

The assessment of spatial features and kinematics of characters: an analysis of subjective and objective measures

Anne Hillairet de Boisferon, Jeremy Bluteau and Edouard Gentaz
*Laboratoire de Psychologie et NeuroCognition, Grenoble Université
France*

1. Introduction

Handwriting is a complex daily activity that involves attention, memory, linguistic, cognitive and perceptual-motor skills. As motor act, writing a letter requires to retrieve the letter storing in memory, to access the corresponding motor program (letter global shape, relative size of letter strokes), to set the parameters for the program (absolute size of letters and writing speed) and to execute the program (muscles recruitment) (Ellis, 1988; Van Galen, 1991). The letter to be traced and the corresponding graphic motion are intimately related in handwriting activity. Because both reading and writing are learned simultaneously in school, it may be assumed that letters are both coded visually and under a sensorimotor form (Hulme, 1981). From this point of view, the most notable example is probably Chinese or Japanese ideograms, which are composed of a number of strokes that must be written in a precise order when learning to read and write. The metaphor “grammar of action” was proposed by Goodnow & Levine (1973) to define stroke composition rules. Subsequently, this order is used as a cue to retrieve the ideograms from memory (Flores d’Arcais, 1994), suggesting that the motor schema specific to each ideogram may be an essential component of their representation. Many others arguments in favour of the tight coupling between the visual and sensorimotor representations of letter shapes can be advanced. First, Anderson et al., (1990) describe the case of a patient whose inability to write letters can be associated with deficits in the visual identification of letters. By contrast, she could easily read all numbers and nonverbal symbols, and she was equally able to write numbers and perform written calculations without difficulty. In a same way, writing movements can help alexic patients whose reading abilities are impaired. When they were asked to trace the outline of the letters with their fingers, they sometimes succeeded in recognizing letters they were not able to recognize only visually (Bartolomeo et al., 2002; Seki et al., 1995). To go further on, some researchers investigated the question of the presence of a global cerebral network including visual and sensorimotor components and mediating a multimodal representation of letters. Longcamp and colleagues (2003) showed activations of a part of the left premotor area during passive observation of isolated letters although no motor response was required. The same zone was also activated when the

participants were writing the letters. They suggested that handwriting motions might therefore be activated in memory by the visual presentation of letters. Moreover, this multicomponent neural network could be built up while learning concomitantly to read and write.

In this framework, handwriting acquisition consists in learning the visual representations of letters, which are used to guide their production, and the motor representations (motor programs) specific to each one. Longcamp, et al. (2005) have studied two groups of preschool children (aged 3–5 years) who were learning letters either by handwriting or by typing, and compared letter recognition performances one week later. Results showed that in the older children, handwriting training gave rise to a better letter recognition performance than typing training. Further research in multisensory-training protocols, as opposed to unisensory protocols, produce greater and more efficient learning. These results indicate that multisensory training promotes more effective learning of the information than unisensory training (Ernst & Bühlhoff, 2004; Bluteau et al., 2008; Fredembach et al., 2009). The benefit of multisensory exposure is persistent even when information is gathered from unisensory condition (Shams & Seitz, 2008). These previous results suggest that character recognition abilities are somehow dependent on the way we learn to write and to read. To acquire proficient handwriting is required to produce legible texts to be read, and to a large extent, to communicate. Indeed, we write words manually to be read, character production is mainly guided by this implicit subjective recognition goal. However, handwriting acquisition is neither trivial nor effortless, and it takes many years of instructions to master this skill. Difficulties are also observed in adults involved in learning new handwriting systems. At the beginning of learning, movements are slow and guided by visual and kinaesthetic feedback resulting in letter forms not yet mastered. With practice, writing becomes more automatic and the control of movement is mostly proactive, that is to say, based on an internal representation of motor acts. The developmental changes in the product and the process of handwriting could be the consequence of a change from retroactive control of movement (based on sensorial, visual and kinaesthetic feedback) to proactive control (Zesiger, 1995; Bara & Gentaz, 2010; Hillairet et al., 2007).

Considering this whole framework, one can understand that handwriting recognition presents a challenge for most researchers working on letter perception. Indeed, how can people accurately discriminate letters given the important variability in handwritten forms? Classically it is assuming that to recognize handwritten letters people must be able to accept distortions on the standard letter. Freyd (1983) proposed an alternative to classical view, and demonstrates that handwriting recognition makes use of information about how the letters are formed. Specifically, perceivers could spontaneously infer the underlying dynamics pattern of motor movements used for a particular handwritten letter by applying their own knowledge of the production processes to its static trace.

To sum up, subjective handwriting recognition evaluation is accurate and relevant for character recognition analysis. A better understanding of the types of information (shape, kinematics, motor internal simulation) used by the perceiver could enhance the development of almost essential computerized character recognition methods. The finding that readers spontaneously extract production information from static handwritten characters may have implications beyond handwriting recognition. Essentially, in developing a handwriting recognition interface one should take into account static input as well as dynamic characteristics of handwriting. As we have seen, character production and

recognition both deal with static and dynamic features of letters, because spatial shape and kinematics of production are intimately related in handwriting activity. Finally, the question of “how quantifying handwriting and its static and dynamic characteristics?” despite a large extend of researches is not trivial and remains crucial. The choice and relation between subjective/objective, static/dynamic evaluation criteria is a decisive factor for character recognition.

2. Measure of writing performance

2.1 Historical approach

Over the years, many methods were developed for the evaluation of handwriting proficiency. Since academic instruction aims to write legibly and rapidly, quality and rapidity criteria seemed sufficient to evaluate handwriting. Thus, most of evaluations are based on analyzing the handwritten product and speed. Nevertheless, authors who work on handwriting acquisition run up against the problem of assessment. Since decades, researchers continuously tried to develop and improve standardized evaluations and proposed numerous tools, which can be classified according to whether they involve qualitative or quantitative measures, global or analytic scales, or measures of the handwriting product or process (for a review, see Rosenblum et al. 2003).

Historically, first evaluations were dedicated to an overall judgment of readability of written products. Handwriting production was first evaluated for its “global quality” or “legibility” before researchers developed more analytic evaluations based on predetermined criteria considered as important factors in written products quality. One of the first scales devoted to assess global quality of written product was the Thorndike Scale for Handwriting of Children (1910) (from fifth to eighth grade) based on the rating of “general merit”. After handwriting was evaluated for its “general merit”, authors proposed scales based on the attribution of an average score assigned by a group of judges who compared written texts to handwriting samples previously graded from “readable” to “unreadable”.

In the aim to provide less subjective judgments, some authors proposed to replace global scales by using clearly defined criteria to grade handwriting samples. Analytic scales then gradually replaced the earlier global evaluations. With analytic scales, the various characteristics of handwriting considered as playing a role in the overall quality of written product are rated individually. The most common criteria used to judge writing legibility are letter form, size, slant, spacing, and line straightness. For example, Freeman (1959) scale included the following five criteria: tilt, height, shaping of letters, line quality and an overall score representing the general merit. These parameters still lead to recent development of character production analysis software (Guinet & Kandel, 2010). One of the more used analytic scales is The Concise Evaluation Scale for Children’s Handwriting (Dutch abbreviation BHK; Hamstra-Bletz et al., 1987) first developed to examine the readability and speed of writing performance in young dysgraphic children. It should be noted that quantitative measures are sometimes preferred over qualitative ones because it is easier to quantify fluency (e.g., the number of characters a child is able to write in one minute or the total time taken by a child to complete a given text of a predefined number of characters) than legibility (scoring readability of handwriting products requires judges expertise and laborious comparison to numerous standards). For the BHK evaluation, children are asked to copy a standard text that is presented to them on a card for five minutes. The first five

sentences (grade 3 level of reading) are evaluated by judging deviations of the child's writing from the standard handwriting text according to 13 criteria (e.g. global size, line straightness, spacing, letters joins, letters distortions, ambiguous shapes, overlaps between letters, wavering and trembling). 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). Copying speed is calculated according to the number of letters written in five minutes. The BHK diagnostic sensitivity, the development of norms and use among children in various populations explain the extensive use of the BHK in studies and clinical practice (Blöte & Hamstra-Bletz, 1991).

