Usability of text input interfaces in smartphones

Tom Page

Electronic Product Design,
Loughborough Design School, LDS.1.18,
Loughborough University,
LE11 3TU, UK
Fax +44 (0) 1509-223999
E-mail: T.Page@lboro.ac.uk

Abstract: This paper reviews a number of different text input methods on smartphones and proposes how smartphone typing methods may inhibit user interaction rather than enhance it. It explores design and ergonomic considerations underpinning the use of six commonly used smartphone text input methods in order to determine and compare their usability. Furthermore, it discusses the results from testing smartphone users to assess which method is considered most efficient. It concludes that alternative input methods such as Swype and SwiftKey offer substantial benefits to users and are comparable with common typing speeds found on computer keyboards.

Keywords: smartphones; user interface design; ergonomics.

Reference to this paper should be made as follows: Page, T. (2013) 'Usability of text input interfaces in smartphones', *J. Design Research*, Vol. 11, No. 1, pp.39–56.

Biographical notes: Tom Page mainly teaches electronic product design and has been working between the fields of interaction, physical computing and electronic product design for the last couple of years. His main interests lie in the area of technology mediated learning, value management and electronic product design. He has been a consultant for many small and medium sized enterprises (SMEs) and large-scale manufacturing and service enterprises in the UK, in engineering design and electronic product design. His work has been widely published in the form of journal papers, book contributions, refereed proceedings, refereed conference papers and technical papers.

1 Introduction

1.1 Background

Interaction design has played a central role in the design of small touchscreen devices as a consequence of advances in display technology (Saffer, 2008). Furthermore, since the launch of the Apple iPhone in 2007, touchscreen devices have become the essential interfaces in smartphone design. In fact, many people nowadays have a 'smartphone mindset' and this is of increasing importance to interaction designers as "*smartphones keep getting smarter, more of us are leaving our laptops at home*" [Clark, (2010), p.32].

In essence, smartphone interfaces today have been designed in completely different ways as users are starting or seeking to replace laptops and computers with small screen highly portable devices. "Slowly but surely, we're learning to get stuff done on the small screen" [Clark, (2010), p.32] and this provides significant insight into the transforming psychology of smartphone users. According to Friedman (2007), users have had to adapt to new devices but in doing so we have 'constrained our abilities'.

1.2 Aim of this research

The aim of this research is to review text input methods on smartphone soft keyboards in order to evaluate their usability. The inherent usability of an object may be defined as the quality of use that a user experiences when in contact with the object (Bevan, 1995). Furthermore, this work identifies theories that underpin what might be considered as good or poor interface design and therefore attempts to define usability in the context of smartphone soft keyboards.

1.3 Objectives of the research

The objectives required to fulfil this aim required an understanding of how usability is influenced by different text input styles on smartphones. Further to this, a review of literature was undertaken in order to bring together the principles and guidelines relating to soft keyboard text input and product usability. The literature combined with various psychological mental models or personas provided a basis to analyse such products [Goodwin, (2009), p.229]. This contributed to the development of a research methodology for data collection and analysis which formed a framework and criteria for evaluation of the text entry styles in current smartphones.

2 Review of text input methods on smartphones

There are a number of different methods for text input on smartphones and the most commonly found is the QWERTY soft keyboard layout. Other, more ergonomic soft keyboard layouts such as DVORAK and ABCDE improve on typing comfort and speed (Clawson et al., 2010). Significantly, other text input methods exist such as: OPTI, 8pen; Swiftkey; Swype; Keypurr and thick buttons. Their developers claim that such text input methods improve upon text input as the QWERTY soft keyboard is considered inaccurate on very small smartphone soft keyboards [Shneiderman and Plaisant, (2005), p.325]. Furthermore, advances in text prediction and error correction are considered as equally important due to inaccuracies of text input on such small smartphone soft keyboards (Cooper et al., 2007).

2.1 The soft keyboard

Fundamentally, the success or failure of any new interactive technology or text input method such as soft keyboards is determined by its usability. Usability may be described as "*the extent to which a product can be used by specified users to achieve specific goals with effectiveness, efficiency and satisfaction*" (ISO, 1998). The "*most important* *principles of interface design are to deliver intuitive and efficient interfaces*" (Page, 2011).

An intuitive interface may be considered as one that requires minimal cognitive load and feels intuitive and natural for the user to use (Shneiderman and Plaisant, 2005). An efficient soft keyboard layout may be considered as one that produces the least amount of errors and allows the quickest text entry rate words per minute (WPM) resulting in minimal overall completion time. Relating to efficiency, Clawson et al. (2010) suggest that "*Fitts' Law, which describes the relationship between speed of movement, target size, and accuracy, implies that users will type less accurately as they increase their rate of text input*" (Hirche et al., 2008). In other words, as users become more efficient in their text input speed they will reduce accuracy, produce more errors and therefore reduce text input efficiency.

