Jones, D., and Christensen, C. A. (1999). Relationship between automaticity in handwriting and student's ability to generate written text. *J. Educ. Psychol.* **91**: 44–49.
- Jones, J., Trap, J., and Cooper, J. (1977). Technical report; Student's self recording of manuscript letter strokes. *J. Appl. Behav. Anal.* **10**: 509–514.
- Kaminsky, S., and Powers, R. (1981). Remediation of handwriting difficulties, a practical approach. *Acad. Ther.* **17**: 19–25.
- Krupa, M. (1991). The Effects of Two Writing Aids on the Handwriting of Two Children with Learning Disabilities. Thesis, University of Puget Sound, Tacoma, WA.
- Krzesni, J. S. (1971). Effect of different writing tools and paper on performance of the third grader. *Elem. Engl.* **48**: 821–824.
- Larsen, S. C., and Hammill, D. P. (1989). *Test of Legible Handwriting*, Pro-Ed., Austin, TX.
- Laszlo, J. I. (1990). Child perceptuo-motor development: Normal and abnormal development of skilled behaviour. In: Hauer, C. A. (ed.), *Developmental Psychology: Cognitive, Perceptuo-Motor and Neuropsychological Perspective*, North Holland, Amsterdam.
- Laszlo, J. I., Birstow, P. J., and Bartip, J. (1988). A new approach to treatment of perceptuo-motor dysfunction: Previously called clumsiness. *Support Learn.* **3**: 35–40.
- Laszlo, J. I., and Broderick, P. (1991). Drawing and handwriting difficulties: Reasons for and remediation of dysfunction. In: Wann, J., Wing, A. M., and Sovik, N. (eds.), *Development of Graphic Skills*, Academic Press, London, pp. 259–280.
- Lewis, E. (1964). An Analysis of Children's Manuscript Handwriting. Unpublished doctoral dissertation, University of California, Berkeley.
- Levine, M. D. (1993). *Developmental Variation and Learning Disorders*, Educators Pub., Cambridge, MA, pp. 308–345.
- Lifshietz, N., and Parush, S. (1996). *Screening Assessment for Children with Handwriting Deficits*, Unpublished manuscript, Jerusalem: School of Occupational Therapy, The Hebrew University.
- Lilly, L. (1987). The Effects of Service Delivery Models on Children's Handwriting. Thesis, University of Illinois, Chicago.
- Lindsay, G. A., and McLennan, D. (1983). Lined paper; It's effects on the legibility and creativity of young children's writing. *Br. J. Educ. Psychol.* **53**: 364–368.
- Longstaff, M. G., and Heath, R. A. (1997). Space-time invariance in adult handwriting. *Acta Psychol.* **97**: 201–214.
- Maeland, A. F., and Karlsdottir, R. (1991). Development of reading, spelling and writing skills from third to sixth grade in normal and dysgraphic school children. In: Wann, J., Wing, A. M., and Sovik, N. (eds.), *Development of Graphic Skills*, Academic Press, London, England, pp. 179–184.
- Marr, D., and Cermak, S. (2001). Consistency of handwriting development in the early elementary years: A literature review. *Isr. J. Occup. Ther.* **10**: E109–E129.
- McCutchen, D. (1996). A capacity theory of writing: Working memory in composition. *Educ. Psychol. Rev.* **8**: 299–325.

- McHale, K., and Cermak, S. A. (1992). Fine motor activities in elementary school: Preliminary findings and provisional implications for children with fine motor problems. *Am. J. Occup. Ther.* **46**: 898–903.
- Miller, L. T., Missiuna, C. A., Macnab, J. J., Malloy-Miller, T., and Polatajko, H. J. (2001). Clinical description of children with developmental coordination disorder. *Can. J. Occup. Ther.* **68**: 5–15.
- Modlinger, I. (1983). *Disorders of Written Language: Diagnosis and Remedial Teaching*, Sifriyat Hapoalim, Tel Aviv, Israel.
- Mojet, J. W. (1989). *Kemmerken van schrijfvaardigheid: Praespecten van het schrijven bitot twaalfjarigh* (Characteristics of the handwriting skill: Process aspects in 6–12 year olds), Academisch Boeken Centrum, De Lier: Academic Boeken Centrum.
- Mojet, J. W. (1991). Characteristics of the developing handwriting skill in elementary education. In: Wann, J., Wing, A. M., and Sovik, N. (eds.), *Development of Graphic Skills*, Academic Press, London, England, pp. 53–74.