Although most researchers agree upon criteria used in analytic scales (Bruinsma & Nieuwenhuis, 1991), approaches used to collect handwriting products vary across studies and factors such as the nature of the handwriting task (e.g. copying or writing from memory tasks), given instructions (e.g. fast or slow handwriting) and graphic workspace (e.g. school writing paper or not) can influence written products quality. Moreover, scales are also designed in different aims, including handwriting difficulties assessment, detection of children being at risk to develop handwriting difficulties, developmental changes assessment, etc. Methodological variations hinder direct comparisons between analytic scales and limit the development of optimally effective handwriting assessment. Nevertheless, global and analytic scales permit an analysis of the handwriting final product. Considering the fact that handwriting is a highly dynamical process; it appears that evaluation of handwriting products do not provide many information about the underlying handwriting process. With the development of computerized measures, it is possible to assess the handwriting process while children are writing. Because handwriting movements require a precise organization in time and space, and proper control over pressure, spatial, and temporal measures during writing supply information about the degree of handwriting proficiency. In literature, we face an increasing number of measures dedicated to the analysis of specific aspects of the handwriting process (e.g. average velocity, production time, movement fluency) and dedicated to more static or global criteria (e.g. recognition rates, height, strokes number).

In conclusion, both approaches which assess the product or the process of handwriting have their advantages and inconveniences. Subjective evaluations (when someone has to judge the quality of the handwriting product) suffer from limited accuracy, sensitivity, and reliability, but are simple to implement and nearer to the natural situation of handwriting in the classroom. Moreover, in many cases, for example in presence of motor production noise in learning process or in handwriting troubles which deteriorated presentation conditions, experts still give their preference to subjective judgments of the quality even if human expert can also make mistakes. "In most applications, the machine performances are far from being acceptable, although potential users often forget that human subjects generally make reading mistakes" (Barrière & Plamadon, 1998). The more objective, computerized analyses make possible to evaluate handwriting dynamics by providing more accurate and more reliable data by means of rapid, automated procedures, but, the practical applications (clinical or educational issues) are still limited because the lack of global decisions about the legibility of a written product. In mainly cases of character recognition, subjective judgments remain more accurate than their computerized corresponding scores.

2.2 Human performance

The measure of performance in the execution of a task raises a number of issues, largely present in the literature under the term “motor learning and human performance” (Schmidt & Wrisberg, 2000). This measure depends on three main factors: the subject, the task and the environment.

First of all, the subject which is the fundamental element of any condition of motor performance comes with a number of characteristics: inherent abilities, cognitivo-motor knowledge, sociocultural context and level of motivation. These characteristics can influence the way of carrying out the task and the performance. The estimation of these capabilities such as the level of subject expertise and the nature of the population to be tested may help up in the understanding of this influence. As an example, several normalized tests are applied to a limited population (child, in remedial persons, etc.).

The second factor in motor learning and human performance is the environment. It can affect the production of a task by the application of temporal constraints (limited time to complete a task), or spatial constraints. In the area of character recognition, the limitation by writing templates lines is an example of these environmental spatial constraints.

The last factor influencing the measure of performance is the task. The nature of the task directly affects the demand for performance and achievement. Certain tasks have high sensory demands, such as detection of an approaching ball to return to tennis. Others have high cognitive demands for action, planning and implementation of action. Finally, we can consider the competence necessary to execute the task, that is to say, the ability of the person to make the right move. In some tasks, only one of these factors will determine the performance of the person, but it is more often a succession of analysis, planning, decision and implementation of a gesture that is indicative of the degree of success or failure of the task. A first answer is to classify the task according to the progress of the task: discrete action, actions in series or ongoing activities. Discrete actions are generally fast and well defined from the beginning to the end. A pointing task belongs to this category. Actions in series are a succession of several discrete actions, connected in sequence and whose order is crucial in the successful accomplishment of the task. Some writing of characters (especially composed by several strokes like Chinese characters) obeys to this classification. Finally, the continuous actions are defined by the absence of precise start and end, and are generally repetitive. At a certain level of handwriting expertise, when subject has internalized and automated the motor act, handwriting can be classified as continuous action. Another classification of the tasks is to determine the sensorimotor load and cognitive load. For a beginner writer, drawing a character is mainly interfered with the control of its movements. Gestures are slow and guided by visual and kinaesthetic feedbacks. Children have to constantly check on their handwriting trace in order to guide their fingers in the right way. Then, with fluency, writing access the stage of cognitive treatments, by concatenating letters to make words, and then sentences with meaning. As an example, we can also cite the learning of Japanese or Chinese characters. At the end of the first elementary cycle, the students have to know a small set of 1006 kanji (denoted as gakushuu kanji) ordered by increasing level of difficulty. Finally, tasks can also be classified according to the degree of predictability of the environment. A task performed in a changing, unpredictable and open environment, requires an adaptation of the person. Writing a phone number while someone is driving on a bumpy road is a quite unpredictable environment. A closed task will be realized in a stationary environment and predictive. The person could then plan ahead the

completion of the task. That fact explains the needs of closed tasks for most of the normalized test or objective measure of character recognition to be effective. Gentile (Gentile, 1987) proposed a classification in two dimensions, depending on the predictability of the environment and the sensorimotor and cognitive load of the task. The change of type of task highly influences the intra individual performance of the task (Higgins & Spaeth, 1972; Franks et al., 1982), as illustrated in figure 1.

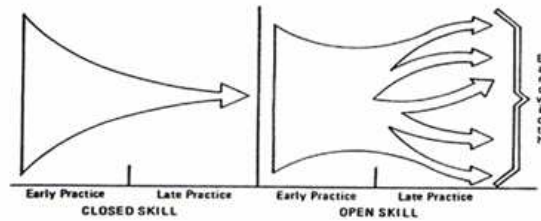


Fig. 1. Change of intra individual performance variability during closed or open skill environment. From Higgins & Spaeth (1972).

In this context, the question of which measure criteria should be used to evaluate motor performance remains. To clarify the multitude of possible measures, authors working in the field of human performance list three main types of measures (Guthrie, 1952):

- 1) **the maximum certainty of achieving the goal;**
- 2) **the minimum energy expenditure,** and
- 3) **the minimum completion time.**

Their variation is seen as a variation of the degree of achievement of performance, from performance considered as beginner level to expert level (see figure 2). Performers who are more proficient in movements designed to achieve a particular goal usually demonstrate one or more of the qualities mentioned previously.

Reference	Early stage of learning	Later stage of learning
Fitts and Posner (1967)	Cognitive (trial and error), associative (homing in)	Autonomous (free and easy)
Adams (1971)	Verbal motor (more talk)	Motor (more action)
Gentile (1972)	Getting the idea of the movement	Fixation and diversification (closed or open skill)
Newell (1985)	Coordination (acquire the pattern)	Control (adapt the pattern as needed)
Associated Motor Performance Characteristics		
Early learning		Later learning
Stiff-looking	More relaxed	Automatic
Inaccurate	More accurate	Accurate
Inconsistent	More consistent	Consistent
Slow, halting	More fluid	Fluid
Timid	More confident	Confident
Indecisive	More decisive	Certain
Rigid	More adaptable	Adaptable
Inefficient	More efficient	Efficient
Many errors	Fewer errors	Performer recognizes errors

Fig. 2. Theoretical depictions of the stages of motor learning and associated motor performance characteristics. From Schimdt & Lee (1987).

The maximum certainty of goal achievement implies that a person is able to meet a performance goal regardless to the situation, on demand and without luck. This criterion can often be seen as a combination of low variability in task performance regarding to a predefined performance level. We can notice that variability in the required movements may help acquiring a maximum certainty of goal achievement as demonstrated in other motor learning fields such as sports (Barlett et al., 2007). The minimum energy expenditure is a consequence of a low noise action, realized without unwanted and unnecessary movements. Finally, the minimum completion time supposes that a skilled movement has a higher level if its duration is shorter than another movement, with the same level of precision. A contradiction remains with the certainty of goal achievement criterions while speed and accuracy are antinomic. Fortunately, humans seem to have the capability to swap speed for accuracy, depending on the task requirements.

The way to access to the three main types of measure of performance is founded on two types of information. The first class of information is directly accessed by internal states of sensorimotor system. Observations such as physiological parameters (ECG, EMG, etc.) or subjective evaluations about the performance belong to this group. These parameters are rich but often suffer from a complex analysis mainly due to the motor noise and complexity of the system. Furthermore, external states such as traces, on-line recording (cf. figure 3.a) or off-line recording (cf. figure 3.b) and observer evaluations can be seen as a second group of measurements. These former parameters suffer also from computation noise and/or subjectivity of observers, but are quite easier to analyse due to the recording of only a part of the system. In lots of performance evaluations, these criterions are preferred and proved their efficiency in constrained domains such as postal addresses (Cohen, 1991), bank check or census forms reading. In character recognition and performance measurement, both internal and external states information can be used simultaneously: for the writing task, EMG would be valuable for the estimation of minimum muscle energy expenditure and off-line record of several paper trails would be valuable for maximum certainty of achieving the goal. The two ways to access information have both complementary meanings. The special case of on-line recording allows access to the internal states of the system (pressure, forces, velocity, angles...) through the development of new sensors, and can be seen as belonging to both groups. For a detailed survey on off-line and on-line character recognition, see Plamadon (Plamadon, 2000).