2.2 Keyboard ergonomics and how it affects typing on soft keyboards

Soft keyboards on smartphone devices are typically very small compared to what many users are familiar with. A computer keyboard is typically 700 mm in width from tab key to the enter key, and its expected use is two handed. A typical smartphone such as the iPhone or HTC Desire (3.2" screen size) has a portrait screen width of 50 mm and a landscape width of 80 mm. In portrait mode, a button is 4 mm wide whereas in landscape mode, it is 6 mm wide. Smartphone soft keyboards can be used double- or single-handed and neither is more common than the other.

The ergonomic aspects of soft keyboard typing on a smartphone differ greatly from their physical counterpart. The main differences is that in physical keyboard typing, the users' wrists are free and flexible to move around and accommodate the most comfortable and efficient typing posture. When typing on a smartphone QWERTY layout the user's wrist is in a locked position and thus limiting finger and thumb mobility (Baumann and Thomas, 2001). This is another reason why the DVORAK keyboard layout makes ergonomic sense on small soft keyboards as movement is already restricted by having to hold the device still whilst typing (Hirche et al., 2008). Often smartphone users adopt the portrait mode for typing because they can then effectively type using two hands and have full wrist and fingertip flexibility. In this method of typing, users typically hold the device in one hand and then use the other hand to create a stylus with their index finger (Page, 2011). This is the method adopted in the user trials as it allows user maximum flexibility and remains a constant single input whereas the thumb input method will vary as some users will be better at utilising both thumbs.

3 Smartphone input methods

3.1 ITU-T numeric keypad

Smartphone text input interaction is an exemplar of the user-centred design process that many companies now adopt. Presently, two physical hardware input methods are used on smartphones, one is the increasingly popular touchscreen input and the other is the physical keyboard based on the ITU-T keypad. The ITU-T keypad is the traditional keypad with 12 basic keys, that is, number key 0–9, *-key and #-key. ITU-T is a standard of International Telecommunication Union.

Touchscreen keyboards are essentially a software iteration of the physical piece of hardware and this is where the flexibility comes from, for example the Android smartphone operating system (OS) allow for multiple keyboard layouts for a customised typing experience. The default options available to android users and most other smartphone operating systems too are QWERTY, phone keypad (ITU numeric keypad) and compact QWERTY as shown in Figure 1. The latest area of interest and major development to soft keyboards are within the software capabilities and the customisation that can occur. With the Android smartphone operating system being completely open-source it has received the most attention from software developers. The stock Android keyboard layouts shown in Figure 1 can be altered and many different variations of the above are now emerging including some completely re-thought ones.

Figure 1 Android OS smartphone keyboard layouts (see online version for colours)

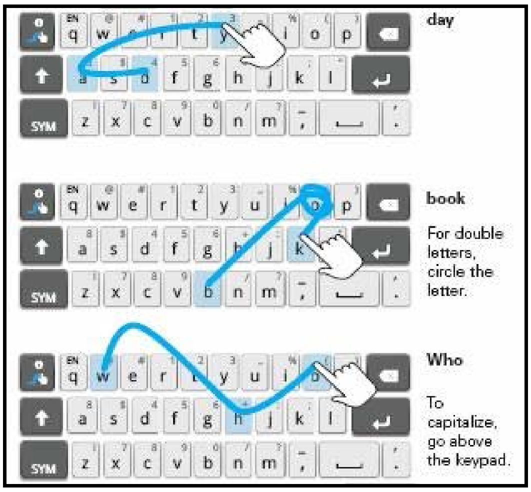
Source: Paul (2009)

3.2 Swype input method

Swype is a gestural input method that claims to improve soft keyboard typing. Swype is commercially available across operating systems such as Android, Blackberry, Windows and Symbian. It is an intelligent software enhancement which analyses a trace or route taken to type, or in this case swipe a word, shown in Figure 2. Its powerful predictive text engine allows for accurate prediction of the typed word even if the trace is not exact, this allows for reduced errors whilst typing and a claimed improvement in speed.

Swype has an innovative way of displaying predicted words allowing auto correction a prediction with a single tap. The only disadvantage to Swype is that its method for text entry is not very intuitive as the majority of users are used to tap typing. Swype claims that users can achieve over 40 WPM with increased accuracy over the QWERTY keyboard.

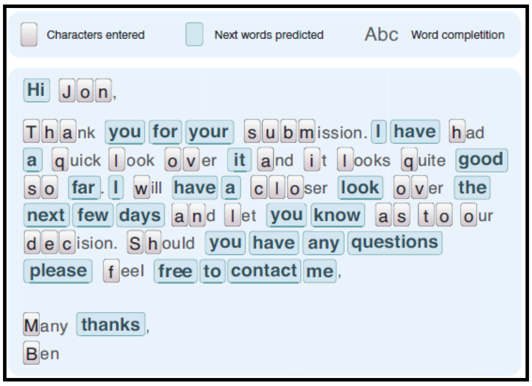
Figure 2 Swype input method (see online version for colours)



3.3 Swiftkey input method

SwiftKey is another very useful input method, and uses their patent-pending Fluency™ technology to automatically predict the next word a user will type, before they have even started typing the next word, shown in Figure 3. Their user testing trials claim that a third of the next word predictions are correct, and usually 80% are predicted within two or fewer key presses. SwiftKey learns the user's style of writing and vocabulary used and its AI (artificial intelligence) grows with every word and sentence written. SwiftKey has other features such as gestural support aimed at improving interaction, using gestures the user can minimise the keyboard (swipe down), delete previous word (swipe across screen from right to left) and shift to upper case (swipe up). As is typical with most of these advanced input method apps, SwiftKey offers automatic features too such as auto completion if space is pressed and automatic spelling and grammar.

