Effects of Pen Design on Drawing and Writing Performance

Ravindra S. Goonetilleke, Errol R. Hoffmann and Ameersing Luximon

Human Performance Laboratory
Department of Industrial Engineering and Logistics Management
Hong Kong University of Science and Technology
Clear Water Bay, Kowloon, Hong Kong
Abstract

Two experiments are reported with the aim of determining the effect of pen shape and size on two different types of task: drawing and writing. Experiment 1 attempted to determine the optimum shape and size of shank for ball-point pens used to perform an accurate drawing task. Twenty-seven participants used a total of nine different pens. Drawing performance was measured by having the subjects follow mazes of different size. The dependent variables were drawing movement time and drawing accuracy. The results indicate that the pen with an equivalent diameter of 8 mm had the best accuracy during drawing, even though it had the lowest speed. The results confirm the applicability and validity of the Drury tracking model for such a task. Even though users tended to prefer larger pens, their accuracy with such pens tended to be low.

In a second experiment subjects performed a writing task (with no accuracy constraints) and rated their preference for types of pen. An attempt was made to determine the preferred dimensions of ball-point pens for writing in Chinese and English. A total of 36 “bare-bodied” pens and 20 Chinese subjects were used to evaluate the effects of shape, size and weight on time to write a sentence. The Writing ability, Comfort and an Overall rating were given by the subjects. The results indicate that the Chinese subjects preferred a circular shaped pen for writing in both languages. A factor analysis showed that speed, comfort and writing ability are independent parameters in pen evaluations.

Keywords: Pen, pen size, pen cross-section, design, shank-design, human-performance, writing, implements
1. INTRODUCTION

The design of handwriting instruments has been based primarily on touch, feel, aesthetics, and muscle exertion (Kao, 1977, 2004), and has mostly been related to writing instruments that are commercially available. Kao (1976) has investigated the effect on hand-writing quality of lead pencils, ball-point pens and fountain pens. In two other studies Kao (1977, 1979), investigated the effects on writing time, writing pressure, and writing efficiency of ball-point pens, pencils, felt pens, fountain pens and of pen point shape variations. The primary conclusions of these studies were that:

1. Ball-point pens yield the fastest and fountain pens the slowest writing speed.
2. Writing pressure was highest with ball point pens and lowest with felt-tip pens.
3. In terms of the tip design, the pen contact time was lowest for tilted points followed by straight points and then curved points.

Thus, it is clear that different pen characteristics have to be considered along with hand-instrument interaction in the design of writing instruments.

The “dynamic-tripod” is one of the most common ways to hold a writing instrument (Wynn-Parry, 1966; Wu and Luo, 2006c). This is where the thumb, index finger and middle finger grasp the writing instrument so that so they function together. Rosenbloom and Horton (1971) found that such a grip requires fine motor coordination. Callewaert (1963) described a grip where the writing instrument is held between the index and middle fingers with the wrist more canted. He claimed that such a grip was superior as the muscles would be more relaxed. Some researchers (Sassoon et al., 1986; Jacobsen and Sperling, 1976; Ziviani and Elkins, 1986) have shown that pen grip does not influence the speed of writing.
Handwriting performance however, appears to be directly affected by finger pressure on the writing instrument, point pressure of the writing implement on the writing surface and the pressure of the hand on the writing surface (Bailey, 1988). High point pressure is related to high grip pressures (Herrick and Otto, 1961) even though the amount of finger flexion and the pressure on the index finger when writing tends to reduce in children with increasing age (Ziviani, 1982). Repetitive stress injuries are well known to relate to the risk factors of force, frequency and posture (Putz-Anderson, 2006; Drury, 1987). Hence it is not surprising that some studies have been primarily driven by the need to better understand disorders such as mogigraphia, a special case of tenosynovitis better known as writer’s cramp (Chakarov et al., 2006; Udo et al., 2000). Researchers such as Udo et al. (2000) have attributed writer’s cramp to the increased muscle force in the forearm needed to hold the pen without slipping and to generate the force at the pen tip required to write. The authors of that paper propose that increasing the gripping area, increasing the pen diameter in the gripping area and/or increasing the friction coefficient between the pen and the hand can reduce the muscle load. But Gross et al. (1996) have shown that smaller-handed writers prefer smaller pen sizes. They also showed that those with larger hands generate less thenar EMG (electromyograph) activity when writing. Relaxation training can improve writing performance, as it helps reduce muscular tension while writing (Jackson et al., 1980; Carter and Synolds, 1974). The muscular stress can also be reduced through modifications to the writing instrument.

Alston (1986) found that children preferred a conventional six-sided pencil to a triangular cross-sectioned pencil primarily because they were more used it. Based on Alston’s findings, Bailey (1988) pointed out that further investigation may be necessary to investigate the performance effects of varying the cross-sections of writing instruments. The effect of shank design on writing performance is important if pen-point pressure is a factor
related to writing quality. The pressure exerted on the writing surface depends, to a large extent, on the pinch grasp used by a subject. Even though hand grip strength variations with grip size have been extensively investigated (Greenberg and Chaffin, 1977; Ayoub and Lo Presti, 1971), the effect of pinch grasp force on writing pressure has yet to be determined. Kao (1979) has mentioned that pen shafts with larger diameters tend to be preferred. Udo et al. (2000) have reported that gripping pressure can be reduced by using a pen 12 to 14 mm in diameter. Gross et al. (1996) concluded that pen preference is related to a pen’s texture, balance, grip diameter and length, and that ease of writing measured in terms of ink flow also appears to contribute to pen preference. Wu and Luo (2006b) have mentioned that the diameter, length and shape of a touch-pen (stylus pen) affect handwriting performance and efficiency. Other parameters that have been shown to have an effect on performance and comfort are the taper on the shank and the friction coefficient of the material used in the shank to reduce grip force (Udo et al., 2000).