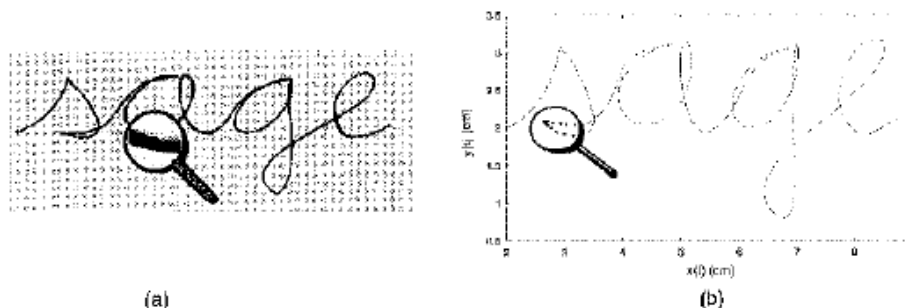


Fig. 3. a) Off-line word. The image of the word is converted into grey-level pixels using a scanner. b). On-line word. The x,y coordinated of the pen tip are recorded as a function of time by a digitizer (haptic device, tablet,...). From Plamadon, 2000.

3. Challenges: Which criteria do I have to choose for what?

Regarding human performance knowledge, a set of criteria chosen in the three main types (maximum certainty of achieving the goal; minimum energy expenditure, and minimum completion time) correctly depict the analysis of human performance. Nevertheless, the authors working on the acquisition of handwriting still face the problem of evaluation of character recognition. What class of criteria should we adopt in assessing the performance of manual gestures such as writing? Another way to classify character recognition performance criteria is based on the separation between objective and subjective measures. Subjective (qualitative) criteria result from a judgment, and objective (quantitative) criteria refer to a computerized numerical analysis. Creating tools to assess the quality of writing (i.e. character recognition) remains difficult despite the presence of many standardized assessments, and many measures that can be classified according to their consideration of qualitative or quantitative parameters. Historically, the production of writing was first evaluated on subjective criterions. Technological advances in computation and preference for objective scientific method has reversed the trend. However, the use of subjective criteria is still relevant for character recognition performance analysis and stay the final goal of character production. Indeed, we write words manually with the final objective to be read (i.e. give a trace to be visually evaluated). The character production is mainly guided by this implicit subjective recognition goal. In addition, two levels are accessible for such manual gesture and are offering to evaluate the performance: the writing process (gesture) and production (static trace) (Rosenblum et al., 2003). Each level includes different information respectively kinematics of gesture and static final quality of character. The fact that writing is a highly dynamic process with the support of strong relations between action and perception, allow us to study the process of production (gesture) to enrich the product evaluation (static trace). Technological advances made possible to study and quantify the links between spatial accuracy and kinematics of handwriting. But we face with a plethora of measures in the literature dedicated to the analysis of specific aspects of the writing process (average speed, production time, fluidity of movement...) and static global criteria based on the production (recognition rate, size of letters, number of strokes,...) with few links among them.

To clarify criteria choice for both gesture and handwriting products evaluation and quantification, we propose to identify the relation between static (product) and dynamic (handwriting process) measures in an objective and subjective evaluation of on-line acquisition of writing.

As suggested by several researches presented in §1, many connections between perception and action can be observed in humans. As the judgment of product and process involve human perception, correlation amongst static subjective criteria (mainly related to shape) and dynamic subjective criteria (related to kinematics) should appear (Hypothesis 1). Literature tends to distinguish in one hand static objective criteria related to shape and in the other hand, dynamic objective criteria, related to kinematics of production. Then, we should observe some strong links between static measures, and strong links between dynamics criteria. However, this hypothesis has to be contrasted in regards to existing tight coupling between spatial shape and kinematics in handwriting production. So, we also assume some correlation between static and dynamic objective measures (Hypothesis 2). The obvious example of mean velocity computation (distance divided by time) tends to suggest a relation between static objective measure (distance) and dynamic objective measure (time). At last, comparison between subjective judgments and objective measures

should be somehow related, justifying the current use of objective criteria instead of the classically performed subjective character recognition and evaluation (Hypothesis 3).

4. Experimental evaluation of kinematics and spatial features: an analysis of subjective and objective measures

To access relations between spatial criteria and kinematics of handwriting and relations between subjective and objective judgments, we designed a two step experiment. The first phase consisted in the acquisition of children and adults handwriting. The second phase consisted in an evaluation of handwriting with an objective computation of criteria and subjective judgments.

4.1 Acquisition of handwriting

The acquisition phase differs in children and adults. By choosing two different populations, we wanted to access whether traditional handwriting evaluations could be generalized to pre-scripser children, who present more variable handwriting and are more susceptible to noise generated by the establishment of fine motor control. In addition, we wanted to generalize this evaluation with the learning of novel trajectories with adults. In traditional handwriting of cursive Latin characters, adults are considered as expert-scripser and a clear ceiling effect would have occurred on each evaluation criteria. Indeed, to avoid this effect, unknown trajectories were proposed to the participants. This choice is also related to the characteristics of subjects (developmental difference in sensory motor control) which have to be considered as proposed by human performance researches (cf. §2).

4.1.1 Participants

Forty-four children between the ages of 4.9 and 5.9 months (21 boys and 23 girls, mean age: 5.3) from two senior kindergarten classes in Grenoble participated in this study. All participants spoke French as their first language and no child had a statement of special educational needs. Permission for recruitment was gained from the head teacher of the school, and written informed consent for the participation of the children was obtained from their parents. At the same time, we asked 23 adults participants aged between 18 and 26 (including 13 girls, mean age: 21.3 ± 2.5) to participate in this study. All adult participants were unfamiliar with Arab or Japanese languages, and none of them had known motor trouble or neurological dysfunction. Their participation was done after their informed written consent, in respect with Helsinki declaration.

4.1.2 Method

The acquisition methodology differs in children and adults in order to take these two population specific needs.

On the one hand, children were seated comfortably in front of a table, upon which a digital tablet (Wacom®) was placed. In this measuring system, the positions of the pen were sampled at a frequency of 100 Hz and at a spatial resolution of .008 cm. The pen used in order to write on the tablet was a ball-point pen (Intuos Ink Pen, Wacom®) allowed to receive feedback of the written samples. A white paper was placed on the digital tablet. We asked children to copy on the paper the 26 cursive letters of the alphabet. Each letter was presented separately on a

paper placed in front of the child (for example see fig 4). There were no time and size constraints. The order of letter presentation was counterbalanced across participants. The test lasted approximately 10 min by participants. We randomly chose 250 trajectories over the 1144 collected for the incoming subjective judgments and objective analysis.

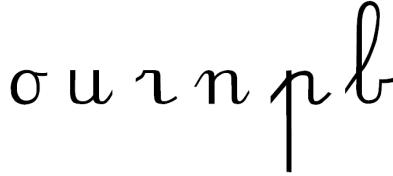


Fig. 4. Examples of standard cursive letters proposed to children participants

On the other hand, adult participants were asked to produce four foreign characters (Japanese inspired and Arabic letters - cf. figure 5). The choice of novel trajectories was a way of experimenting participants with a lower motor skill level, in order to avoid ceiling effect in both handwriting performance and character recognition. The digitalization of their traces was performed using a haptic device (PHANToM Omni® from SensAble). The desired trajectory was displayed on a horizontal screen and the participant's pen trajectory was recorded from the haptic device, placed over the screen (Bluteau et al., 2008). Ergonomics efforts have been made to achieve this virtual co-located configuration, close to the real writing task, allowing standardize protocol of trajectories presentation and recording (Bluteau et al., 2008). As a result, we recorded at 1000 Hz the positions and forces applied during the drawing of characters. Each adult participants has to draw 20 trajectories, given in a pseudo random order (two consecutive and identical letters were not allowed). We kept only 250 of the 460 trajectories, randomly chosen for our analysis.

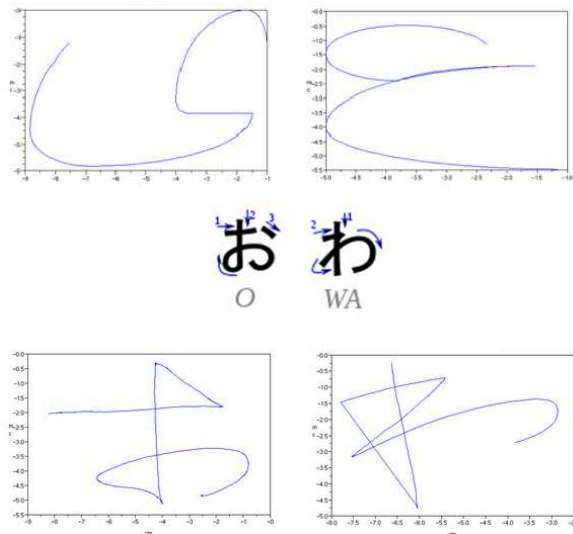


Fig. 5. Trajectories proposed to adult participants. The two upper letters were issued from Arabic alphabet. The two lower letters were issued from Japanese hiragana alphabet, with the order of drawing indicated above.