Figure 3 Swiftkey input method (see online version for colours)



3.4 Thick buttons input method

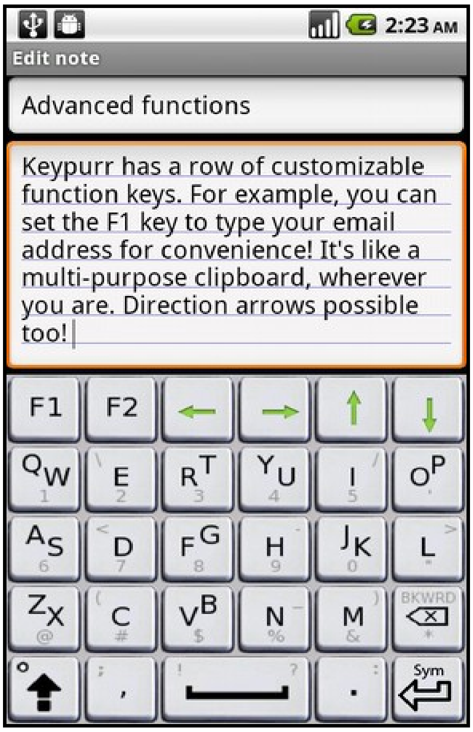
Thick buttons address the source of most errors that people experience with conventional stock keyboards, this being that keys are too small and often users press the wrong letter. The buttons actually change shape based on prediction, so that they get larger and easier to hit. Moreover, as the user types more letters of a word the prediction narrows down the possibilities and this too also increases input speed as the predicted letters are highlighted and enlarged, shown in Figure 4. This is an adaptive keyboard if people are looking to improve upon the accuracy of their touch typing experience without investing much time on learning a new keyboard layout or input method. However, it is hard to see how this improves text entry speed as the key location is constantly changing therefore using the visual scan time model this app would if anything slow typing speed down. The developers do however claim a 31% increase in typing speed and 24% fewer errors.

Figure 4 Thick buttons input method (see online version for colours)

3.5 Keypurr input method

This keyboard is similar to thick buttons in that it uses letter frequency to create a custom keyboard layout. The main aim of this input method is to improve upon the keyboards usability and enable single handed input to be possible. Although it is a custom keyboard it is still based on the QWERTY layout and thus requires minimal learning effort and time. In this custom QWERTY layout there keys are only 17 keys whereas a conventional QWERTY layout has 26, therefore each key is larger and the most frequently used letters have a single key whereas the less frequently used letters share a key, shown in Figure 5. This input method also has an improved text prediction engine allowing users to select the finished word before typing the whole thing. Keypurr's main goal is to prevent users from having to flick through screens to get to numbers or special symbols and this is the main driving force behind the new layout as all common symbols are also located on the keyboard sharing keys with letters thus allowing quicker operation and improved usability. Users can still access a special number and symbol screen which contains far more symbols than other keyboards such as copyright and trademark symbols. The selection of the desired character on a key depends on the gesture used when typing for example to select the upper letter the user selects the key and then swipes up.

Figure 5 Keypurr input method (see online version for colours)



4 Research methodology

The research methodology applied appropriate empirical tests based on the usability measures of Nielsen (1993). The focus of the methodology was to establish why one text input method is favoured over another. For the purpose of this study, six text input methods were chosen from a total pool of 15 methods. This was done because these six methods were all available as downloadable mobile applications and configurable for the chosen smartphone handset the HTC Sensation XL.A representative sample of the UK population of smartphone users were considered in order to: measure the task performance and user satisfaction of six different text input methods; understand why smartphone users encounter difficulties in the execution of text input tasks; and understand how a text input method can affect the perceived and inherent usability of a smartphone. A user segmentation strategy was used in this study to determine the level of experience of an individual by understanding their technology consumption, smartphone usage habits, attitudes and lifestyles. Identifying the experience level will determine if the learning barriers present in a text input method are due to the lack of usability in a handset or the inexperience of a user.

4.1 Procedure

Thirty two smartphone users aged between 18 and 52 years participated in the questionnaire, and all were familiar with the basic functionalities of e-mail, IM (instant messaging), internet browsing and application procurement (app downloading). For the validity of questionnaire, none of the participants held any technical training within the smartphone industry. Of the 32 participants, 14 were in the age group 18–29 years old, 12 were the age group 30–40 years old and six in the age group 41–52 years old. The questionnaire provided a method of screening in which the participants were subsequently placed into one of the three user segments based on first-time users, typical users and expert users. The inexperienced user is characteristic of low technology consumption, and no experience of using different text input methods. The moderate user; represented the majority of the population where smartphone usage is dominated by short messaging service (SMS) and call making. The experienced user; would easily converge between different text input methods, and values efficiency over intuitiveness (Kiljander, 2005) smartphone usage in applications such as e-mail, WAP, app downloading and VoIP.