- Moxley, R. A., et al. (1990). *Self Recording and Discussion in Young Children's Formative Self-Evaluation of Their Writing*, EDRS, New York.
- Otto, W., Askov, E., and Cooper, C. (1967). Legibility rating for handwriting samples: A pragmatic approach. *Percept. Mot. Skills* **25**: 638.
- Parush, S., Levanon-Erez, N., and Weintraub, N. (1998a). Ergonomic factors influencing handwriting performance. *Work* **11**: 295–305.
- Parush, S., Pindak, V., Han Markowitz, J., and Mazor-Karasenty, T. (1998b). Does fatigue influence children's handwriting performance. *Work* **11**: 307–313.
- Phelps, J., and Stempel, L. (1988). The children's handwriting evaluation scale for manuscript writing. *Read. Improv.* **25**: 247–255.
- Phelps, J., Stempel, L., and Speck, G. (1985). The children's handwriting scale: A new diagnostic tool. *J. Educ. Res.* **79**: 46–50.
- Reinders-Messelink, H., Schoemaker, M. M., Goeken, L., van den Briel, M., and Kamps, W. (1996). Handwriting and fine motor problems after treatment for acute lymphoblastic leukemia. In: Simner, M. L., Leedham, C. G., and Thomassen, A. J. W. M. (eds.), *Handwriting and Drawing Research*, IOS Press, Amsterdam, pp. 215–225.
- Reisman, J. E. (1991). Poor handwriting, who is referred? *Am. J. Occup. Ther.* **45**: 849–852.
- Reisman, J. E. (1993). Development and reliability of the research version of the Minnesota Handwriting Test. *Phys. Occup. Ther. Pediatr.* **13**: 41–55.
- Rogers, D., and Found, B. (1996). The objective measurement of spatial invariance in handwriting. In: Simner, M. L., Leedham, C. G., and Thomassen, A. J. W. M. (eds.), *Handwriting and Drawing Research*, IOS Press, Amsterdam, pp. 3–13.
- Rubin, N., and Henderson, S. E. (1982). Two sides of the same coin: Variation in teaching methods and failure to learn to write. *Spec. Educ. Forward Trends* **9**: 17–24.
- Sassoon, R. (1997). Dealing with adult handwriting problems. *Handwriting Rev.* **11**: 69–74.
- Sasson, R., Nimmo-Smith, I., and Wing, A. M. (1986). An analysis of children's penholds. In: Kao, H., Van Galen, G., and Hossain, R. (eds.), *Graphonomics: Contemporary Research in Handwriting*, Elsevier Science, North Holland, The Netherlands, pp. 93–106.
- Sasson, R., Nimmo-Smith, I., and Wing, A. M. (1986). An analysis of children's penholds. In: Kao, H., Van Galen, G., and Hossain, R. (eds.), *Graphonomics: Contemporary Research in Handwriting*, Elsevier, Amsterdam, pp. 93–106.
- Scardamalia, M., Bereiter, C., and Goleman, H. (1982). The role of production factors in writing ability. In: Nystrand, M. (ed.), *What Writers Know: The Language Process and Structure of Written Discourse*, Academic Press, New York, pp. 173–210.
- Shneck, C. M. (1998). Clinical interpretation of "test-retest" reliability of the evaluation tool of children's handwriting manuscript. *Am. J. Occup. Ther.* **52**: 256–258.
- Shoemaker, M. M., Shellekens, J. M. H., Kalverboer, A. F., and Kooistra, L. (1994). Pattern drawing by clumsy children: A problem of movement control. In: Simner, M. L., Hulstuijn, W., and Girouard, P. (eds.), *Contemporary Issues in the Forensic, Developmental, and Neurological Aspects of Handwriting* (Monograph of the Association of Forensic Document Examiners, Vol. 1). Association of Forensic Document Examiners, Toronto, pp. 43–64.

- Shoemaker, M. M., and Smits-Engelsman, B. C. M. (1997). Dysgraphic children with and without a generalized motor problem: Evidence for subtypes? In: Colla, A. M., Masulli, F., and Morasso, P. (eds.), *IGS 1997 Proceedings: Eight Biennial Conference*, The International Graphonomics Society, Nijmegen, The Netherlands, pp. 11–12.
- Simner, M. L. (1982). Printing errors in kindergarten and the prediction of academic performance. *J. Learn. Disabil.* **15**: 155–159.
- Simner, M. L. (1985). *Printing Performance School Readiness Test*, Guidance Centre, Faculty of Education, University of Toronto, Toronto.