Most previous studies evaluating pens have had confounding effects. Hence this study focused primarily on evaluating shank shape, shank diameter, pen weight and pen performance in terms of writing speed, errors and subjective preferences.

Handwriting in most languages is a result of two primary movements, one horizontal from wrist motion and one vertical from finger movements, with the addition of a rightward translation of the whole arm (Hollerbach, 1981; Sallgoity et al., 2004; Singer and Tishby, 1994). Drawing is similar except that when drawing, horizontal and vertical hand motion is driven primarily by arm translation. In terms of motion, this makes drawing and handwriting somewhat different (Wu and Luo, 2006a). The primary goal of this research was to examine the effect of shank design (shape and size) and pen mass on “accuracy”, speed of drawing
and on preferences when writing in English and Chinese. At the same time, Drury’s (1971) model of constrained paths was applied in an attempt to model the movement time and errors when using different types of pen. Experiment 1 related to drawing through a maze in the x- and y-directions, while experiment 2 evaluated writing in English and Chinese. All pens in the experiments were longer than 115 mm so that the pen extended beyond the hand, as the 50th/95th percentile hand width of the Hong Kong population is 80/90 mm for males and 70/80 mm for females (Pheasant, 1988).

2. EXPERIMENT 1

2.1 Method

2.1.1 Pen Design

A total of nine pens were fabricated. These pens had three different cross sections: circular, elliptical and hexagonal resembling common drawing instruments. Each of the shapes was made in three sizes: large (cross sectional area of 320 mm²), medium (cross sectional area of 180 mm²), and small (cross sectional area of 50 mm²). The equivalent circular diameters were 20, 15 and 8 mm, respectively. These pen sizes were chosen to cover the range of sizes in common use. In addition the range of sizes was chosen so that the results could be mathematically modelled to cover a wider range of pen sizes. Wu and Luo (2006b) had previously found that a diameter of 8 mm is the most suitable for stylus-pens and hence the other sizes were larger than 8 mm, especially since Peck et al. (1980) found that 7.9 mm diameter pencils were the most suitable for nursery school students. Identical ink inserts with a tip size of 0.5 mm was used in each pen. The importance of this size will be seen in later analysis.
2.1.2 Participants

Twenty seven undergraduate students at the Hong Kong University of Science and Technology served as subjects in the experiment. Two of the subjects were left-handed and all others were right-handed. The finger dimensions of the participants are given in Table 1.

| Insert Table 1 about here |

2.1.3. Maze tracing task

The Archimedes spiral has been commonly used for testing tremor (Cooper et al. 2000, Lakie et al., 1994), and hence would be suitable for determining accuracy in tracing a path. Wu and Luo (2006a) used a circle 90 mm in diameter and a square with sides of 90 mm, each having a track 3 mm wide for their drawing task. In this experiment, three different maze sizes were used to test for any relationship between maze size and pen performance in terms of drawing accuracy and speed. The mazes are illustrated in Figure 1. The track widths of the three mazes were 6, 3 and 1.5 mm, in order of increasing accuracy requirements for path control (Table 2). The largest maze measured 52.5 mm across its width; the second and third mazes were one-half and one-quarter the size of the first maze. The path length traversed at the maze centerline was 404 mm for the largest maze, 202 mm for the middle-sized and 101 mm for the smallest maze. Hence the largest to smallest mazes are designated “4”, “2” and “1”.

The mazes required path movement in two perpendicular directions, each with relatively tight path constraints due to the limited width. This form of task thus required the use of different muscle groups used in accurate drawing. The structure of the maze was simple as the task simply required right-angle turns at each maze corner. That is, there was no requirement for the subjects to solve problems of required motion direction.
2.1.4 Procedure

The order of pen presentation to each subject was completely randomized, with each subject testing all of the nine available pens. To minimize any order effects, the three maze sizes were grouped on three sheets which had the mazes in a pre-determined order as follows: group 1: large, medium, small; group 2: medium, small, large; group 3: small, large, medium. The 27 subjects were randomly assigned to one of these three groups to eliminate any maze order effect.

Each subject performed three trials with each size of maze. So each subject used each pen to draw on a total of 9 mazes (3 different sizes of maze with three repetitions of each size). Subjects were instructed to draw as fast as possible but to maintain accuracy.

Movement time was measured with a stopwatch.

2.2 Results

The two dependent measures were the time taken to complete the maze and the drawing accuracy. Movement time was the time to complete the maze moving from the outside entry point, back to the outside through the center of the maze. Accuracy was defined as the number of cross-over points with the printed lines of the maze; a touch was considered as one error while a cross-over was considered to be two errors.
A within-subject ANOVA was performed on the movement time data. This yielded a significant main effect of Maze size \( F(2,52) = 33.97; p < .001 \). Pen size \( F(2,52) = 2.87; p = 0.066 \) and pen shape were not significant, and there were no significant interactions among the maze size, pen size and pen shape variables. The mean movement time data are given in Table 2 and plotted in Figure 2. Here the ‘4’ refers to the largest maze, ‘2’ to the middle-sized and ‘1’ to the smallest maze. Post-hoc tests using the Newman-Keuls procedure showed that all comparisons between maze sizes were significantly different (the smallest had longer movement times than the larger two \( p < .01 \) and the larger two being different at \( p < 0.05 \). Figure 2 shows clearly that the pen size effect was small. The times for the 50 and 180 mm\(^2\) cross-sectional area pens differed by less than 0.1 seconds, while the 320 mm\(^2\) pens had a movement time about 0.3 seconds longer (which did not quite reach statistical significance).