4.2 Evaluation of handwriting

In this second phase, subjective judgments and objective evaluations were performed. As we wanted to study the relationship between human evaluations and objective measures, the recordings from children and adults productions went through these two evaluations.

4.2.1 Subjective evaluation

Method

First of all, a “judgments” software (NoteSub) was developed to normalize the presentation of trajectories and gather the judgments. Two different display methods were proposed to get static (accuracy of the trace mainly based on spatial characteristics) and dynamic judgments (kinematics of the motor production) (cf. figure 6). The static display in which each of the shapes appear simultaneously on a computer screen was used to assess the quality of the product. The dynamic display, in which letters were printed on a computer screen, according to the cinematic of production of the writer, was used to evaluate the process of writing (kinematics). Two orders were given for either static presentation: “judge the graphical quality”; and dynamic presentation: “judge the quality of movements” of the presented letters. The letters were presented randomly in blocks of 50 letters for a static or dynamic judgment, in order to avoid a fatigue effect. The order of these blocks was balanced using a standard Latin square protocol. In total, judges rated two times each of the 250 letters (a static judgement and a dynamic judgment). Each judge had to recognize the character before evaluating the quality (rate from 1 – lowest quality, to 10 – highest quality) and were asked to describe the underlying criteria on which they based their judgment. In addition, a judgment was considered by the software as valid only when an “active” displacement of the rating cursor was performed (in order to avoid by default judgment). Finally, we obtained two subjective ratings for each trace: a score based on the product of handwriting, closely related to the shape or spatial criteria and a score of production process, closely linked to kinematics.

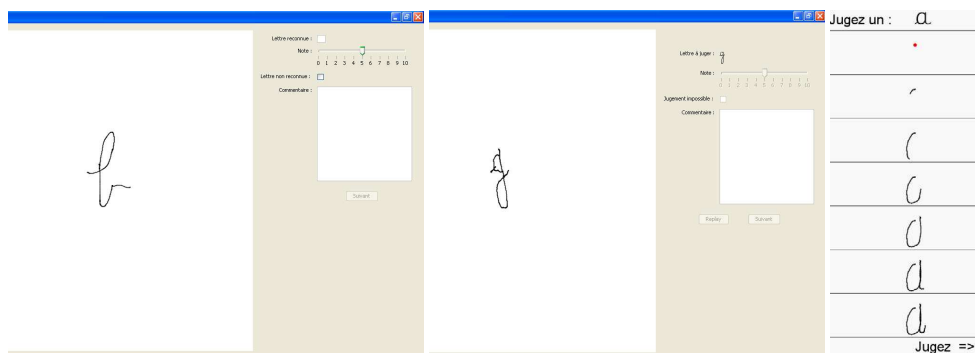


Fig. 6. Screen capture of NoteSub, the subjective evaluation software. (Left) Static Presentation of letters. (Center) Dynamic presentation of letters. (Right) Example of dynamic presentation, the letter appears gradually, in respect with subject movement velocity, pauses and pen lift up.

Participants/Judges

Ten judges evaluated the quality of character and writing process. Four of them were asked to judge the 250 children productions and six were asked to judge the 250 adult productions. Concerning adults production judgments, evaluators mean age was 32.6 years (± 9.6); concerning children production judgments, evaluators mean age was 28.2 years (± 2.8).

Subjective criteria

As previously describe, we asked judges to rate the “graphical quality” of the handwriting product (static trace) and the “movement quality” of the handwriting process (dynamic). In addition, for each set of letters, judges were asked to report underlying criteria of their judgments

a. Static presentation

Basing their assessment on a static presentation (called “static judgment” later in this chapter) judges had access to traditional presentation of character. Their character recognition performance and evaluation should include several spatial criteria such as tilt, orientation and shaping of letters encountered in literature. However, we can hypothesize that judges could also access to cinematic characteristics of the production given its subtle link to the written shape. For example, static clues such as curvature of letters could call to judge corresponding knowledge of rule production, the well known “two-third power law” (Viviani & Terzuolo, 1982) and allow access to cinematic parameters such as velocity. But we believe that their static judgment will be mainly based on spatial shape information.

b. Dynamic presentation

In the dynamic presentation, the judge had access to kinematics of production. This evaluation (called “dynamic judgment” later in this chapter) allows the extraction of cinematic parameters induced by pauses, accelerations and order of strokes. In order to avoid judgment based on shape’s features, the character immediately disappears after dynamic presentation. We make the assumption that their dynamic judgment will be mainly based on kinematics of production. To our knowledge, this kind of kinematics judgments has not been implemented in literature but appears to be a substantial source of information.

4.2.2. Objective evaluation

Method

In parallels with subjective evaluation, the computation of objective measures was done using a normalized Matlab script. For each of the 250 adults and children trajectories, we calculated a number of objective measures, inspired by the literature. Quantitative spatial measurement considered were the number of strokes, the distance and a score of similarity between the “experimental trajectory” and “theoretical trajectory”. This score (Dynamic Time Warping - DTW) provides access to a criterion of similarity of form. The quantitative cinematic measures considered include the duration of movement, duration of pauses, average speed and the number of velocity peaks. Their calculation was performed by a Matlab™ script, from the acquisition of positions at 1000 Hz for the adult production and 100 Hz for the children productions. In addition for further analysis, we computed predictors (or control variables) by taking in account the total distance of the track, standing for the difficulty of trajectories and the level of initial motor skills for children by a figure copy task. These various objective criteria are described in details below.

Objective criteria

a. Spatial shape

Number of Strokes and Number of pen up/pen down

We defined a stroke as a continuous drawing of trajectory according to pen up/pen down actions. These measures are linked and are both indicators of the difficulty of the letter and the global formation of the shape of the letter. By counting the number of strokes, we are able to extrapolate the degree of fluency of writing for a specific letter. For example, in children above 7-years-old, the number of pen up/pen down (or number of strokes) is usually larger than the theoretical number of strokes required to trace the letter. This is due to absence of achieved motor program and efficient control of the trajectory (in charge of topokinesis and morphokinesis). Due to one-stroke design of adult trajectories, these criteria were only computed for children.

Distance of trajectories

We computed the total distance of the trajectories of our participants. Distance criterion could be considered as size information, one of the most common criteria used to judge writing legibility. This criterion is particularly used in children evaluation of written letters because size variations are important during handwriting acquisition. Young children's handwriting is often characterized by the production of large letters. When children better master fine movements required in writing, letters size decrease (Blöte & Hamstra-Bletz, 1991).

Dynamic Time Warping (DTW)

The computation of a distance score between two curves is a usual way of quantifying the differences or to put a figure on likeness. Classical distance measures include point-to-point distance quantification (also known as Euclidean distance), point-to-closest point or even unidimensionnal distance (known as Manhattan distance or nearest prototype). In 1983, Joseph Kruskal and Mark Liberman introduced a new technique to calculate the distance between two curves. This technique, called time warping, proposes to match the two curves by distorting time axis (or "warping" as called by its authors). This means that variation in writing speed is considered as noise and then will be deleted (or at least decreased) by the algorithm. This algorithm has been applied to many fields, including speech recognition, handwriting pattern recognition, video analysis, quality of cursive character in reference to a standard (Niels, 2004) and also sequence alignment. We will take this last example to detail the algorithm. In genetics, sequence alignment consists in transforming one sequence into another using edit operation that replaces, inserts or removes an element. Each operation has an associated cost, and the final alignment will be given by the lowest cost standing the sequence of editing operations. Note that the lengths of the two sequences do not have to be equal. The Dynamic Time Warping (DTW) belongs to dynamic programming methods, that solve complex problem by breaking it into simpler steps, and provides solutions to such genetic questions. The problem can be stated naturally as a recursion, a sequence A is optimally edited into a sequence B by either:

1. inserting the first character of B , and performing an optimal alignment of A and the tail of B
2. deleting the first character of A , and performing the optimal alignment of the tail of A and B
3. replacing the first character of A with the first character of B , and performing optimal alignments of the tails of A and B .

The same reasoning is valuable for comparison of distance between two sequences of points ($P1$ and $P2$), where step 1 stands for the computation of distance between $P1(i)$ and $P2(i+1)$, step 2 stands for the computation of distance between $P1(i+1)$ and $P2(i)$ and finally, step 3 stands for the distance computation between $P1(i)$ and $P2(i)$. The partial alignment of the two sequences (or curves) can be tabulated in a matrix, where $cell(n,m)$ contains the cost of the optimal alignment of $A[1..n]$ to $B[1..m]$ (or $P1(1..n)$ to $P2(1..m)$). The cost in $cell(i,j)$ can be calculated by adding the cost of the relevant operations to the cost of its neighbouring cells, and selecting the optimum. In other words, the global DTW cost is given by “finding the way of the valley of minimum cost into the cost matrix” (cf. Figure 7).