- Simner, M. L. (1986). Further evidence on the relationship between form errors in preschool printing and early school achievement. In: Kao, H. S. R., Van Galen, G. P., and Hoosian, R. (eds.), *Graphonomics: Contemporary Research in Handwriting*, Elsevier, Amsterdam, pp. 107–120.
- Simner, M. L. (1986). Further evidence on the relationship between form errors in preschool printing and early school achievement. In: Kao, H. S. R., Van Galen, G. P., and Hoosian, R. (eds.), *Graphonomics: Contemporary Research in Handwriting*, North-Holland, Amsterdam.
- Simner, M. L. (1990). Printing performance school readiness test. *Acad. Ther.* **25**: 369–375.
- Simner, M. L. (1996). The use of handwriting legibility scales in grade one to help identify children at risk of early school failure. In: Simner, M. L., Leedham, C. G., and Thomassen, A. J. W. M. (eds.), *Handwriting and Drawing Research*, IOS Press, Amsterdam, pp. 197–202.
- Sims, E., and Weisberg, P. (1984). Effects of page prompts on beginning handwriting legibility. *J. Educ. Res.* **77**: 360–365.
- Smits-Engelsman, B. C. M., Niemeijer, A. S., and Van Galen, G. P. (2001). Fine motor deficiencies in children diagnosed as DCD based on poor grapho-motor ability. *Hum. Mov. Sci.* **20**: 161–182.
- Smits-Engelsman, B. C. M., Schomaker, M. M., Van Galen, G. P., and Michels, C. J. M. (1996). Physiotherapy for children's writing problems: Evaluation study. In: Simner, M. L., Leedham, C. G., and Thomassen, A. J. W. M. (eds.), *Handwriting and Drawing Research: Basic and Applied Issues*, IOS Press, Amsterdam, pp. 227–240.
- Smits-Engelsman, B. C. M., and Van Galen, G. P. (1997). Dysgraphia in children: Lasting psychomotor deficiency or transient developmental delay? *J. Exp. Child Psychol.* **67**: 164–184.
- Smits-Engelsman, B. C. M., Van Galen, G. P., and Portier, S. J. (1994a) Psychomotor aspects of poor handwriting in children. In: Simner, M. L., Hulstijn, W., and Girouard, P. L. (eds.), *Contemporary Issues in the Forensic, Developmental and Neurological Aspects of Handwriting* (Monograph of the association of forensic document examination, Vol. 1), Association of Forensic Document Examiners, Toronto, pp. 17–44.
- Smits-Engelsman, B. C. M., Van Galen, G. P., and Portier, J. (1994b). Psychomotor development of handwriting proficiency: A cross-sectional and longitudinal study on developmental features of handwriting. In: Faure, C., Keuss, P., Lorette, G., and Vinter, A. (eds.), *Advances in Handwriting and Drawing. A Multidisciplinary Approach*, Europa Press, Paris, pp. 187–205.
- Smits-Engelsman, B. C. M., Van Galen, G. P., and Michels, C. G. J. (1995). De leekrach beoordeeld: Inschatting van schrijfvaardigheidsproblemen en motorisch achterstand bij basisschool leerlingen (Prevalence of poor handwriting and the validity of estimation of motor proficiency and handwriting performance by teachers). *Tijdschrift voor Onderwijsresearch* **20**: 1–15.
- Smits-Engelsman, B. C. M., Van Galen, G. P., and Shoemaker, M. M. (1998). Theory-based diagnosis and subclassification in the developmental coordination disorder. In: Rispens, J., Van Yperen, T. A., and Yule, W. (eds.), *Perspective on the Classification of Specific Developmental Disorders*, Academic Publisher, London, England, pp. 245–264.
- Sovik, N., Arntzen, O., and Thygesen, R. (1987a). Relation of spelling and writing in learning disabilities. *Percept. Mot. Skills* **64**: 219–236.

- Sovik, N., Arntzen, O., and Thygesen, R. (1987b). Writing characteristics of "normal," dyslexic and dysgraphic children. *J. Hum. Mov. Stud.* **31**: 171–187.
- Sovik, N., Arntzen, O., Samuelstuen, M., and Heggberget, M. (1994). Relations between linguistic wordgroups and writing. In: Faure, C., Keuss, P., Lorette, G., and Vinter, A. (eds.), *Advances in Handwriting and Drawing. A Multidisciplinary Approach*, Europa Press, Paris, pp. 231–246.
- Starch, D. (1919). Educational research and statistics: A scale for measuring handwriting. *Sch. Soc.* **9**: 155–158.