The ANOVA showed a highly significant effect of maze size on the movement time. This was not expected, as Drury’s tracking model for movement times between constraining paths would predict that the movement time in each task would be the same (at a given pen size). This is because the \((A/W)\) ratio was the same for each of the mazes, where \( A \) is the distance to be moved and \( W \) is the track width. Drury’s model is (Drury, 1971; Drury et al., 1987)

\[
MT = a + b (A/W) \text{ or } V = c + d W, \text{ where } V \text{ is the speed of movement.}
\]

Further analysis of the effect of maze size was clearly required, as this is the factor that determined the speed and accuracy of control with the different pens.
The effect of the maze size seen in the ANOVA can be analyzed by using an ‘effective track width’ \( W_{\text{eff}} \) – in this case by subtracting 0.5 mm from each track width to obtain the actual tolerance, since the width of the pen tip was 0.5 mm. That is, \( W_{\text{eff}} = W - 0.5 \). This ‘effective track width’ is given in Table 2. The task was essentially a series of Drury-type tracking tasks where there were changes in direction as the maze was traversed. The total length traversed in each maze was 404, 202 and 101 mm, respectively, for the largest to the smallest maze. For purposes of this analysis, only the amplitude \( A \) of the first leg of the maze will be used – the total lengths are all proportional to this length. The task was not a continuous tracking task, but tended to be performed as a series of discrete movements that ended at each 90 degree bend in the maze. Each leg was uni-directional. Thus the total amplitude is not relevant if this task is considered as the sum of a number of constrained movements, each of which is a Drury-type task. These values are given in Table 2. A regression in the form of the Drury model,

\[
\begin{align*}
\text{MT}(320) &= 3.06 + 0.300 \frac{A}{W_{\text{eff}}}; \quad r^2 = 0.99 \\
\text{MT}(180) &= 2.84 + 0.289 \frac{A}{W_{\text{eff}}}; \quad r^2 = 0.99 \\
\text{MT}(50) &= 2.75 + 0.293 \frac{A}{W_{\text{eff}}}; \quad r^2 = 0.99
\end{align*}
\]

where the 320, 180 and 50 are the cross-sectional areas of the pens. The maze size effect is thus completely accounted for by Drury’s tracking model and the use of an effective path width. The data are plotted in Figure 3. The Drury model thus shows a separate effect of pen design on path speed, that of the size of the pen tip when accurate control is required.

Insert Figure 3 about here
2.2.2 Errors in traversing the maze

The errors were analyzed using a within-subject ANOVA. There were significant main effects of maze size \[F(2,52) = 208.7; \ p < .0001\] and pen size \[F(2,52) = 14.59; \ p < .0001\]. There was also a significant interaction between maze size and pen size \[F(4,104) = 12.67; \ p < .0001\]. Newman-Keuls post-hoc tests showed that, for each pen size, all maze sizes were significantly different, with errors increasing significantly with decreasing maze size (\(p < 0.01\)). The significant interaction is better shown by breaking down the data into the different maze sizes. At the largest maze size, there were no differences in errors with any of the pen sizes. For the middle-sized maze, the smallest pen had significantly fewer errors than the largest pen (\(p < 0.05\)). However with the smallest maze, all comparisons between pen sizes were significant (\(p < 0.01\)). The data are illustrated in Figure 4 as dependent on the ratio, A/W, with the errors increasing with pen size.

---

Insert Figure 4 about here

---

Errors for the individual pen sizes are given by,

\[
\text{Errors}(320) = -8.97 + 1.32 \frac{A}{W_{\text{eff}}}; \quad r^2 = 0.99
\]  
\[\text{Errors}(180) = -6.27 + 0.976 \frac{A}{W_{\text{eff}}}; \quad r^2 = 0.99
\]  
\[\text{Errors}(50) = -3.47 + 0.629 \frac{A}{W_{\text{eff}}}; \quad r^2 = 0.98
\]

Including the interaction between the pen and maze size variables, the data is well represented by,
Errors = - 6.24 + 0.800 (A/W_{eff}) + 0.0122 (A/W_{eff})D_{equiv} ; R^2 = 0.935 \hspace{1cm} (7)

where the equivalent diameter (D_{equiv}) is the circular diameter having the same cross-sectional area as the experimental pens.

As the mazes differed in the length traversed, it may be better to look at the error rate in terms of the distance moved, rather than in terms of the A/W ratio. Figure 5 is a result of this transformation.

Stepwise regression analysis of this data showed that the dominant term describing the error/distance relationship was the reciprocal of the square of effective target width (1/W_{eff})^2. This was the case for all three pen sizes. In each case, the intercept was non-significant, so the regressions were forced through the origin to yield,

\[
\begin{align*}
\text{Error/distance}(320) &= 0.0667 (1/W_{eff})^2 \hspace{1cm} (8) \\
\text{Error/distance}(180) &= 0.0528 (1/W_{eff})^2 \hspace{1cm} (9) \\
\text{Error/distance}(50) &= 0.0395 (1/W_{eff})^2 \hspace{1cm} (10)
\end{align*}
\]

The gradient is strongly linearly related to the size of the pen (using the equivalent diameter D_{eq}) by Gradient = 0.0212 + 0.00220 D_{eq}; \ r^2 = 0.98, so that the error is well described by

\[
\text{Error/ Unit distance} = (0.0212 + 0.00220 D_{eq}) / (W_{eff})^2 \hspace{1cm} (11)
\]

This equation describes the interaction between path width and pen size found in the analysis of variance. Note that, as was found in the experiments of Goonetilleke and Hoffmann (2008), the errors are related to the inverse of the squared effective track width. Goonetilleke
and Hoffmann have suggested that the errors depend on hand tremor while traversing such a constrained path.