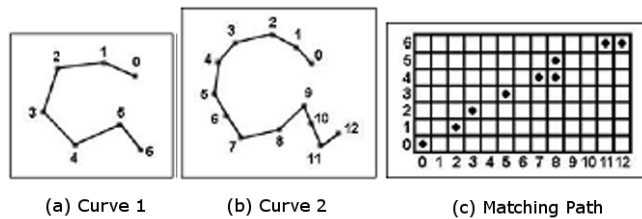


Fig. 7. Example of Matching path. (c) shows a possible matching path of the curves shown in (a) and in (b). Curve 1 is on the vertical axis and curve 2 is on the horizontal axis of the matching path. If we had filled the matching matrix with distances between points, this matching path would have been the visualisation of the way of the valley of minimum costs.

After the definition of the matching path, a backtracking algorithm allows visual checking of the alignment of the two sequences (cf. figure 8).

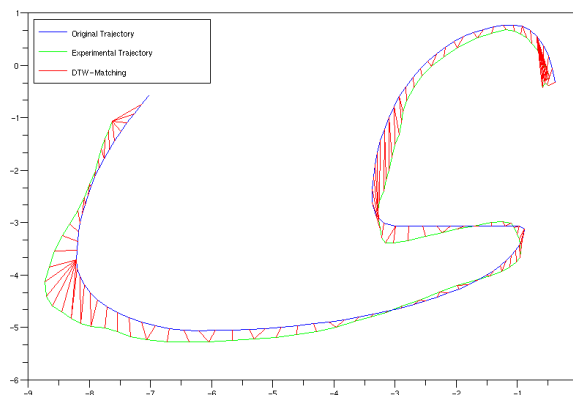


Fig. 8. Application of Dynamic Time Warping algorithm on a trajectory used in our experiment with adults and visualisation of the alignment using a backtracking algorithm. Optimal alignment (red) computed between theoretical trajectory (blue) and experimental trajectory (green) by a temporal distortion corresponding to a delay or advance between the two traces.

DTW allows comparison of sequences without re-sampling and has been demonstrated to be successful for comparison of discrete and online handwriting acquisitions (Niels, 2004;

Di Brina, 2008; Bluteau et al., 2008). This algorithm has a real application in character recognition as shown by this citation from Di Brina and colleagues (2008): "By objectively analyzing the spatial-temporal patterns, DTW captures the essential character of writing, i.e., the overall shape of its graphic output".

b. Cinematic criteria

Number of velocity peaks

This measure is classically seen as an indicator of writing fluidity and is related to the number of accelerations and decelerations during production. A movement is seen as jerky as soon as the number of velocity peaks is high. The number of velocity peaks is given by counting the number of zero crossing of the acceleration ($\partial v = 0$). The velocity profile has to be filtered to reduce acquisition noise. The filtering process implies different values in adults and children to access the number of velocity peaks. In our experiment, we used a third order Butterworth filter with a cut-off frequency of 12Hz to filter the adult productions (down-sampled at 100Hz); and we used a third order Butterworth filter with a cut-off frequency of 6Hz to filter the children productions (sampled at 100Hz). We chose Butterworth filter for its large use in literature and its intrinsic parameters (slow roll offs around cut-off frequency and no ripples) compare to other quicker roll offs cutting filters (Chebishev filter, Elliptic filter, etc.).

Mean Velocity

This measure is also a traditional criterion to evaluate the fluidity of writing. Normalized tests have integrated it by counting the number of words or characters copied in a certain amount of time (i.e., fluency); Other researchers have used it to assess the level of expertise and writing performance with adults (Bluteau et al., 2008) or with children (Palluel-Germain et al., 2007). Many researchers claim that children's competence in writing depends, in part, on their mastery of handwriting. They found that handwriting skills, particularly handwriting fluency, improve with age and schooling (Graham et al., 1998; Hamstra-Bletz & Blote, 1990) and these individual differences in handwriting fluency predict how much and how well children write (Graham et al., 1997). This measure enters in the minimum completion time class of human performance classification of criteria as soon as the required trajectory has a fixed length (distance).

Duration measures

Total duration of the trajectory

We computed the total duration to draw a character. This criterion gives an indication of the temporal performance on the path. It is linked to the average velocity and belongs to the same class (minimum completion time) of human performance classification.

Pen up duration

This criterion corresponds to the duration when the pen was up during the production of the character. An increase of this last measure indicates augmentation of breaks, and thus would reveals lacks in motor production of the character.

Number of pauses in character production

We defined a pause as a minimum of 150 ms period while the pen was down on the paper but no active movement was performed. This measure is an important indicator especially for children who often make pause while writing to visually check the model or their production in case of substantial breaks (i.e., allowing retroactive control based on visual and sensory motor feedbacks). This measure can also be an indicator of jerky handwriting in case of shorter breaks. This criterion is correlated with the level of expertise of the task.

Number of Strokes and Number of pen up/pen down

These criteria were already presented as shape specific criteria but they implicitly belong to dynamic criteria category. Indeed, the number of strokes of a writing production is intuitively correlated to cinematic criteria. A very high number of strokes reveal a quite bad cinematic of writing, thus referring to cinematic criteria. The same remark is valuable for the number of pen up/pen down criterion. This dual membership is also linked to the relation between action and perception. By the perception of some characteristics (such as the number of strokes), humans are able to deduce the actions performed to generate this product, and thus, to access the cinematic of production.

c. Predictors or Control variables

Distance of trajectories

This measure, previously classified as static, appears in the computation of the mean velocity and is implicitly linked to the difficulty of the trajectory. In a way, this criterion can be taken as a normalization index that could explain relations between kinematics and shape characteristics, especially in adults' productions.

Designs copying task (NEPSY)

We evaluate children motor and visuospatial skills using the designs copying subtest of the NEPSY – A Developmental Neuropsychological Assessment (Korkman et al., 1998). This subtest is an untimed two-dimensional constructional task that requires the integration of visuospatial analysis and graphomotor skills. Children have to reproduce paper-and-pencil copies of geometric designs of increasing complexity. These copies are then rated according to a set of indicators. Final normalized scores are comprised between 1 and 19. This measure does not count for character production evaluation but can be seen as a predictor variable of handwriting abilities.

4.3 Results

4.3.1 Preliminary analysis

After the second step of subjective judgment, we retrieved a static and a dynamic score for each of 250 letters from children or adults, associated to the character actually recognized. We kept the recording if the required character matches the recognized one, and only if all judges recognized the letter. This work has to be done to avoid nonsense statistical means. This cleaning process leads to 204 trajectories for the children set and 221 trajectories for the adult set. We performed inter-judges correlation using reliability test (calculating Cronbach's alpha) that informed us about the homogeneity of scores through judges ($\alpha = .797$ for static judgments and $\alpha = .772$ for dynamic judgments of adults characters; $\alpha = .824$ for static judgments and $\alpha = .851$ for dynamic judgments of children characters). Regarding this preliminary analysis, we averaged static scores of judges for each character in a new variable, called *mean static* later in this chapter. The same operation was done for dynamic judgments in a new variable called *mean dynamic*.

4.3.2 Criteria's descriptive data

The Table 1 represents mean results and standard deviations obtained for each subjective and objective criteria collected for both children and adults handwriting products and process.

Criteria	Children		Adults	
	Mean	SD	Mean	SD
Mean static	5.34	1.80	4.22	1.36
Mean dynamic	5.80	1.90	4.84	1.20
Total duration (s)	6.24	3.75	18.63	11.19
Pen up duration (s)	1.56	2.28	-	-
Mean velocity (cm/s)	1.84	1.02	1.37	0.66
Number of velocity peaks	12.14	7.53	5.44	2.49
Number of pauses	1.48	2.99	-	-
Number of strokes	1.96	1.45	-	-
Number of pen up	0.96	1.45	-	-
Distance (cm)	9.57	4.88	20.23	5.36
DTW	3.27	2.43	7.26	4.41
Designs copying	12.23	2.51	-	-

Table 1. Means and standard deviations (SD) of each static and dynamic criteria in children and adults.

Descriptive data reveal that despite mean distance, mean total duration and DTW score are more important and mean velocity is lower in adults trajectories, adults produced less acceleration and deceleration (i.e., number of velocity peaks) while they traced letters. These findings have to be related to the size of required trajectories in adults (cf. figure 5).

