Usability News

The Effects of Perceptual Grouping on Text Entry Performance

By Christopher J. Hamblin, Michael Bohan, & Alex Chaparro

Summary: One of the primary challenges confronting designers of mobile computing devices is the issue of efficient text entry. One potential solution is to group multiple letters onto single keys, similar to the T9 keyboard currently used on telephones. Two experiments examined the effects of perceptual grouping on soft keyboard transcription rates. Results from Experiment 1 showed significantly slower transcription rates for QWERTY keyboards with grouped keys. Results from Experiment 2 showed various levels of perceptual interference due to the different Gestalt grouping effects. These results indicate that perceptual grouping can negatively affect text entry performance, and placing multiple letters onto single keys reduces the speed at which users can transcribe words.

INTRODUCTION

Technological advances in the computing industry have lead to a dramatic increase in the accessibility of mobile computing devices such as pocket PCs, personal digital assistants, pen tablets, web-enabled cellular phones, and two-way pagers (Mobile.com, 2002; Legard, 2000). As these devices continue to provide increased functionality while at the same time becoming progressively smaller, data entry for these devices becomes a critical issue.

Previous research has shown that novice users are typically faster and more accurate using a QWERTY keyboard than with alternative keyboard design (e.g. Bohan, Phipps, Chaparro, & Halcomb, 1999; McQueen, MacKenzie, Nonnecke, & Riddersma 1994; MacKenzie, Zhang, & Souckoreff, 1999). Thus, shrinking the QWERTY keyboard to accommodate mobile devices provides familiarity at the expense of speed.

Altering keyboards to optimize data entry often results in a loss of immediate usability. Several "high-performance" keyboard designs such as the OPTI (MacKenzie & Zhang, 1999) and the Fitaly (Broida, 1998) have been proposed to minimize travel distance between keys. MacKenzie and Zhang demonstrated that users were able to achieve tapping rates upwards of 44 wpm on the OPTI keyboard after approximately 15 hours of use compared to 40 wpm for the standard QWERTY keyboard with the same amount of practice. However, initial performance of novice typers was considerably faster with the more familiar QWERTY (28 wpm) than the OPTI keyboard (17 wpm).

Additional keyboards have been proposed that attempt to achieve a balance between more efficient

input and ease of use. Hashimoto and Togasi (1995) attempted to maximize the performance of a virtual QWERTY keyboard by adjusting its shape to better accommodate the biomechanical attributes of the user's wrist and fingers. The authors demonstrated that participants could point to the designated keys significantly faster on the oval keyboard than on the standard rectangular keyboard layout.

The T9 keyboard (Tegic communication, 1998) incorporates the familiar alphanumeric layout of touchtone telephones where 3-4 letters are grouped onto a single button in alphabetical order. Applying Fitts' Law (Fitts & Peterson 1964; Fitts, 1954), this keyboard is potentially better suited for touch tapping in small spaces because it reduces the number of keys in addition to reducing travel time between keys. In this case, tapping sequential letters that are grouped on the same key would minimize travel time. Bohan, et al. (1999) investigated the immediate usability of the T9 keyboard relative to the standard QWERTY keyboard of the same overall size. It was expected that participant's performance would be faster and more accurate using the T9 keyboard due to the larger blocked keys and the ability to make multiple taps on a single key when selecting letters within the same block; however, transcription rates were actually higher for the QWERTY keyboard (23 wpm) than for the T9 keyboard (17 wpm).

There are at least two interpretations of Bohan, et al.(1999) results. One interpretation is that the slower performance of the T9 keyboard relative to the QWERTY keyboard reflects differences in the participant's familiarity with the respective keyboard. Alternatively, performance may have been hindered due to perceptual interference in accordance with the Gestalt principle of common region (Palmer, 1992) which states that elements will be perceived as a group together if they are located within a common region of space or an enclosed boundary. This grouping effect may in turn make it more difficult to process the individual elements within a group. The letters on the telephone keypad are arranged on the buttons in groups of three or four, which will cause them to be perceptually grouped. The resulting group would then have to be decomposed into its requisite component letters before a target letter could be identified. If perceptual grouping does negatively impact transcription performance in this way, the effects may be strong enough to counteract the motor advantages supported by larger keys (Bohan, et al. 1999). The following experiments address both explanations.