2.3 Discussion of Experiment 1

In these experiments, the three mazes were geometrically similar and were simply scaled in the ratio of 1:2:4 from the smallest to the largest. Thus, for all mazes, the ratio of the path width to the length of the path traversed was identical. The expectation was that, according to Drury’s law for tracking within a constrained path, the movement times should be the same. Such an expectation does not take into account the ‘effective track width’ which results from the pen having a finite tip size. At the smaller track widths, the proportion of the total track width taken up by the pen tip was relatively large, but this proportion decreased as the track width increased. Thus it is necessary to use a modification of the Drury relationship in which the actual track width is replaced by the true tolerance available to the subject. This ‘effective track width’ has been used in a number of different applications of Drury’s model such as steering control of motor vehicles (DeFazio et al., 1992) and tracking through parallel constraining paths (Hoffman, unpublished data). The results of this experiment show that for an understanding of pen performance when tracking through a constrained path, it is necessary to make this pen tip correction. The resulting regressions in terms of Drury’s law then show the classical form for this type of task.

The interaction between pen size (defined by the equivalent diameter) and the track path width is complex, but it is well described by the relationship

\[
\text{Error/ Unit distance} = \frac{(0.0212 + 0.00220 \ D_{eq})}{(W_{eff})^2}
\]  

(11)
This equation involved a term that is proportional to the reciprocal of the square of effective track width and a second term that has an interaction between this term and the pen’s equivalent diameter. The constant 0.0212 was in fact non-significant, but was retained for better prediction; a regression forced through the origin gave a coefficient for the interaction term of 0.00351. At constant track width (accuracy), the errors were linearly related to pen size. This suggests that some obscuring of the track may occur and that this increases with pen size, so the errors increase. The increase in errors is unlikely to be due to problems of manipulating the pen, as it would be expected that the grip and control of the pen would be better at the larger sizes (Kao, 1979; Udo et al., 2000). As Goonetilleke and Hoffmann (2008) have suggested, the denominator of the error relationship may arise from tremor in the hand and wrist controlling the motion of the pen. Here the level of tremor should have remained approximately constant, so the error increases may have arisen from the variations in the width of the maze track.

Of major interest is that the tracking performance (at least in terms of errors) was strongly dependent on the size of the pen shank, with errors increasing with pen size. Although the increase was not statistically significant, the movement time also increased with increasing pen size. These results contradict those of Wu and Luo (2006b) who found that drawing time and errors were significantly lower for 11 and 15 mm diameter pens when compared to 5.5 mm diameter touch-pens. A possible reason for this difference is that Wu and Luo adopted a fixed track width of 3mm in their experiment whereas it is clear that track width and pen size have a complex interaction, as shown in equation (11). Kao (1977, 1979) reported that ballpoint pens deliver the fastest times, and this may be why differences in the speeds observed in this experiment were too small to be significant. The smallest pen (8mm equivalent diameter) yielded the smallest number of errors and was also the slowest in
performing the task. In addition, the results do show a speed-accuracy tradeoff. But, as shown by researchers such as Udo et al. (2000), smaller grips tend to impose a larger muscle load, so smaller grips may possibly be related to better control. Thus, the smaller pens may have reduced errors. There is an apparent linear relationship between all the error rates and the corresponding movement times (Figure 6). At the lower end of this relationship is the smallest pen with the widest path, and at the upper end is the largest pen with the narrowest path. There is thus a trend for the errors to increase as the task difficulty (as measured by the movement times) increases. Within Figure 6 there are three sets of data points for each maze size. For the largest and middle-size mazes, there is little variation of the error rate with movement time; with the smallest maze there is considerable variation (top three points in Figure 6).

3. EXPERIMENT 2

Whereas Experiment 1 required accurate control in a drawing task, this experiment was aimed at determining the optimum form of pen for freehand writing, and thus a wider range of pens was tested. Wu and Luo (2006b) concluded that writing quality depends on a pen’s diameter rather than its length. The hypotheses investigated in this study were whether the cross-sectional shape of a pen, its cross-sectional size and its weight affect writing time and subjective ratings of writing ability, comfort and overall perceptions.

3.1 Method
3.1.1 Equipment

Thirty six pens were fabricated from three sizes of circular aluminum stock. Six different shapes of pens were made: circular (c), hexagonal (h), octagonal (o), triangular (t), elliptical (e) and square (s). The triangular shank was chosen as it can assist with the tripod grip commonly adopted by many (Wynn-Parry, 1966), even though Alston (1986) found no significant differences between triangular and conventional hexagonal pencils. Each of these shapes was made in three sizes (Table 3): small (cut from 7.5mm diameter stock), medium (cut from 9 mm diameter stock) and large (cut from 10.5mm diameter stock) especially since Wu and Luo (2006b) found that touch-pens 11 and 8 mm in diameter generated fewer writing errors than those 15 and 5.5 mm in diameter. The design of each pen was such that the middle portion of the pen shank could be fitted with one of two different (light and heavy) weighted sections (Table 4, Figure 8). Hence only eighteen of the pens are shown in Figure 7.

3.1.2 Participants

Twenty undergraduate and postgraduate Chinese students (16 from Hong Kong and 4 from the People’s Republic of China) from the Hong Kong University of Science and Technology participated in the experiment. All subjects received a 5% credit in a related course for their participation.

3.1.3 Procedure
The order of pen presentation was completely randomized but counterbalanced to reduce order effects. Each subject was asked to write one English sentence, “This is the Department of Industrial Engineering and Engineering Management” and one Chinese sentence, “歡迎您到香港來!” using each of the 36 pens. Copying texts has been common when evaluating handwriting (e.g., Gross et al., 1996). Even though brief texts are not considered suitable for studies that evaluate EMG (Udo et al., 2000), studies of a short duration ought to be appropriate for evaluating speed and subjective impressions of a pen design.