The questionnaire was divided into three categories; technology background, attitudes and behaviour and technology consumption. Technology background was developed to understand the participants' smartphone history. Attitudes and behaviour were developed to identify the individuals' habits of smartphone usage. The technology consumption category was used to identify the amount of technology each respondent has been exposed to, shown in Table 1. Understanding the respondents' exposure to technology can represent their level of expertise [Rubin, (1994), p.38]. For example, an individual's exposure to a variety of electronic devices would allow users to converge between milti-modal text input interfaces with ease.

Table 1 The device ownership of the six selected individuals

Participant	*Sat-Nav*	*iPod*	*Laptop*	*MP3*	*Tablet*	*Camera*	*Smartphone*	*Total*
6	1	1	1	0	1	0	1	5
17	0	1	1	1	0	1	0	4
9	0	0	1	1	0	0	1	3
22	1	1	1	1	1	1	1	7
14	1	0	0	0	0	0	0	1
29	1	0	1	1	1	1	1	6

Table 2 illustrates that participants 6, 17, 9, 22, 14 and 29 were selected for empirical testing out of the 32 individuals. Participants 22 and 29 had the highest total count of device ownership and can be deemed expert users. Participants 14 and 9 had the lowest total count can be deemed first time users. The mean total for device ownership was a count of 4; individuals selected with a total count of 4 can be considered as typical users

4.1.1 User profiles derived from the screening questionnaire

Table 2 details the six user profiles selected for usability testing. The profiles were derived from the answers to the questionnaire and selected based on the author's judgement. The screening questionnaire has increased the validity of testing as each profile is a representative in age, gender and experience of the UK population of smartphone users (Nielsen, 1993).

Table 2 User profiles selected for text input usability testing

User profiles	*Age*	*Gender*	*Smartphone background*	*Interaction style experience*
1 First-time	48	Male	LG Optimus 2X	No experience with capacitive or resistive touch screen on smartphones. Text input via ITU keypad.
2 First-time	26	Female	Apple iPhone 4	Familiar with the standard ITU keypad. Text entry has been with an ITU keypad for the last four to six years.
3 Typical	43	Male	HTC Evo 3D	Familiar with ITU keypad, and resistive touch. Familiarity with five-way navigation, and text entry via a standard ITU keypad.
4 Typical	38	Female	BlackBerry Torch 9860	Familiar with ITU keypads, five-way navigation, and an extensive user of shortcuts. Predictive text user.
5 Expert	22	Male	Motorola Defy	Experience with PDA and standard smartphones. Text entry via a full QWERTY, ITU and DVORAK keypads. Familiar with resistive touch screens.
6 Expert	19	Male	Samsung Galaxy Note	Experience with PDA and standard ITU keypads. Text entry via full QWERTY, ITU keypad, stylus and capacitive touch.

4.2 Text input testing

The objectives of text input testing are to measure the participants' performance against the usability metrics of Ryu and Smith-Jackson (2006). The six individuals were instructed to undertake four representative tasks of typical text input. The tests were conducted individually during 15 minute sessions using a smartphone. The sessions were voice-recorded, paying particular attention to the smartphone screens. Each individual occupied one of the three user segments; first-time, typical, and expert users (Table 2). The four tasks were representative of typical smartphone operations performed on a frequent basis by the majority of smartphone users.

The smartphone used for testing was the HTC Sensation XL and it was selected on the basis of its ability to incorporate all six individual styles of text input. None of the participants had prior experience with the model. The test consisted of the completion of four text input tasks by the six selected users using an application called myTextSpeed downloaded from the android market for data recording. This application recorded all relevant measures that traditionally would have to been recorded using a video camera setup and then the recorded footage post-analysed (Weiss, 2002). These recordings comprised completion time, WPM, characters per minute, word accuracy, character accuracy and highlighted all the users text input errors in an analysis screen.

4.3 Empirical tests

Prior to the tasks, the users were given the smartphone for two to three minutes to familiarise themselves with its keypad and UI before beginning each task. To answer the study objective '*What are the attributes of a smartphone text input method that contribute to inherent usability in text input?*' the procedure was developed to assess two measures of task performance; time taken in seconds to complete a task and the input count of the task. These measures provided an indication of task efficiency and identified errors in the performance (Ryu and Smith-Jackson, 2006).The phrases were chosen because of the content of grammar, punctuation and capital letters within the body of text. This content requires the user to input text, symbols, reverse any necessary actions and evoke a range of sub-menus. This will assess the usability metrics of efficiency and error count.

4.3.1 Task 1

This task duration was approximately 20 seconds and comprised a standard usability pangramme to simulate quick instant messaging (IM) responses or texting. The users inputted 'The quick brown fox jumps over a lazy dog.' This test comprised ten words (including punctuation); nine words (excluding punctuation); and 42 characters (including spaces).