- Stott, D. H., Henderson, S. E., and Moyes, F. A. (1987). Diagnosis and remediation of handwriting problems. *Adapt. Phys. Act. Q.* **4**: 137–147.
- Stott, D. H., Moyes, F. A., and Henderson, S. E. (1984). *Diagnosis and Remediation of Handwriting Problems*, Hayes Publishers, Burlington, ON.
- Stowitschek, J., Ghezzi, P., and Safely, K. (1987). "I'd rather do it myself"; Self evaluation and correction of handwriting. *Educ. Treat. Child.* **10**: 209–224.
- Sudsawad, P., Trombly, C. A., Henderson, A., and Tickle-Degen, L. (2001). The relationship between the Evaluation Tool of Children's Handwriting and teachers perceptions of handwriting legibility. *Am. J. Occup. Ther.* **55**: 518–523.
- Teulings, H. L. (2001). Optimization of movements duration in accurate handwriting strokes in different directions in young, elderly, and Parkinsonian subjects. In: Meulenbroek, R. G. J., and Steenbergen, B. (eds.), *Proceedings of the Tenth Biennial Conference of the International Graphonomics Society*, University of Nijmegen, IGS, The Netherlands, pp. 40–45.
- Teulings, H. L., and Thomassen, A. J. W. M. (1979). Computer aided analysis of handwriting movements. *Vis. Lang.* **13**: 218–231.
- Thorndike, E. L. (1910). Handwriting. *Teach. Coll. Rec.* **11**: 83–175.
- Trap-porter, J., Gladden, M. A., Hill, D. S., and Cooper, J. O. (1983). Space size and accuracy of second and third grade students cursive handwriting. *J. Educ. Res.* **76**: 231–233.
- Tseng, M. H., and Cermak, S. H. (1991). The evaluation of handwriting in children. *Sens. Integr. Q.* **XIX**: 3–6.
- Van Galen, G. P. (1991). Handwriting: Issues for a psychomotor theory. *Hum. Mov. Sci.* **10**: 165–192.
- Van Galen, G. P., and Morasso, P. G. (1998). Neuromotor control in handwriting and drawing: Introduction and overview. *Acta Psychol.* **100**: 1–7.
- Van Galen, G. P., Meulenbroek, R., and Bouwhuisen, C. (2001). Neuromotor noise as a window upon cognitive effects in motor control. In: Meulenbroek, R. G. J., and Steenbergen, B. (eds.), *Proceedings of the Tenth Biennial Conference of the International Graphonomics Society*, University of Nijmegen, IGS, The Netherlands, pp. 68–73.
- Van Galen, G. P., Portier, S. J., Smits-Engelsman, B. C. M., and Shomaker, L. R. B. (1993). Neuromotor noise and poor handwriting in children. *Acta Psychologica* **82**: 161–178.
- Waber, D. P., and Bernstein, J. H. (1994). Repetative graphomotor output in L.D. and nonlearning-disabled children. The Repeated Patterns Test. *Dev. Neuropsychol.* **10**: 51–65.
- Wallen, M., Bonney, M., and Lennox, L. (1996). *The Handwriting Speed Test*, Helios, Adelaide, Australia.
- Wann, J. P. (1987). Trends in the refinement and optimization of fine-motor trajectories: Observations from an analysis of the handwriting of primary school children. *J. Mot. Behav.* **19**: 13–37.
- Wann, J. P., and Jones, J. G. (1986). Space-time invariance in handwriting. *Hum. Mov. Sci.* **5**: 275–296.
- Wann, J. P., and Kadiramanathan, M. (1991). Variability in children's handwriting computer diagnosis of writing difficulties. In: Wann, J., Wing, A. M., and Sovik, N. (eds.), *Development of Graphic Skills*, Academic Press, London, England, pp. 223–236.
- Ziviani, J. (1984). Some elaborations on handwriting speed in 7- to 14 years old. *Percept. Mot. Skills* **58**: 535–539.
- Ziviani, J. (1996). Use of modern cursive handwriting and handwriting speed for children aged 7 to 14 years. *Percept. Mot. Skills* **82**: 282.

- Ziviani, J., and Elkins, J. (1984). An evaluation of handwriting performance. *Educ. Rev.* **36**: 249–261.
- Ziviani, J., and Elkins, J. (1986). Effects of pencil grip on handwriting speed and legibility. *Educ. Rev.* **38**: 247–257.
- Ziviani, J., and Watson-Will, A. (1998). Writing speed and legibility of 7–14 year old school students using modern cursive script. *Aust. Occup. Ther. J.* **45**: 59–64.