4.3.3 Correlation between static and dynamic subjective scores

We performed for each subjective judgment, correlations (Bravais-Pearson r) between static score and dynamic score attributed to children and adults handwriting products and process. Results showed a mean correlation coefficient of $r = .56$ (.55 to .60) in the judgment of children characters and a mean correlation coefficient of $r = .51$ (.35 to .61) in the judgment of adults characters.

4.3.4 Inter-correlation of static objective scores

We performed for each static objective measures inter-correlations (Bravais-Pearson r). Results are presented in table 2.

Criteria	Children			Adults	
	(1)	(2)	(3)	(1)	(2)
(1) Distance (cm)	-	.42**	.17*	-	-.33**
(2) DTW	-	-	.27**	-	-
(3) Number of strokes	-	-	-	-	-

* $p < .05$ significance, ** $p < .01$ significance

Table 2. Correlations between objective static criteria in children and adults

Table 2 indicates mean levels of correlation between the three objective static criteria scores in children. The closest link is between distance measures and dynamic time warping (DTW) scores (.42). The relations between DTW and number of strokes scores and between distance and number of strokes scores (respectively .27 and .17) are weaker. All correlations

are significant and positive. The longer the trajectories are, the more the number of strokes is susceptible to increase, and the more the gap between the theoretical trajectory and the effective trajectory is susceptible to increase. In adults, as letters were always produced within one stroke, only two criteria were taken into account. The longer trajectories are, the smaller the gap between the theoretical trajectory and the effective trajectory tends to be.

4.3.5 Inter-correlation of dynamic objective scores

We performed for each dynamic objective measures inter-correlations (Bravais-Pearson r). Results are presented in table 3. This table indicates strong to weak levels of correlation between the six objective dynamic criteria scores in children and the three dynamic criteria scores in adults. Number of velocity peaks, total duration and mean velocity scores are correlated with most of objective criteria in children as well as in adults. Compare to previous criteria, pen up duration and number of pauses show less significant correlations with other objectives measures. Finally, number of pen up shows almost no correlation with other measures. More precisely in children, the closest links are between total duration and the number of velocity peaks (.89) and between number of pen up and the total duration (.77). Weaker correlations can be observed in this population between the number of velocity peaks and number of pauses (.50), the number of velocity peaks and mean velocity are negatively correlated (-.47) and finally low correlation appears between the number of velocity peaks and the pen up duration (.14). We can notice that mean velocity score is negatively correlated with the majority of dynamic criteria (total duration, pen up duration and number of pauses).

In adults, the closest links are found between total duration and the number of velocity peaks (.70), and between total duration and the mean velocity (-.71) which are negatively correlated. Moderate negative correlation also appears between the number of velocity peaks and the mean velocity (-.35).

Criteria	Children						Adults		
	(1)	(2)	(3)	(4)	(5)	(6)	(1)	(2)	(3)
(1) Number of velocity peaks	-	-.47**	.89**	.14*	.50**	.08	-	-.35**	.70**
(2) Mean velocity (cm/s)	-	-	-.49**	-.002	-.33**	.08	-	-	-.71**
(3) Total Duration (s)	-	-	-	.09	.77**	.04	-	-	-
(4) Pen up duration (s)	-	-	-	-	-.02	.64**	-	-	-
(5) Number of pauses	-	-	-	-	-	.04	-	-	-
(6) Number of pen up	-	-	-	-	-	-	-	-	-

* $p < .05$ significance ** $p < .01$ significance

Table 3. Correlations between objective dynamic criteria in children and adults

4.3.6 Correlation between static and dynamic objective scores

We performed for each objective measures, correlations (Bravais-Pearson r) between static score and dynamic score attributed to children (table 4) and adults (table 5) handwriting products and process.

Static/Dynamic	Number of velocity peaks	Mean velocity (cm/s)	Total Duration (s)	Pen up duration (s)	Number of pauses	Number of pen up
Distance (cm)	.47**	.32**	.51**	.12	.28**	.17*
DTW	.29**	.003	.33**	.25**	.23**	.27**
Number of strokes	.08	.08	.04	.64**	.04	-

* $p < .05$ significance, ** $p < .01$ significance

Table 4. Correlations between static and dynamic criteria in children

In children, some moderate positive correlations can be observed amongst static and dynamic criteria. The closest link is observed between the number of strokes and the pen up duration (.64). It is interesting to note that the number of strokes is only correlated with this dynamic measure. Medium positive correlations are observed between distance and total duration (.51), between distance and the number of velocity peaks (.47). Weaker correlations implying distance are observed with mean velocity (.32), the number of pauses (.28) and the number of pen up (.17). We observed that the DTW (considered as a static indicator in literature) is correlated to most of dynamic scores (except mean velocity).

Should be noticed that despite inter-correlation between all static measure (table 2) and a majority of inter-correlation between dynamic measures (table 3), some correlation between static and objective measure seem important. For example, the number of pen up only correlated with the corresponding dynamic measure of pen up duration (.64) but correlated with two static criteria, the distance (.17) and the DTW (.27) scores

Static/Dynamic	Number of velocity peaks	Mean velocity (cm/s)	Total Duration (s)
Distance (cm)	.37**	.40**	.10
DTW	-.27**	.04	-.24**

* $p < .05$ significance, ** $p < .01$ significance

Table 5. Correlations between static and dynamic criteria in adults

Concerning adults characters analysis, we also observed strong correlations between static and dynamic scores. Medium and positive correlations are observed between distance and the number of velocity peaks (.37), and between distance and mean velocity (.40). The DTW measure is found to be negatively correlated with the number of velocity peaks (-.27) and with the total duration (-.24). These correlations remain moderates. As in children, some correlation between static and objective measure seem important but results also show that strongest links are found in adults between dynamic criteria (table 3), the total duration and the number of velocity peaks (.70) and the mean velocity along the path (-.71)

4.3.7 Correlation between objective and subjective evaluations in each population

To analyse the possible existing relation between subjective and objective evaluations, we performed forward stepwise regressions using subjective judgments as references. This analysis computes step by step linear regression, by including in each step predictor variable (objective criterion) with the highest shared amount of variance with the predicted variable (subjective criterion). At each stage in the process, after a new variable is added, a F-test (Fisher-Snedecor) is made to check if some variables can be deleted without appreciably increasing the residual sum of squares (RSS). The procedure terminates when the measure is maximized, or when the available improvement of the model falls below some critical value.

Regarding static criteria in children, the stepwise regression analysis reveals that number of strokes, distance and designs copying taken together explain 11% of the mean static judgment score [$F(2,200)=9.33;p=.034$]. Analysis of partial correlations indicates that number of strokes explain a unique amount of mean static judgment's variance [5%, $t(200)=-3.36,p<.001$], designs copying explain also a unique amount of variance [2.5%, $t(200)=2.24,p<.05$] and finally, distance criteria explains a unique amount of variance too at a level of 2.3% [$t(200)=-2.19,p<.05$]. On the dynamic criteria side, the analysis reveals that pen up duration, DTW and number of pauses taken together explain 38% of the mean dynamic judgment score [$F(3,200)=43.19,p<.01$]. Analysis of partial correlations indicates that the pen up duration explains a unique amount of mean dynamic judgment's variance [22%, $t(200)=-7.49,p<.001$], DTW explains a unique amount of variance [12%, $t(200)=-5.24,p<.001$] and finally, the number of pauses explain a unique amount of variance too [4.2%, $t(200)=-2.96,p<.01$].

Concerning static criteria in adults, the analysis reveals that DTW and mean velocity taken together explain 13% of the mean static judgment score [$F(2,218)=3.91,p<.05$]. Further partial correlations analysis reveals that the DTW [12%, $t(218)=-5.51,p<.001$] and mean velocity [1.8%, $t(218)=-1.98,p<.05$] both explain a unique amount of variance of the mean static judgment. The analysis of the dynamic judgment score in adults reveals that the number of velocity peaks, DTW, mean velocity and total duration taken together explain 16% of the mean dynamic judgment scores [$F(4,216)=4.14,p<.05$]. Partial correlations analysis reveal that the number of velocity peaks [3.9%, $t(216)=-2.97,p<.01$], the DTW [10.2%, $t(216)=-4.95,p<.001$], the mean velocity [4.5%, $t(216)=-3.18,p<.01$] and the total duration [1.9%, $t(216)=-2.03,p<.05$] explain each a unique amount of variance of the dynamic judgment score.

5. Discussion and Conclusion

This study was designed to identify the relationship between spatial features and kinematics of handwriting process through static and dynamic criteria in an objective and subjective evaluation of handwriting. The underlying purpose was to clarify and help criteria choice for both gesture and handwriting products evaluation and quantification.