EXPERIMENT 1

The first experiment compared the transcription rates of participants using a single key (Qs) keyboard and a modified keyboard (Qm) in which the arrangement of the letters were not altered; however, keys were merged so that the width of the keys were increased by approximately four fold with up to three letters on a single key (see Figure 1). Participants also transcribed words using single- and multiple-key versions of the ABC keyboard layout (As and Am, respectively). It was reasoned that perceptual grouping would have a greater impact on transcription performance when participants are less familiar with the spatial arrangement of a particular keyboard thus having to rely a greater extent on visual scanning to find the target letter. Whereas with experience the participants may rely more on remembered spatial locations and kinesthetic positional cues rather than visual scanning to find the target letter.

Participants transcribed words presented at systematically varying rates using soft keyboards graphically displayed on a touch-screen. The average maximum transcription rate for each keyboard was calculated and used to compare text entry performance. According to the perceptual grouping hypothesis, transcription rates for the Qs should be as high or higher than those for the Qm due to the additional visual processing. The alternative hypothesis is that grouping letters on a single key does not introduce additional visual processing resulting in significantly faster transcription rates for the Qm due to the key's larger target size and the ability to make multiple taps on a single key. In addition, transcription rates for the QWERTY keyboards would be higher than those of the ABC keyboards due to difference in familiarity. In addition, transcription rates for the As keyboard should be higher than those for the Am keyboard.

Figure 1. Qm keyboard with multiple letters grouped onto single keys.

METHOD

Calculation of Transcription Rate

To measure participant performance, a mathematical formula originally developed to measure reading rates was used (Legge, Pelli, Rubin, & Schleske, 1985). Transcription rates were calculated by multiplying the percentage of correctly transcribed words by the text presentation rate (wpm). Maximum average transcription rates were estimated by fitting each participant's performance curve with a second order polynomial function (i.e. parabola). The peak of the function was estimated by solving for the first-order derivative of the function. This method minimizes the potential impact of experimenter bias in estimating the peak when the function was relatively flat. Maximum average transcription rates were used to measure the mean differences between each of the keyboard formats.

Participants

Ten right-handed participants, ages 28-45 (M = 34) participated in the study. All participants reported having 20/20 corrected/uncorrected vision. All participants had extensive computer experience and four had used a PDA.

Apparatus

A custom computer program written in Microsoft Visual Basic 6.0 was created to display the various keyboards and collect the data. A land-on touch selection strategy was used to activate the "buttons" of the soft keyboard. For these experiments the actual size of the keys were not a concern as the purpose of the experiments was to determine if there was an effect for grouping; as such the size of the keys were four times larger than a traditional keyboard in an effort to minimize inadvertent errors. With participants seated approximately 24 cm from the screen single key targets subtended 2.5° visual angle with alphanumeric characters within the keys subtending 1.67° visual angle. Grouped keys employed the same size characters; however, the key target width subtended to 6.95° visual angle to include the additional letters. The program ran on a 233 MHz IBM-compatible computer and stimuli displayed on a 15-inch diagonal 3M Microtouch CRT Touch Monitor (1024 x 768 pixel resolution) with a resistive touch-sensitive screen.