Four dependent variables were used: (i) Time to complete writing the sentence, as measured with a stop-watch; and subjective evaluations of (ii) Overall Comfort, as experienced by gripping the pen during the task -- subjects generally commented on the security of the grip and any discomfort produced by the edges of the pen in the grip area; (iii) Writing Ability, which subjects described as largely an issue of control of the pen during the writing task; and (iv) an Overall Evaluation, making an attempt to include all factors in an omnibus evaluation. The subjective evaluations were rated for each pen and writing language using the following 7-point scale.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>very dissatisfied</td>
<td>quite dissatisfied</td>
<td>neutral</td>
<td>satisfied</td>
<td>quite satisfied</td>
<td>very satisfied</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Speed has always been used as a performance measure in relation to writing (Ayres, 1912; Burt, 1921; Freeman, 1959; Grof, 1961; Gross et al., 1996; Peverly, 2006). Legibility was not assessed, as there does not seem to be any consensus about its evaluation as researchers such as Ayres (1912), Alston and Taylor (1984), Burt (1921), Freeman (1959), Feldt (1962), Gipps (1984), Helwig et al. (1976), Rarick and Harris (1963), Rubin and Henderson (1982),
Ziviani (1982) and Ziviani and Elkins (1984) have all proposed different ways to evaluate legibility.

3.2 Results and Analysis

The Overall Rating was significant correlated \( (p \leq 0.001) \) with Writing Comfort \( (R=0.82) \) and Writing Ability \( (R=0.82) \). The relationship between Writing Comfort and Writing Ability was also significant \( (R=0.60) \). There was no significant correlation between Writing Time and any of the subjective ratings \( (r^2 < 0.1) \). A regression analysis showed that the overall rating was a composite of writing comfort and writing ability, as shown below.

\[
\text{Overall Evaluation} = 0.052 + 0.463(\text{Overall Comfort}) + 0.516(\text{Writing Ability}) \quad R^2 = 0.84
\]

A factor analysis with varimax rotation showed the emergence of three dominant factors (Table 5). The first factor was heavily weighted by the two measures, \textit{Overall Comfort} and \textit{Overall Evaluation}. The second factor related to \textit{Writing Ability} and \textit{Overall Evaluation}. The third factor was almost solely dominated by the \textit{Time} taken for writing. The three factors explained 94.3\% of the total variation. Since \textit{Overall Evaluation} was strongly dependent on \textit{Overall Comfort} and \textit{Writing Ability}, it appears that \textit{Overall Comfort}, \textit{Writing Ability} and speed or \textit{Time} were three distinct constructs in the evaluation of a pen in these writing tasks.

Analyses of variance were carried out with the (a) Comfort ratings (b) Writing ability (c) Overall ratings and (d) Writing times for given passages of text. These analyses were done with (i) Language (English and Chinese), (ii) Type of pen (bh, bl, mh, ml, sh, sl, where b,m,s refer to big, medium and small and h,l refer to heavy and light) and (iii) Shape of the pen grip (circular, elliptical, hexagonal, octagonal, square and triangular) as independent factors.
Each grip design was machined from the same size stock, so the b,m,s refer to the stock size used for machining. As there was considerable variation in the mass of the pens with the size, so the six combinations were analyzed as six levels of the independent variable ‘type’. Table 4 shows that there was little variation of pen mass within a mass category, and hence it was considered adequate to use these categories for the analyses. In each case, a within-subjects design was used, with subjects as a random factor. In the case of writing time, it was necessary to perform separate analyses for the two languages, as the text to be written was different in each.

3.2.1 ANOVA of pen comfort ratings

There was no significant main effect of language. The significant main effects were Pen Type [F(5,95)=6.09; p< 0.0001] and Pen Shape [F(5,95)=32.86; p< 0.0001]. There was also an interaction between Pen Type and Language [F(5,95)=3.06; p < 0.05]. Newman-Keuls post-hoc tests showed a ranking on shape of triangular, square, elliptical, hexagonal, octagonal, circular in order of comfort ratings. The e, h and o shapes were not significantly different but all the other comparisons were significant at p< 0.01. The notable feature here is that, as the shape became more circular, the comfort rating improved consistently. This is simply explainable in terms of the decrease in sharp edges as the shape becomes more circular. A grip used one experimental subject is shown in Figure 9.

-----------------------------------------------------
Insert Figure 9 about here
-----------------------------------------------------
The interaction between Type and Language was more complex. In writing English, the ‘bh’ pen had a significantly lower comfort rating than all the others, which were not significantly different from each other. The heavy large pen was the most uncomfortable to use. In writing Chinese, the ‘bh’ pen was again the least comfortable, but the groupings were somewhat different. The ranking was bh, bl, sh, mh, sl, ml. The bh and bl were not significantly different from each other, as were the groups bl, sh, mh and sh, mh, sl, ml. There was thus greater discrimination among pen types in writing Chinese than in writing English, although the groupings, when taking into account the non-significant differences, are essentially similar. Large and heavy pens were less comfortable than the smaller pens, which contradicts the findings of Udo et al. (2000), who suggested that muscle load will be less with larger pens.

Alternatively, the interaction can be viewed by breaking down the data into Types for each of the languages. Here, for two pen types bl and sl, the rating of comfort for writing in Chinese was lower than that for writing in English.

3.2.2 ANOVA of pen writing ability

There were significant main effects of Type \( [F(5,95)=4.42; p < 0.01] \) and Shape \( [F(5,95)=14.5; p<0.0001] \). There was no effect of language and there were no significant interactions. Newman-Keuls post-hoc tests showed that the ‘bh’ pen had the lowest rating; all the others were not significantly different from each other. This is similar to the result for the comfort rating.
The Shape factors were ranked t,s,e,o,h,c. Again, as the shape approached circular, the writing ability rating improved. The worst cases were those with sharp edges, the triangular and square grips. The groupings which were not significant in this case were t,s; s,e and o,h. The other comparisons were different at the p \leq 0.05 level or less.