4.3.2 Task 2

This task duration was approximately 60 seconds and aimed to analyse complex symbol and punctuation performance. The users inputted 'The amount to pay is 17.5% of: £1523.47 please e-mail to test@gmail.com!'. This test comprised 14 words (including punctuation); 12 words (excluding punctuation); and 71 characters (including spaces). This content requires the user to input text, symbols, reverse any necessary actions and evoke a range of sub-menus. This will assess the usability metrics of efficiency and error count.

4.3.3 Task 3

This task duration was approximately 120 seconds and aimed to simulate replying to an e-mail. The users inputted 'Therefore during the modification of the descendants of any species, and during incessant struggle of all species to increase in numbers, the more diversified these descendants become, the better their chance of succeeding in the battle of life.' This test comprised 40 words (including punctuation); 38 words (excluding punctuation); and 245 characters (including spaces).

4.3.4 Task 4

This task duration was approximately 135 seconds and aimed to simulate entering or editing a word-processed document. The users inputted 'There was a bell beside the gate, and Dorothy pushed the button and heard a silvery tinkle sound within. Then the big gate swung open and they all passed through and found themselves in a high arched room, the walls of which glistened with countless emeralds'. This test comprised: 50 words (including punctuation); 46 words (excluding punctuation); and 260 (including spaces).

During the testing and at the end qualitative data was recorded as to interesting insights to problems occurred and opinions of the different input methods. Finally, each user was then asked which input method they preferred once all test were completed. The outcome of the user testing was to determine whether the current trend towards touchscreen typing on 'soft keyboards' is inhibitive and if the alternative input methods were actually more efficient than the stock input method as they claim to be.

4.4 Measures of task performance

The total input count of a task represents the total number of key presses from start to the finish. This method draws upon the keystroke levels developed in HCI to measure task performance (Card et al., 1980). The input count obtained from the six users is measured against the optimum input. This represents the shortest key press sequence to complete a task. The shortest path to task completion was calculated for each text input method using shortcut keys where appropriate. This measure of task performance draws upon the usability metrics established by Moon and Millison (2000) and provides an indication of the barriers to learning for the text input method. It may be assumed that the higher the input count the more difficulty the individual has experienced in the execution of the task. For instance, a count above the optimum indicates a high error rate as the user had to reverse their actions several times or had digressed from the correct menu sequence. The second measure of task performance is the time taken to complete a task. The task time can reveal several learning barriers in a text input method. If a menu structure is unorganised and complex, the time it takes for a user to navigate through a text input method will be high (Hick, 1952). The input counts were taken for each user, by reviewing the MyTextSpeedlog after each task. The input count of a task is assumed to offer more validity than the task time. This measure provides a better indication of the users' performance as the speed of text input can vary between each user. An expert user may produce speeds lower than first-time users without encountering a difficulty in the task. This may be due to personal preference, fear of making a mistake or psychological limitations of text entry, such as dyslexia.

4.5 Limitations of the procedure

One limitation within the procedure may be assumed where the last task performed by a user on a text input method will instigate the shortest task time. This is because the user will be more confident with the text input method having undergone the previous three tasks. This limitation was reduced by employing a balancing order effect. All four tasks were reordered for each individual. The task order for each user is outlined in Table 3. A further limitation of testing is that some users have had prior experience with the brands of the test phones. This may subjectively influence user satisfaction when determining perceived usability.

Table 3 Task order for text input testing

User	*Task order*
1	1,2,3,4
2	3,4,1,2
3	2,3,4,1
4	3,4,1,2
5	1,4,2,3
6	2,3,1,4

5 Results and analysis of usability testing

The data obtained in empirical testing was analysed by cross-examining the input count and time taken to complete tasks 1 to 4. Each participant's performance was analysed to identify if any learning barriers existed in the text input method and if they were a result of the users' lack of experience or elements within the text input method.

5.1 Data analysis of task 1

Results of task1 demonstrated that users with no experience of capacitive touch screens had difficulty with the text entry (Table 4). Users 1 and 2 produced the longest times for text entry on the all input methods with the exception of Keypurr. The reason for this trend can be explained in the lack of confidence whilst entering text. A contributing factor was the lack of tactile feedback for confirmation of a key press. This result causes the user to momentarily check that the onscreen letter corresponds to the last key pressed, causing a slight time delay and reducing the fluency of the task (Hoggan, et al., 2005). In the case of the Keypurr, this effect is heightened, as the user waits for the visual feedback of a key press. Users 1 and 2 have had no experience with capacitive touch screens, thereby their lack of convergence to a fully touch based device is apparent. Interestingly users 2, 3 and 4 were very close to each other in terms of words per minute. The Swiftkey text input method provided a notable range in terms of words per minute. The feedback from the users was that the key size of the QWERTY layout was too small and the use of a stylus for navigation and text entry was a tedious task.