First, statistical analysis of inter-correlation of subjective judgements puts forward the existence of a link between the static and dynamic judgments. High scores based on a dynamic presentation (and thus involving kinematics) correspond to high scores based on static one (involving characteristics related to accuracy of trace) and inversely. This result supports our first hypothesis (cf. §3) based on the action/perception and product/process links described in

literature. As suggested by Freyd (1983), human could extract dynamic information in the perception of static forms. For recall, the author asked subjects to learn some artificial characters, drawn in real time, and then to recognize distorted versions of these characters presented statically. In accordance with her theory, subjects were faster on static character recognition when the distortion was consistent with the drawing method (Badcock & Freyd, 1988; Freyd, 1983). Our results showed the same underlying process and extended it to the measure of the product quality. Characters subjectively considered as well produced leads to a good rate of the character product. In other words, when we perform character evaluation of the quality in a static manner, we could also access to kinematics information and take it into account in our final judgment. The analysis of judges' comments on criteria they used during the subjective evaluation emphasizes the supposed combination between static and dynamic components for both static and dynamic presentations. As example in static judgments, the underlying criteria are mainly related to the shape ("a space between the center loop and the end is too big", "Dissymmetry between top and bottom") and aesthetic ("precise, nice character", "clear and precise") but are also present some dynamic components ("jerky writing", "several breaks in the letter"). In dynamic judgments, same dual process exists, the underlying criteria are mainly related to kinematics ("correct movement", "regular motion", "too quick/slow", "wrong strokes order") but also related to shape criteria ("too titled", "an additional stroke") and aesthetics ("nice realisation"). Same confusing frontier exists in normalized test such as BHK, where shape criteria (size, tilt, orientation) are combined with kinematics criteria (jerkiness, speed indicators, hesitance, etc.). To conclude, in accordance with our first hypothesis, a significant overlap between criteria initially considered as purely static and purely dynamic judgments exists.

Secondly, we tried to verify the feasibility of the classic distinction proposed in literature between static objective criteria related to shape and dynamic objective criteria related to kinematics of production (hypothesis 2, cf. §3). Our second set of results reveals that in majority, dynamic criteria are correlated together (e.g. number of velocity peaks, mean velocity, duration and number of pen up) the mean correlation coefficient is $r=.53$ in children and $r=.59$ in adults. Static criteria are also correlated together (Number of strokes, distance and DTW) the mean coefficient is $r=.29$ in children and $r=.33$ in adults. Due to the influence of gesture production on the product, correlation showed that several static criteria are linked to dynamic measures (e.g. distance with number of velocity peaks, DTW and total duration of the trajectory, etc.). The mean coefficient of correlation between static and dynamic objective criteria is about $r=.32$ in children and about $r=.34$ in adults. These second set of results emphasize the difficulty to classify criteria in the two main category used in literature (static or dynamic; product and process). Moreover, interesting difference in correlations between criteria is observed in children and adults. The correlation between distance and total duration disappears in adults production. We can suggest that this effect is mainly due to absence of isochrony in children production. The isochrony law was formulated for the first time by Binet & Courtier (1893). They found that the speed of movement remains constant regardless of size variations of trace to be produced. In adults, we observed behaviour in accordance with isochrony law: the more letter size is important, the more velocity increases to remain constant in gesture duration. In children, even if we observed an increase of the velocity to preserve duration of the production on larger letter, we still notice a correlation between distance of the path and the corresponding duration. This specific correlation suggests that children do not master all fine motor control

mechanisms and cannot already respect all handwriting motor rules. In regard with this result, one should understand that scripiter expertise has to be carefully taken into account in character recognition and evaluation.

Then, we evaluated the link between subjective judgments and objective measures. We used the subjective judgments, traditionally considered as more accurate than their computerized corresponding scores, as criterion to be predicted in stepwise regression analysis. In adults, the resulting model for static judgments (mean static) reveals two predictive criteria: the DTW, classically considered as static information, and the mean velocity along the path, considered as dynamic information. The model for subjective dynamic judgment (mean dynamic) reveals three objective dynamic criterion as predictors, the number of velocity peaks, the mean velocity and the total duration of the path, and one static criteria as predictor, the DTW score. These models raise a number of issues. First of all, the mean velocity, often used in the literature for cinematic analysis, can explain the static judgment. Partial correlation reveals that the more the mean velocity of the production was important, the more the static judgment was low. Several assumptions can be made. (1) The judges have access to the dynamic component of the course by analyzing visual cues present in the trace. This hypothesis is supported by studies which showed activations of motor areas during observation of human movements (Saygin et al., 2004) or during observation of static pattern previously learned to be traced (Longcamp et al., 2003). Theories of motor simulation (Jeannerod, 2009) could also follow our direction. At last, Viviani (2002) assert that access to dynamic properties could occur through a motor representation of the act of writing. (2) In the static presentation modality, the judges have access to the size of the letter; they could deduce the calculation of mean velocity (by divided the distance by the duration). Nevertheless, participants had no access to duration information and the involvement of the distance was not notice in regression model. (3) This third assumption, referred to the close link between the spatial shape and the kinematics of production. In adults writing news characters, higher velocity production could mean poor letter shape. In this case, an obvious correlation would appear between dynamic and static criteria with no need of simulation or knowledge of handwriting motor rules.

Another significant result concerns the link between DTW and both static judgment and dynamic judgement. It seems that this objective measure, generally regarded as a method of static analysis of production, could also be taken into account for kinematics assessments of characters. Partial correlation reveals that the less the distance is important between the standard trajectory and the effective trajectory, the more important are the static or dynamic judgment scores. The link between DTW score and dynamic judgment score is not surprising. Indeed, the application of this algorithm is limited by computational initial condition such as the matching of the starting point and pen-up/pen down sub sequence computation. Thus, order of strokes production (not always respected in children handwriting, seen as a dynamic characteristic, is taken into account in the DTW computation. Same kind of relation between DTW methods and kinematics has been found by Di Brina (2008). We make the assumption that this index could be used as an overall indicator of the quality of handwriting production in is double assessment of spatial characteristics and kinematics features, as already suggested by several researches (Niels, 2004; Di Brina, 2008).

As introduced previously, developmental differences could also influence the choice of character recognition and evaluation criteria. In children, the resulting model for static

judgments (mean static) reveals two predictive criteria: the number of strokes and the distance classically considered as static information, and the designs copying scores used as an indication of children visuospatial and motor skills. Prediction from DTW on static judgment is not shown in children. This could be due to the fact that children do not always produce characters in the correct order. Indeed, DTW score is quite sensible to the order of strokes production but in static presentation this information is not available and could not be used by judges. The model for subjective dynamic judgment (mean dynamic) reveals two objective dynamic criteria as predictors, the number of pauses and the number of pen up, and one static criterion as predictor, the DTW score. As children movements are classically slower than movement that would have been performed by expert scriber, we suggest that pen up duration and number of pauses are probably the more prominent criteria for perceptual judgements. In contrast, in adults, the following objective dynamic predictors, the number of velocity peaks, mean velocity and duration seem to be more distinct indicators for handwriting fluency.

In conclusion, this research highlights the difficulty to classify criteria as clearly static or dynamic. Correlations amongst and between static and dynamic criteria, objective and subjective, are observed in both children and adult populations. These results moderate existing classifications considering criteria as more related to static or more related to dynamic information. Moreover, dynamics criteria considered in children and adults production are not the same and do not have the same meaning, probably due to the differences in motor production skills. Finally, subjective evaluation which can be seen as more relevant in character recognition and evaluation tasks is linked with objective criteria but differs amongst tested populations. In definitive, the best way to perform handwriting character recognition and evaluation would be an evaluation including a large sample of both spatial and dynamic criteria as suggested by human performance classification (taking into account each type of performance criterion of maximum certainty of achieving the goal type, minimum energy expenditure type, and minimum completion time type) in either objective or subjective evaluations.

6. Acknowledgments

The authors gratefully thank all participants (scribers and judges) for their implication in this research. We also thank Sébastien Boisard for the development of the subjective data collection software and his help in data analysis, and Francois Branchon and Heather Larin for their english review.