3.2.3 ANOVA of Overall ratings

There were significant main effects of Type [F(5,95)=5.50; p< 0.001] and Shape [F(5,95)=26.37; p< 0.0001]. The pattern of significant differences shown by Newman-Keuls tests was similar to that for the Comfort and Writing Ability ratings. The ‘bh’ pen had the lowest overall rating, but the others were not significantly different when Types were compared. Again, the ratings for the Shape increased as the pen shape approached a circle. The rankings in this case were t,s,e,h,o,c. There was no significant difference between s and e and between h and o.

Subjective ratings for Comfort, Writing Ability and Overall are shown in Figures 8 and 9 for the English and Chinese languages, respectively.

----------------------------------
Insert Figures 10 and 11 about here
----------------------------------

3.2.4 ANOVA of Writing Times

Separate analyses were carried out for the English and Chinese languages due to the different text in each case. In neither case was there a significant effect of Shape or Type of pen, nor were there any significant interactions.
3.3 Discussion of Experiment 2

With this experimental population there was little effect of the language in which the writing was being performed. There were no main effects of language in any of the subjective measures and only an interaction with the Type of pen in the Comfort ratings. Thus, the design of pen should be the same, independent of the intended language used.

There was also no effect of pen design on the time taken for writing the given sentences. At least for this short-term task, even the more uncomfortable pens did not slow performance. This is in agreement with most previous studies where writing time or speed has consistently failed to show any differences among pens or pencils (Alston 1986; Gross et al. 1996). Udo et al. (2000) in their experiment with a conventional pen and a newly designed pen did not see any differences in the number of letters written within a one hour period even though they recorded differences in the EMG activity of the flexor pollicis brevis muscle. There may of course be effects in more long-term tasks. For example, the heavy-large pen, which was found to be the least comfortable, may in the long term produce poorer performance.

Thus, selecting a pen for writing may be based purely on subjective preferences. In these trials, the circular pen was always the pen of choice. Apart from the large-heavy pen, there was little difference in the ratings of the pens on any of the three subjective rating scales. These results are in contrast to those for the accuracy task of Experiment 1, where there was a strong impact of pen size on drawing performance.
4. CONCLUSIONS

4.1 Accuracy

(i) Although the pen shank shape did not significantly affect either movement time or error scores, the minor differences that did occur showed that the hexagonal cross-section was the worst for accuracy, when compared to the elliptical and circular cross-sections.

(ii) The smaller pen (8mm equivalent diameter) had the lowest error rate measured by the number of excursions outside the marked track per unit length of track. Errors increased with increasing pen size.

(iii) The time for performing the drawing task increased as the pen size decreased. Larger pens were faster in this task.

4.2 Writing

(i) Pen grips should be circular or close to circular. Pens that had sharp edges (triangular and square) gave poor subjective ratings of Comfort and Writing Ability. Subjective ratings improved as the pen shape approached a circular cross-section.

(ii) Large diameter, heavy pens had low Comfort and Writing Ability ratings.

(iii) Writing speed was not affected by pen design in the short-term tests used here, but it may be affected when the pen is used over longer time periods. This possibility needs further investigation.

Most daily writing requires speed rather than accuracy, and in fact may not require any large degree of accuracy in control. It appears from these data that the selection of a pen based on speed is not important, but can be important when accuracy in control is required.
(as in the drawing task of Experiment 1). The preferences seen in previous studies (Kao, 1979) are consistent with the lack of need for close control and the desire for speed. An interesting experiment would be to get subjects to write while controlling the amplitude of the script to within tight constraints. In such circumstances smaller pen shanks may show better accuracy, although at lower writing speed.

It is in tasks where there is a need for accuracy, such as engineering drawing, that the smaller shank sizes may be superior. Hence, in conclusion, it may be said that a specification based on a set of standardized tasks may be helpful in writing instrument selection. Such a specification could be valuable to artists and designers who need a specified level of accuracy during tasks that they perform often. All the pens tested in this study were bare-bodied with hard surfaces to eliminate any influence of the material. However, a soft grip surface may influence performance. Hence pen surface texture effects, not studied in these experiments, is an area that may need further study.

5. REFERENCES


Gross, C. M., Lloyd, J. D., and Tabler, R. E., 1996. Ergonomic analysis of pen comfort and
wrist dynamics while writing. Unpublished report, Center for Product Ergonomics,
College of Public Health, University of South Florida.

Publishing.

Helwig, J.J., Johns, J.C., Norman, E. E., Cooper, J. O., 1976. The measurement of manuscript

Herrick, V. E. and Otto, W., 1961. Pressure on point and barrel of a writing instrument.

Hollerbach, J.M., 1981. An oscillation theory of handwriting. Biological Cybernetics, 39,
139-156.

cursive handwriting of fourth-grade students. Perceptual and Motor Skills, 51, 1215-
1221.

Medicine, 198, 395-398.


52.


TABLES AND FIGURES

List of Tables

Table 1. Finger dimensions of Experiment 1 participants
Table 2. Geometry of the three mazes used in the pen trials, along with movement times (MT) in seconds and errors.
Table 3. Cross-sectional area (mm²) of different shape and size pens
Table 4. Weight (grams) for the different size and shape of pens
Table 5. Factor Analysis with varimax rotation of the variables Writing Comfort, Writing Ability, Overall Rating, and Writing Time. The factor loadings greater than 0.5 are shown in bold.
List of Figures

Figure 1. The three mazes used for drawing. They are geometrically similar: Maze (b) is half the size of Maze (a); Maze (c) is one-quarter the size of Maze (a).