Table 4 Words per min speed for task 1

Input method	*1*	*2*	*3*	*4*	*5*	*6*
ITU-T	22	25.4	36.7	26.5	43	40.4
Swype	32.2	35	47.7	54	48.5	57.5
Swiftkey	40.4	64	72.3	83	109	116.9
Thick buttons	34.4	21.8	21.5	27	33.6	22.9
Keypurr	19	20.9	17.5	22.4	16.6	24

5.2 Data analysis of task 2

Results of task 2 demonstrate that users 1 and 2 who showed difficulty with text entry, user 2 also produced the longest times with the text input exercise in task 2. Here, users 2, 4, 5 and 6 produced the lowest task times for all text input methods (McDaniel, 2009). All four users were familiar with ITU based keypads and therefore these results are as expected. Analysis of the input counts for 2 provides a deeper insight into task performance. Table 5 illustrates that all users were able to complete the sentence.

Table 5 Words per min speed for task 2

Input method	*1*	*2*	*3*	*4*	*5*	*6*
ITU-T	18.3	22	24	22.8	22	26.5
Swype	23.7	27.5	20.9	28	24.8	27.4
Swiftkey	55.2	82.9	46.2	58.5	62.5	81.7
Thick buttons	11.9	13.4	13.3	11	12	14.4
Keypurr	23.7	27.5	20.9	28	24.8	17.5

5.3 Data analysis of task 3

Swiftkey produced the highest speed of text input for task 3. The low performance of Thick buttons is also reflected in the input counts of the users (Table 6) revealing that all individuals experienced a long interaction sequence. Users 1, 3, and 4 produced very close task times for all text input methods. Due to the amount of content in the device the site map is very large requiring deep menu selections for the execution of simple tasks. The main reason for the lengthy input counts of thick buttons is due to the menu presentation after entering a character.

Table 6 Words per min speed for task 3

Input method	*1*	*2*	*3*	*4*	*5*	*6*
ITU-T	25.6	24	25	25.7	30	32.4
Swype	28.3	32.5	30	32.5	24.6	31.2
Swiftkey	62	94	69	63.5	52.4	92.3
Thick buttons	14.6	18.9	15.6	15.7	16.8	17.1
Keypurr	26.8	32	30	31.5	25.6	21.6

5.4 Data analysis of task 4

Thick buttons produced the lowest overall text input speed for task 4 (Table 7). Observation of task 4 revealed that the majority of users spent a large proportion of their overall time selecting options in sub-menus and then reverting back to the main menu. The reason for the low input counts can be explained by the visual simplicity present in the text input method. The single select key and uncluttered text input method on screen have helped the majority of users to navigate through the sub-menus in task 4 with little difficulty. The text input method is consistent as it presents the user with a limited number of options on screen at any one time. This reduces the complexity and cognitive load indecision making (Hick, 1952), increasing the perceived and inherent usability of the smartphone.

Table 7 Words per min speed for task 4

Input method	*1*	*2*	*3*	*4*	*5*	*6*
ITU-T	20	23.3	24	23.2	28.2	30.2
Swype	29.1	32	27	30	22.9	29.6
Swiftkey	66.3	92.3	61.2	61.7	50	89.6
Thick buttons	14.4	17.6	12.6	15	15.2	15.9
Keypurr	26	30.3	24.8	29.4	22.9	20.2

Observation of the empirical tests reveals that the graphical user interface of the smartphone's home screen had caused confusion with some of users as there are a variety of different metaphors of interaction; icon representation, numeric symbols and two softkey options. All representations are similar in size and colour; there is no major distinction for prioritisation, this reduces the important presence of the start menu which is not distinguishable.

5.5 Comparison of qualitative data and empirical testing

Analysis of the participants' feedback revealed that the Swiftkey input method produced the highest level of user satisfaction and the second lowest times for all tasks. The data provides an insight into perceived usability as it is subjectively obtained from the users and also inherent usability as the questions were directly related to ease-of-use attributes.

Swiftkey's good performance was due to all users having experience with an ITU keypad and the similarity of the interface with entry-level devices aiding inexperienced individuals. The data reveals that when the users were asked directly to state which input method they perceived most usable the majority selected Swiftkey. This underlines that attributes of ease-of-use were not affecting the perceived usability. The aesthetic nature of Swiftkey was subjectively influencing the users' perception of usability.

5.6 Discussion and evaluation of data collection and segmentation strategy

The text input methods were judged against their compliance with the usability principles and interface guidelines outlined in Section 2.2.A segmentation strategy was implemented prior to testing. This involved recruiting people with different levels of experience with electronic devices. It would be unjust to declare that a text input method had a negative effect on usability if it was only judged by inexperienced users. This was employed to understand if the learning barriers existed because of the user interface design or the inexperience of the user. The segmentation strategy identified three levels of experience; first-time, typical and expert users. Each of the three segments was developed by understanding the participants' technology consumption and smartphone background. Although the data gathered about these categories allows a valid judgement of the participants' experience to be made, the data is reliant on their honesty. For instance, a person could have answered inaccurately to questions for fear of judgement. As a consequence this has reduced the validity of the results, as those placed into the expert segment may have the skills of an inexperienced user.