7. References

- Anderson, S. W., Damasio, A. R. & Damasio, H. (1990). Troubled letters but not numbers: Domain specific cognitive impairments following focal damage in frontal cortex. *Brain*, 113, 749-766.
- Babcock, M. K. & Freyd, J. J. (1988). Perception of Dynamic Information in Static Handwritten Forms. *The American Journal of Psychology*, 101(1), 111-130.
- Bara F. & Gentaz, E. (2010) Apprendre à tracer les lettres : une revue critique. *Psychologie Française*, 55 (2), 129-144

- Barriere, C. & Plamondon, R. (1998). Human identification of letters in mixed-script handwriting: an upper bound on recognition rates, *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, 28(1), 78--81.
- Bartlett R., Wheat J. & Robins M. (2007). Is movement variability important for sports biomechanists? *Sports Biomechanics*, 6(2), 224-243.
- Bartolomeo, P., Bachoud-Lévi, A.-C., Chokron, S. & Degos, J.-D. (2002). Visually- and motor-based knowledge of letters: Evidence from a pure alexic patient. *Neuropsychologia*, 40, 1363-1371.
- Binet, A. & Courtier, J. (1893). Sur la vitesse des gestes graphiques [On the speed of voluntary movements]. *Revue Philosophique*, 35, 664-671.
- Blöte, A. & Hamstra-Bletz, L. (1991). A longitudinal study on the structure of handwriting. *Perceptual and Motor Skills*, 72, 983-994.
- Bluteau, J., Coquillart, S.; Payan, Y. & Gentaz, E. (2008). Haptic guidance improves the visuo-manual tracking of trajectories. *PLoS ONE*, 3(3), e1775.
- Bruinsma, C. & Nieuwenhuis, C. (1991). A new method for the evaluation of handwriting material. In A. M. Wann, Wing & N. Sovik (Eds.), *Development of Graphic Skills* (pp. 41-51). New York: Academic Press.
- Cohen, E. (1991). Understanding handwritten text in a structured environment: determining zip codes from addresses', *International journal of pattern recognition and artificial intelligence*, 1(2), 221--264.
- Di Brina, C. D., Niels, R., Overvelde, A., Levi, G. & Hulstijn, W. (2008). Dynamic time warping: A new method in the study of poor handwriting. *Human Movement Science*, 27(2), 242 - 255.
- Ellis, A. W. & Young, W. (1988). *Human Cognitive Neuropsychologie*. London: Lawrence Erlbaum Associates Publishers.
- Ernst, M. O. & Bühlhoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences*, 8(4), 162-169.
- Flores d'Arcais, G. B. (1994). Order of strokes writing as a cue for retrieval in reading Chinese characters. *European Journal of Cognitive Psychology*, 6, 337-355.
- Franks, I. M., Wilberg, R. B. & Fishburne, G. J. (1982). Consistency and error in motor performance. *Human Movement Science*, 1(2), 109 - 123.
- Fredembach B., Hillairet de Boisferon A., Gentaz E. (2009) Learning of Arbitrary Association between Visual and Auditory Novel Stimuli in Adults: The "Bond Effect" of Haptic Exploration. *PLoS ONE* 4(3): e4844. doi:10.1371/journal.pone.0004844
- Freeman, F. N. (1959). A new handwriting scale. *Elementary School Journal*, 59, 218-221.
- Freyd, J. J. (1983). Representing the dynamics of a static form. *Memory and Cognition*, 11, 342-346.
- Gentile AM. Skill acquisition: action, movement, and neuromotor processes. In: Carr JH, Shepherd RB, editors. *Movement science foundations for physical therapy in Rehabilitation*. 2nd ed. Gaithersburg: Aspen, MD; 2000. p.111-187.
- Goodnow, J. J. & Levine, R. A. (1973). "The grammar of action ": Sequence and syntax in children's copying. *Cognitive Psychology*, 4, 82-98.
- Graham, S., Berninger, V., Abbott, R., Abbott, S. & Whitaker, D. (1997). Role of mechanics in composing of elementary school students: A new methodological approach. *Journal of Educational Psychology*, 89, 170-182.

- Graham, S., Berninger, V. W. & Weintraub, N. (1998). The relationships between handwriting style and speed and legibility. *Journal of Educational Research*, 5, 290-296.
- Guinet, E. & Kandel, S. (2010). Ductus: A software package for the study of handwriting production. *Behavior Research Methods*, 42, 326-332.
- Guthrie, E. (1952). *The psychology of learning*. Peter Smith Publication Inc.
- Hamstra-Bletz, L., DeBie, J. & Den Brinker, B. (1987). *Concise Evaluation Scale for children's handwriting*. Lisse, Swets & Zeitlinger.
- Hamstra-Bletz, L. & Blöte, A. W. (1990). Development of handwriting in primary school: A longitudinal study. *Perceptual and Motor Skills*, 70, 759-770.
- Higgins, J. & Spaeth, R. (1972). The relationship between consistency of movement and environmental conditions. *Quest*, 17, 61-69.
- Hillairet de Boisferon, A., Bara, F., Gentaz, E., & Colé, P. (2007). Préparation à la lecture des jeunes enfants: Effets de l'exploration visuo-haptique des lettres et de la perception visuelle des mouvements d'écriture. *L'Année Psychologique*, 107, 537-564
- Hulme, C. (1981). *Reading retardation and multisensory teaching*. Londres: Routledge & Kegan Paul.
- Jeannerod, M. (2009). *Le cerveau volontaire*. Odile Jacob.
- Korkman, M., Kirk, U. & Kemp, S. A. (1998). *Developmental Neuropsychological Assessment*. The Psychological Corporation: San Antonio.
- Kruskall, J. & Liberman, M. (1983). The symmetric time warping problem: From continuous to discrete. In *Time Warps, String Edits and Macromolecules: The Theory and Practice of Sequence Comparison*, 125-161, Addison- Wesley.
- Longcamp, M., Anton, J. L., Roth, M. & Velay, J. L. (2003). Visual presentation of single letters activates a premotor area involved in writing. *Neuroimage*, 19, 1492-1500.
- Longcamp, M., Zerbato-Poudou, M. T. & Velay, J. L. (2005). The influence of writing practice on letter recognition in preschool children: A comparison between handwriting and typing. *Acta Psychologica*, 119, 67-79.
- Niels, R. (2004). *Dynamic Time Warping: An Intuitive Way of Handwriting Recognition?* Master's thesis, Radboud University Nijmegen.
- Palluel-Germain, R., Bara, F., Hillairet de Boisferon, A., Hennion, B., Gouagout, P. & Gentaz, E. (2007). A visuo-haptic device - Telemaque - increases the kindergarten children's handwriting acquisition. *IEEE WorldHaptics*, 72-77.
- Plamondon, R., Srihari, S., Polytech, E. & Montreal, Q. (2000). Online and off-line handwriting recognition: a comprehensive survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1), 63-84.
- Rosenblum, S., Weiss, P. L. & Parush, S. (2003). Product and Process Evaluation of Handwriting Difficulties. *Educational Psychology Review*, 15(1), 41-81.
- Saygin, A. P., Wilson, S. M., Hagler, J., Bates, E. & Sereno, M. I. (2004). Point-Light Biological Motion Perception Activates Human Premotor Cortex. *Journal of Neuroscience*, 24(27), 6181-6188.
- Schmidt, R. A. & Lee, T. D. (1987). *Motor Control and Learning: A Behavioral Emphasis*. Champaign, IL: Human Kinetics.
- Schmidt, R. A. & Wrisberg, C. A. (2000). *Motor Learning and Performance*. Human Kinetics Publishers.

- Seki, K., Yajima, M. & Sugishita, M. (1995). The efficacy of kinesthetic reading treatment for pure alexia. *Neuropsychologia*, 33, 595-609.
- Shams, L. & Seitz, A. R. (2008). Benefits of multisensory learning. *Trends in Cognitive Sciences*, 12(11), 411-417.
- Thorndike, E. L. (1910). Handwriting. *Teacher College Record*, 11, 83-175.
- Van Galen, G. (1991). Handwriting: Issues for a psychomotor theory. *Human Movement Science*, 10(2), 165-191.
- Viviani, P. & Terzuolo, C. (1982). Trajectory determines movement dynamics. *Neuroscience*, 7(2), 431--437.
- Viviani, P. (2002). *Common Mechanisms in Perception and Action: Attention and Performance*. Oxford University Press, New York, chapter Motor competence in the perception of dynamic events: a tutorial, pp. 406-443.
- Zesiger, P. (1995). *Ecrire : Approche cognitive, neuropsychologique et développementale*. Paris: Presses universitaires de France.



Character Recognition

Edited by Minoru Mori

ISBN 978-953-307-105-3

Hard cover, 188 pages

Publisher Sciyo

Published online 17, August, 2010

Published in print edition August, 2010

Character recognition is one of the pattern recognition technologies that are most widely used in practical applications. This book presents recent advances that are relevant to character recognition, from technical topics such as image processing, feature extraction or classification, to new applications including human-computer interfaces. The goal of this book is to provide a reference source for academic research and for professionals working in the character recognition field.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Anne Hillairet De Boisferon, Jeremy Bluteau and Edouard Gentaz (2010). The Assessment of Spatial Features and Kinematics of Characters: an Analysis of Subjective and Objective Measures, Character Recognition, Minoru Mori (Ed.), ISBN: 978-953-307-105-3, InTech, Available from:
<http://www.intechopen.com/books/character-recognition/the-assesment-of-spatial-features-and-kinematics-of-characters-an-analysis-of-subjective-and-objecti>

INTECH

open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2010 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](#), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.