Figure 2. Effect of pen size (Equivalent Diameter) and Maze Size [large (4), Medium (2), and Small (1)] on movement time through the maze.

Figure 3a. Drury’s model for tracking used to explain the effects of a finite pen-tip width on time to move through mazes of varying sizes.

Figure 3b. Same data as in Figure 3a, except plotted in the speed form of Drury’s model.

Figure 4. Tracking errors as a function of maze and pen sizes, showing the correlation in terms of the Drury tracking model and the significant interaction between maze size and pen size. Numbers 320,180, 50 refer to the cross-sectional areas of the pen shafts in mm$^2$.

Figure 5. Tracking errors per unit distance traveled as dependent on the effective path width and the pen size. The numbers 320,180, 50 refer to the cross-sectional area of the pens in mm$^2$.

Figure 6. Errors per unit distance as dependent on movement time for all experimental conditions.

Figure 7. (i) Cross-sectional shapes of the pens used in Experiment 2: Circular (c), Elliptical (e), Hexagonal (h), Octagonal (o), Square (s) and Triangular (t). (ii) Eighteen of the thirty-six pens used in the experiment. The other eighteen differed only in the weight of the middle section.

Figure 8. (a) The pen with the circular section. (b) The three pieces that make up the pen.

Figure 9. An example showing the grip of one particular subject when writing

Figure 10. Subjective ratings when writing in Chinese.

Figure 11. Subjective ratings when writing in English.
Table 1. Finger dimensions of Experiment 1 participants

<table>
<thead>
<tr>
<th>Subject No.</th>
<th>Thumb Width (mm)</th>
<th>Index Width (mm)</th>
<th>Middle Width (mm)</th>
<th>Handedness</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20.42</td>
<td>15.34</td>
<td>15.68</td>
<td>R</td>
</tr>
<tr>
<td>2</td>
<td>21.04</td>
<td>15.64</td>
<td>16.09</td>
<td>R</td>
</tr>
<tr>
<td>3</td>
<td>17.23</td>
<td>15.07</td>
<td>15.08</td>
<td>R</td>
</tr>
<tr>
<td>4</td>
<td>19.32</td>
<td>15.57</td>
<td>15.94</td>
<td>L</td>
</tr>
<tr>
<td>5</td>
<td>22.04</td>
<td>17.23</td>
<td>18.54</td>
<td>R</td>
</tr>
<tr>
<td>6</td>
<td>21.34</td>
<td>15.86</td>
<td>16.28</td>
<td>R</td>
</tr>
<tr>
<td>7</td>
<td>20.37</td>
<td>15.21</td>
<td>15.94</td>
<td>R</td>
</tr>
<tr>
<td>8</td>
<td>21.84</td>
<td>15.47</td>
<td>15.74</td>
<td>R</td>
</tr>
<tr>
<td>9</td>
<td>20.79</td>
<td>15.65</td>
<td>15.84</td>
<td>R</td>
</tr>
<tr>
<td>10</td>
<td>21.31</td>
<td>15.61</td>
<td>16.27</td>
<td>R</td>
</tr>
<tr>
<td>11</td>
<td>21.24</td>
<td>15.24</td>
<td>15.77</td>
<td>R</td>
</tr>
<tr>
<td>12</td>
<td>21.61</td>
<td>15.68</td>
<td>16.81</td>
<td>R</td>
</tr>
<tr>
<td>13</td>
<td>19.46</td>
<td>15.13</td>
<td>15.84</td>
<td>R</td>
</tr>
<tr>
<td>14</td>
<td>21.25</td>
<td>15.27</td>
<td>15.99</td>
<td>R</td>
</tr>
<tr>
<td>15</td>
<td>21.61</td>
<td>15.91</td>
<td>16.11</td>
<td>R</td>
</tr>
<tr>
<td>16</td>
<td>18.54</td>
<td>15.16</td>
<td>15.63</td>
<td>R</td>
</tr>
<tr>
<td>17</td>
<td>19.97</td>
<td>15.05</td>
<td>15.71</td>
<td>R</td>
</tr>
<tr>
<td>18</td>
<td>20.45</td>
<td>15.34</td>
<td>16.00</td>
<td>L</td>
</tr>
<tr>
<td>19</td>
<td>21.30</td>
<td>17.04</td>
<td>17.64</td>
<td>R</td>
</tr>
<tr>
<td>20</td>
<td>20.47</td>
<td>16.07</td>
<td>16.83</td>
<td>R</td>
</tr>
<tr>
<td>21</td>
<td>18.54</td>
<td>15.07</td>
<td>15.21</td>
<td>R</td>
</tr>
<tr>
<td>22</td>
<td>22.01</td>
<td>17.05</td>
<td>18.64</td>
<td>R</td>
</tr>
<tr>
<td>23</td>
<td>20.51</td>
<td>15.47</td>
<td>16.82</td>
<td>R</td>
</tr>
<tr>
<td>24</td>
<td>21.17</td>
<td>16.24</td>
<td>16.41</td>
<td>R</td>
</tr>
<tr>
<td>25</td>
<td>20.86</td>
<td>15.07</td>
<td>15.48</td>
<td>R</td>
</tr>
<tr>
<td>26</td>
<td>21.46</td>
<td>15.61</td>
<td>16.44</td>
<td>R</td>
</tr>
<tr>
<td>27</td>
<td>21.27</td>
<td>15.19</td>
<td>16.02</td>
<td>R</td>
</tr>
</tbody>
</table>
Table 2. Geometry of the three mazes used in the pen trials, along with movement times (MT) in seconds and errors.