6 Discussion of empirical data

There are several limitations present in the empirical data which impact upon the validity of the results. The task performance data was collected in a lit, indoor static environment, usually at a desk. Environmental noise and light levels did not influence the usability of the tasks. The variables of daylight glare on a display and loud external noise will influence smartphone tasks and are more representative of the context of use. To accurately evaluate usability, a variety of test environments must be considered that replicate the dynamic environments in which a smartphone is used (Vaananen-Vainio-Mattila and Ruuska, 2000).

The empirical data was gathered by undertaking four tasks which were believed to be representative of typical smartphone text input. However, it is subjective to define a task as typical because smartphone users and their goals vary so much that predicting how a smartphone will be used for the majority of time is complex (Bevan, 1995). A main user goal could be the ability to read web pages in a legible aspect ratio whilst physically moving and would therefore purchase a smartphone based on the display size at the compromise of other features and functionality. Thereby, it can be argued that usability should be judged on each user's individual goals and not a generic set of typical text input tasks (Nielsen, 1993).

Whilst Swiftkey was perceived to be most usable by the six participants and produced the second lowest task time overall, the interaction style of text entry is through a full QWERTY, the input method ideally requires two-handed use. This is not feasible in the majority of user tasks; a smartphone is for a lot of people secondary use to their current task. If Swiftkey was tested on different smartphone handsets the perceived and inherent usability of this text input method may have been different. Thereby, a more diverse set of user case scenarios should be conducted for a task performance analysis which may include tasks such as e-mail sending, document downloading and VoIP. All of which are likely to be conducted when seated, or stationary, as they permit more time and cognition from the user. A smartphone must be judged usable by its manufacturer's main user cases (Lindholm, et al., 2003). Furthermore the time duration of the tasks has reduced the validity of the results. The four tasks were conducted in 15 minute sessions; this amount of time is not representative of smartphone text input usage.

The contextual framework of the study assumes that an intuitive text input method will reduce the initial learning required of the user as the task performance results represent the evaluation of usability in the initial 15 minutes, there is no consideration of user contact after this period. For instance, the high task times of the users were due to momentary pauses of thought; where a participant would first decipher the next step of navigation before proceeding, rather than time spend producing errors. This is because each task is only performed once during the time frame and as with any task, practice makes perfect. It could therefore be assumed with task repetition a person could resolve their initial misconceptions of the text input method, condensing any initial errors and performing the task in a reduced amount of time.

The test data was based on the performances and feedback of the participants. However, it is believed that a smartphone user is not solely responsible for determining the usability of a text input method. The empirical data measured two characteristic of task performance, the time it takes to complete a task and the input count. Although the count can reveal a task contained with errors, the method of collection provides a degree of ambiguity. The count must include navigational inputs as well as direct key presses. It had proved difficult to count the inputs of navigation from the video footage, due to of the fluency of their use. Furthermore, the aim of scroll bars and gestured UI is to compress what would be several key presses into a single input. These interfaces are correctly merited with lower counts and ultimately help credit a product more usable. However, their accuracy and appropriateness was not considered. For example, scroll bars present in HTC Sensation XL's user interface allow a user to pan a list quickly with a stylus or finger but do not represent the participants' frustration when trying to accurately stop the scrolling. The measures of task performance used in the study do not depict the nuances of interaction which are important to determine inherent usability.

The empirical data depicted HTC Sensation XL as a product which negatively affects usability. It can be argued that the perceived and inherent usability was a consequence of the Windows mobile OS. However, none of the six contributors had experience with Windows Mobile, despite all being familiar with the UI in a Windows PC. All of the individuals found the convergence between the interfaces of entry-level smartphones and Windows operating system (OS) difficult. It would be insightful to test the HTC with users who do have experience with the OS. If then the HTC still performs below the other three styles, it can justly be deemed to have a negative effect on usability.

7 Conclusions

Observation of the smartphone text input methods demonstrated that the principles of HCI have been adopted by manufacturers. It is clear that Schneiderman's guidelines of usability have been applied to the majority of the smartphone text input methods. Hick' laws are reflected in the software and hardware elements of the smartphone chosen for this investigation, the HTC Sensation XL.

The attributes of a smartphone text input method that contribute to inherent usability were evaluated in the four empirical tests. These features do contribute to the inherent usability of smartphone interaction, but as mentioned previously, may not be appropriate or required in all contexts of use. Furthermore, the test data revealed that usability fluctuates within a task and demonstrates the complexity required to define a product as usable (Nielsen, 1993).

Information gained in empirical testing has revealed a selection of attributes that could be used as usability requirements for software and hardware UI. This can be useful for the design and development of novel smartphone interfaces. However, the characteristics must be appropriately integrated into a manufacturer's product portfolio. For instance, a full QWERTY may provide the best performance of text entry, but its cost in terms user interface platform support and engine architecture may be above the budget of entry-level smartphones. Thereby a set of usability principles should be optimally developed for product segmentation and not generalised for all smartphones.