Procedure

The height of the monitor was adjusted to center the screen approximately at eye level. Participants started each trial by tapping a "START" button displayed on the screen. Upon starting the trial ten four-letter words, randomly selected from a database, appeared individually at the top of the screen. Participants were required to transcribe the word using the soft keyboard displayed on the touch-sensitive screen. The participant's transcriptions were displayed in a window directly below the target word. In trials involving grouped keys the software automatically determined the correct letter to display in the transcription window by comparing the sequencing of the tapped keys to the letters in each word. For example, if the word "BIRD" was displayed at the top of the screen, the participant would correctly tap the "VBN" key for the letter B, the "UIO" key for the letter "I" and so on. If the tapped key contained the required letter the program tallied a correct response and displayed the letter in the transcription window. If the tapped key did not contain the required letter the program tallied a miss and a hyphen was displayed in the transcription window. Each trial presented the words at different rates (10, 15, 20, 25, and 30 words per minute). Participants performed 5 trials at each presentation rate for each keyboard resulting in 25 trials per keyboard. The order of presentation rate and order of keyboard layout was randomized for each participant. Each experiment took approximately 25 minutes.

RESULTS AND DISCUSSION

Maximum average transcription rates were subjected to a 2 (grouping) X 2 (keyboard layout) within-subjects ANOVA. Analysis showed a significant main effect for keyboard layout, $F (1, 9) = 437.81$, $p < .001$, as transcription rates were significantly faster for the QWERTY keyboards than for the ABC keyboards. There was not a significant main effect of grouping, $F (1, 9) = 1.23$, $p = .285$; however, there was a significant interaction between keyboard layout and grouping $F (1, 9) = 14.122$, $p = .004$. Paired sample t-tests comparing transcription rates within keyboard layouts showed a significant difference between Qs and Qm keyboards, $t (9) = 3.29$, $p = .009$ (see Table 1).

Table1. p-values for pairwise comparisons of transcription rates in Experiment 1.

	As	Am	Qs
Am	.299		
Qs	<.001	<.001	
Qm	<.001	<.001	.009

The results of this experiment showed significant differences in transcription rates. In the case of the ABC keyboard, the results were likely indicative of a floor effect. That is, the novelty of the ABC keyboard caused search times long enough that any additional searching induced by the grouping effect would not have a significant affect. The results support the hypothesis that perceptual grouping significantly affects transcription rates. The results also support Bohan, et al.'s (1999) previous findings that transcription rates for the QWERTY keyboard were faster than the T9 keyboard despite the larger blocked keys.

EXPERIMENT 2

As a follow-up study, a second experiment was designed to measure the relative effects of different types of Gestalt grouping strategies (Palmer, 1992; Wertheimer, 1923) on text entry performance using soft keyboards. Specifically, participants transcribed words using QWERTY keyboards with letters that were not grouped, grouped by common region, or grouped by proximity. In the case of proximity grouping, letters were segregated by spatial distance between the letters. The purpose of this experiment was to explore the effects of perceptual grouping on text entry performance on soft keyboards.

Participants

Ten right-handed participants ages 29-49 (M = 36) participated in the study. All participants reported having 20/20 corrected/uncorrected vision. All participants had extensive computer experience and three had used a PDA.

Apparatus

The task and apparatus were identical to those used in Experiment 1 with the exception of the keyboards used. Like the first experiment, the single-key and multiple-key QWERTY keyboards were used in addition to a keyboard grouped by proximity. The proximity keyboard (Qp) consisted of letters grouped together in groups of 1 to 3 by spatial distances between keys; that is, there were not box outlines to contain the letters (see Figure 2).

Figure 2. Screen shot of the Qp keyboard with multiple letters keys grouped by spatial proximity.

RESULTS

A significant main effect was found for grouping type, F (1, 9) = 5.87, p = .006. Pairwise comparison t-tests showed significant differences between the Qs keyboard and the grouped keyboards: Qm (p = .05), Qp (p = .043), Qc (p = .021). Significant differences between the grouped keyboards are shown in Table 2. Interestingly, transcription rates for the Qc condition was significantly different from all the other conditions suggesting that the Qc condition has a much stronger perceptual grouping effect that grouping by common border or proximity. The reason for this effect is unknown but will be the focus of continuing research.

Table 2. *p*-values for pairwise comparisons of transcription rates in Experiment 2.