<table>
<thead>
<tr>
<th>Maze Size</th>
<th>Track width</th>
<th>Effective Track Width</th>
<th>First Path Length (mm)</th>
<th>MT (320)</th>
<th>MT (180)</th>
<th>MT (50)</th>
<th>Error (320)</th>
<th>Error (180)</th>
<th>Error (50)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large (4)</td>
<td>6</td>
<td>5.5</td>
<td>47</td>
<td>5.63</td>
<td>5.32</td>
<td>5.25</td>
<td>2.45</td>
<td>2.17</td>
<td>1.81</td>
</tr>
<tr>
<td>Medium (2)</td>
<td>3</td>
<td>2.5</td>
<td>23.5</td>
<td>5.88</td>
<td>5.55</td>
<td>5.51</td>
<td>3.36</td>
<td>2.77</td>
<td>2.56</td>
</tr>
<tr>
<td>Small (1)</td>
<td>1.5</td>
<td>1</td>
<td>11.75</td>
<td>6.59</td>
<td>6.24</td>
<td>6.19</td>
<td>6.63</td>
<td>5.23</td>
<td>3.88</td>
</tr>
</tbody>
</table>

Table 3. Cross-sectional area (mm\(^2\)) of different shape and size pens

<table>
<thead>
<tr>
<th>Shape</th>
<th>small</th>
<th>medium</th>
<th>large</th>
</tr>
</thead>
<tbody>
<tr>
<td>triangular (t)</td>
<td>21.3</td>
<td>27.8</td>
<td>35.5</td>
</tr>
<tr>
<td>square (s)</td>
<td>29.7</td>
<td>40.9</td>
<td>55.9</td>
</tr>
<tr>
<td>elliptical (e)</td>
<td>30.0</td>
<td>43.4</td>
<td>60.1</td>
</tr>
<tr>
<td>Hexagonal (h)</td>
<td>37.1</td>
<td>53.3</td>
<td>70.7</td>
</tr>
<tr>
<td>octagonal (o)</td>
<td>39.0</td>
<td>56.5</td>
<td>75.9</td>
</tr>
<tr>
<td>circular (c)</td>
<td>44.5</td>
<td>64.1</td>
<td>87.2</td>
</tr>
</tbody>
</table>
Table 4. Weight (grams) for the different size and shape of pens

<table>
<thead>
<tr>
<th>Shape</th>
<th>Light</th>
<th>Heavy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>small</td>
<td>medium</td>
</tr>
<tr>
<td>triangular (t)</td>
<td>10.8</td>
<td>16.9</td>
</tr>
<tr>
<td>square (s)</td>
<td>11.0</td>
<td>17.5</td>
</tr>
<tr>
<td>hexagonal (h)</td>
<td>11.7</td>
<td>18.5</td>
</tr>
<tr>
<td>elliptical (e)</td>
<td>11.7</td>
<td>18.4</td>
</tr>
<tr>
<td>octagonal (o)</td>
<td>12.0</td>
<td>18.8</td>
</tr>
<tr>
<td>circular (c)</td>
<td>12.2</td>
<td>19.4</td>
</tr>
</tbody>
</table>
Table 5. Factor Analysis with varimax rotation of the variables Overall Comfort, Writing Ability, Overall Evaluation, and Writing Time. The factor loadings greater than 0.5 are shown in bold.

<table>
<thead>
<tr>
<th></th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
<th>Factor 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Comfort</td>
<td>0.947</td>
<td>0.314</td>
<td>-0.022</td>
<td>0.068</td>
</tr>
<tr>
<td>Writing Ability</td>
<td>0.316</td>
<td>0.946</td>
<td>-0.017</td>
<td>0.069</td>
</tr>
<tr>
<td>Overall</td>
<td>0.627</td>
<td>0.622</td>
<td>-0.007</td>
<td>0.469</td>
</tr>
<tr>
<td>Evaluation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Writing Time</td>
<td>-0.017</td>
<td>-0.013</td>
<td>0.999</td>
<td>-0.002</td>
</tr>
<tr>
<td>Variance</td>
<td>1.389</td>
<td>1.3811</td>
<td>1.000</td>
<td>0.230</td>
</tr>
<tr>
<td>% explained</td>
<td>34.7%</td>
<td>34.5%</td>
<td>25%</td>
<td>5.8%</td>
</tr>
<tr>
<td>variance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cumulative</td>
<td>34.7%</td>
<td>69.3%</td>
<td>94.3%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Variance explained

% explained variance

Cumulative
Figure 1. The three mazes used for drawing. They are geometrically similar: Maze (b) is half the size of Maze (a); Maze (c) is one-quarter the size of maze (a).

Figure 2. Effect of pen size (Equivalent Diameter) and Maze Size [large (4), Medium (2), and Small (1)] on movement time through the maze.
Figure 3a. Drury’s model for tracking used to explain the effects of a finite pen-tip width on time to move through mazes of varying sizes.

Figure 3b. Same data as in Figure 3a, except plotted in the speed form of Drury’s model.
Figure 4. Tracking errors as a function of maze and pen sizes, showing the correlation in terms of the Drury tracking model and the significant interaction between maze size and pen size. Numbers 320, 180, 50 refer to the cross-sectional areas of the pen shafts in mm².
Figure 5. Tracking errors per unit distance traveled as dependent on the effective path width and the pen size. The numbers 320, 180, 50 refer to the cross-sectional area of the pens in mm².

Figure 6. Errors per unit distance versus movement time for all experimental conditions.
Figure 7. (i) Cross-sectional shapes of the pens used in this study Circular (c), Elliptical (e), Hexagonal (h), Octagonal (o), Square (s) and Triangular (t). (ii) Eighteen of the thirty-six pens used in the experiment. The other eighteen differed only in the weight of the middle section.
Figure 8. (a) The pen with the circular section. (b) The three pieces that make up the pen.

Figure 9. An example showing the grip of one particular subject when writing
Figure 10. Subjective ratings when writing in Chinese

Figure 11. Subjective ratings when writing in English