Although smartphones have rapidly developed over the past decade the input method technology has not developed to a sufficient level where Smartphone keyboards are equivalent to computer keyboards. This lack of innovative soft keyboard development is becoming a major weakness of smartphones when compared to computers (discovered from the questionnaire). A good keyboard should be efficient and intuitive to use therefore when using these metrics to compare the current commonly adopted QWERTY keyboard alternative layouts prove to have higher usability.

The test data revealed that usability fluctuates within a task. For instance, a smartphone can present a well organised menu structure, but cause difficulties when entering text due to the limitations of the hardware. Therefore it is hard to measure the direct attributes which relate to inherent usability. These complications are accentuated due to the quantity of smartphone users; their individual needs and tasks differ so much that implementation of a set of user interface alone cannot ensure that a product will be usable (Bevan, 1995). Smartphones should aim to provide multi-modal experiences as multiple methods of inputs in a single handset can offset areas of weakness and increase the performance of individual tasks. Finally, it may be concluded that there is research still being undertaken on touchscreen technology and that capacitive touchscreens still have much development to come, whether it be in the software used to analyse touch instances or text prediction.

References

Baumann, K. and Thomas, B. (2001) *User Interface Design for Electronics*, Taylor & Francis, London.

Bevan, N. (1995) 'Usability is quality of use, human-computer interaction standards', Anzai and Ogawa (Eds): *Proceedings of the 6th International Conference on Human Computer Interaction*, July 1995, Elsevier, Yokohama.

Card, S.K., Moran, T.P. and Newell, A. (1980) 'Computer text editing: information processing analysis of a routine cognitive skill', *Cognitive Psychology*, Vol. 12, pp.32–74.

Clark, J. (2010) *Tapworthy: Designing Great iPhone Apps*, O'Reilly Media, Canada.

Clawson, J., Rudnick, A. and Starner, T. (2010) *Automatic Whiteout++: Correcting Mini-QWERTY Typing Errors Using 1*.

Cooper, A., Reimann, R. and Cronin, D. (2007) *About Face 3*, Wiley Publishing, Indiana.

Friedman, V. (2007, November 26) *Inspiration User Experience of the Future*, Smashing magazine [online] http://www.smashingmagazine.com/2007/11/26/mondayinspiration-user-experience-of-the-future/ (accessed 5 December 2010).

Goodwin, K. (2009) *Designing for the Digital Age: How to Create Human-Centred Products and Services*, Wiley Publishing, Indiana.

Hick, W.E. (1952) 'On the rate of gain of information', *Quarterly Journal of Experimental Psychology*, Vol. 4**,** pp.11–26.

Hirche, J., Bomark, P., Bauer, M. and Solyga, P. (2008) 'Adaptive interface for text input on largescale interactive surfaces', *IEEE International Workshop HCI*, pp.164–164.

Hoggan, E., Brewster, A. and Johnson, J. (2005) 'Investigating the effectiveness of tactile feedback for mobile touchscreens', Glasgow Interactive Systems Group, Department of Computing Science, Glasgow University, Glasgow.

ISO (1998) *ISO 9241-11 International Standard, Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs) Part 11: Guide on Usability*, 1st ed., International Organisation for Standardisation, Genève, Switzerland.

Kiljander, H. (2005) *Evolution and Usability of Mobile Phone Interaction Styles*, Publications in Telecommunications Software and Multimedia, Helsinki University of Technology, Finland.

Lindholm, C., Jarnstorm, J. and Kiljander, H. (2003) *Mobile Usability: How Nokia Changed the Face of the Mobile Phone*, McGraw-Hill, New York.

McDaniel, S. (2009) 'What's your idea of a mental model?', [online] http://www.boxesandarrows.com/view/whats_your_idea_of_a_mental_model (accessed 18 April 12).

Moon, M. and Millison, D. (2000) *Firebrands: Building Brand Loyalty in the Internet Age*, Osborne/McGraw-Hill, California, USA.

Nielsen, J. (1993) *Usability Engineering*, Academic Press, San Diego, California, USA.

Page, T. (2011) 'Interaction and usability considerations in the design of mobile phones', *J. Design Research*, pp.5–6.

Paul, R. (2009) 'Dream(sheep++): a developer's introduction to Google Android', *ArsTechnica*.

Rubin, J. (1994) *Handbook of Usability Testing*, John Wiley & Sons, Inc., New York.

Ryu, Y.S. and Smith-Jackson, T.L. (2006) 'Reliability and validity of the mobile phone usability questionnaire (MPUQ)', *Journal of Usability Studies*, Vol. 2, No. 1.

Saffer, D. (2008) *Designing Gestural Interfaces*, O'Reilly Media, Inc., Sebastapol.

Shneiderman, B. and Plaisant, C. (2005) *Designing the User Interface*, 5th ed., Pearson Education, Boston.

Vaananen-Vainio-Mattila, K. and Ruuska, S. (2000) *Designing Mobile Phones for Consumer Needs at Nokia. Information Appliances and Beyond: Interaction Design for Consumer Products*, Academic Press, San Diego, California.

Weiss, S. (2002) *Handheld Usability*, Wiley& Sons Ltd, Chichester.