	Qs	Qm	Qp
Qm	.051		
Qp	.043	.813	
Qc	.021	.076	.031

DISCUSSION

Tapping on a miniaturized virtual QWERTY keyboard has been shown to be the optimal solution for data entry because of its widespread familiarity (MacKenzie, Zhang & Soukoreff, 1999). Thus, a logical approach to finding a balance between these tradeoffs is to maximize the performance of a familiar text entry method like the QWERTY keyboard (Hashimoto & Togasi, 1995).

Given what is known about speed/accuracy tradeoffs in human performance, one way to improve performance on soft keyboards is to increase the size of the keys. However, size constraints of mobile computing devices prevent this option. Alternatively, placing multiple letters on a single key reduces the number of keys required and minimizes the space required. However, this study suggests that grouping letters onto a single-key may decrease transcription performance due to additional perceptual processing steps despite a four-fold increase in target size.

The results of these experiments illustrate that the need to visually parse groups of letters into their constituents before a target letter can be identified will result in a significant decrease in text entry speed. In addition, comparison of mean transcription rates in the second experiment suggest that strength of the perceptual interference on text input is consistent between the types of grouping effects tested. Ongoing research is currently investigating the effects of other grouping effects.

Note: For additional information, please see the proceedings of the Human Factors and Ergonomics Society's 47th (2003) Annual Meeting.

REFERENCES

Bohan, M., Phipps, C.A., Chaparro, A., & Halcomb, C. (1999). A psychophysical comparison of two stylus-driven soft keyboards. Proceedings of Graphics Interface '99, 92-97.

Broida, R. (1998). Input Alternatives - Fitaly vs. T9. Tap Magazine, Issue 1.6, November 1998 http://www.fitaly.com/

Fitts, P.M. (1954). The information capacity of the human motor system in controlling amplitude of movement. *Journal of Experimental Psychology, 47*, 381-391.

Fitts, P.M., & Peterson, J.R. (1964). Information capacity of discrete motor responses. *Journal of Experimental Psychology, 67,* 103-112.

Hashimoto, M., & Togasi, M. (1995). A virtual oval keyboard and a vector input method for pen-based character input. Proceedings of the ACM Conference on Human Factors in Computing Systems. Retrieved September 11, 2001 from: http://www.acm.org/sigchi/chi95/proceedings/shortppr/mh_bdy.htm.

Legard, David (2000). PDA Sales to Double in 2000. PCWorld, August 28, 2000.

Legge, G.E., Pelli, D.G., Rubin, G.S., & Schleske, M.M. (1985). Psychophysics of reading – I: Normal vision. *Vision Research, 25*, 239-252.

MacKenzie, I. S., & Zhang, S. X. (1999). The design and evaluation of a high-performance soft keyboard. Proceedings of the ACM Conference on Human Factors in Computing Systems - CHI '99, pp. 25-31.

MacKenzie, I. S., Zhang, S. X., & Soukoreff, R. W. (1999). Text entry using soft keyboards. *Behaviour & Information Technology, 18*, 235-244.

McQueen, C., MacKenzie, I. S., Nonnecke, B., & Riddersma, S. (1994). A comparision of four methods of numeric entry on pen-based computers. Proceedings of Graphics Interface '94, pp. 75-82. Toronto: Canadian Information Processing Society.

Mobile.com (2002, November 5). Gartner: PDA Sales Up or Down. News issue 42. Retrieved December 4, 2002, from:
http://www.mobileinfo.com/News_2002/Issue42/PDA_Dataquest.htm

Palmer, S.E. (1992). Common Region: A new principle of perceptual grouping. *Cognitive Psychology, 24,* 436-447.

Tegic Communication (1998). http://www.t9.com/t9_faqs.html

Wertheimer, M. (1923). Untersuchengen zur Leher von der Gestalt. Psychologische Forshung, 4, 301-350. [Reprinted in part in W.D. Ellis (Ed.), A sourcebook of Gestalt Psychology (pp. 71-88). New York: The Humanities Press, 1950].

SUBSCRIBE to *Usability News*!