Ten male colleagues took part in our lab-study about mobile texting

Niels Henze
University of Oldenburg
Escherweg 2
26121, Oldenburg, Germany
niels.henze@uni-oldenburg.de

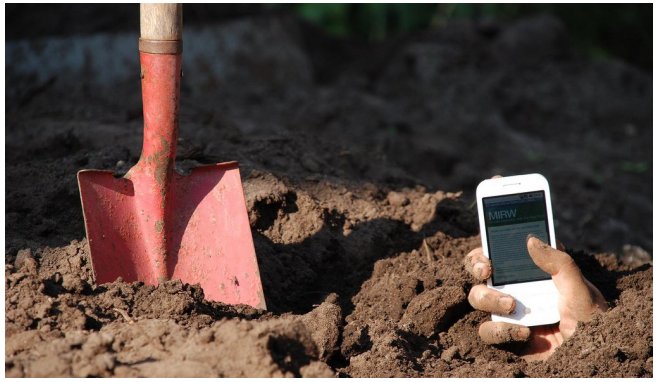

Figure 1: There might be unexpected subjects needing a text entry technique.

CHI 2012, May 5–10, 2012, Austin, TX, USA.
Workshop on Designing and Evaluating Text Entry Methods

Abstract
There will be one billion smartphone users in 2013. About half of them will be female and smartphone adoption in developing countries is one of the market's driving forces. Text entry is a fundamental task to use smartphones and thus a good share of the population will use text entry on a mobile phone. Mobile HCI research, however, focusses on text entry techniques for male students from western countries. As most of our students are young male computer scientists we usually do not study our real target group. As mobile phones are rarely used in labs we also do not study our real context. In this paper we whine about current research practices and describe our approach towards studies with high external validity using games published in mobile application stores. It is argued that existing practices can be supplemented by such approaches.

Keywords
text entry, mobile interaction, validity, large-scale, virtual keyboard, public study, whining

ACM Classification Keywords
H.5.2 [User Interfaces]: Evaluation/Methodology

General Terms
Design, Human Factors, Experimentation

Introduction

Mobile text entry is big. Portio Research, for example, claims that 6.9 trillion SMS messages were sent in 2010 and that SMS traffic is expected to break 8 trillion messages in 2011 [11]. This means that over 250 thousand SMS messages are sent every second. While not all these messages have been handcrafted on a mobile phone it still means a lot of typing. But SMS messages are not the only reason for text entry on mobile phones. Typing mails, typing URLs, entering a Facebook status, searching in the address book – all these activities require text entry and they all are not only common today but will become more and more common. Access to mobile networks is available to 90% of the world population and there was an estimated 5.3 billion mobile subscriptions in 2010 [8]. In particular the developing world is increasing its share of subscriptions from 53% at the end of 2005 to 73% at the end of 2010 [8].

The numbers are clear – mobile text entry is not something that is mainly used by young male geeks that study computer science to get an academic degree. Mobile text entry is a skill that will be required by the majority of the global population in the not so distant future. However, as we will outline in the following section, text entry researchers focus on developing techniques for these geeks. Furthermore, most SMS messages are not sent from HCI research labs and it might be expected that other activities requiring mobile text entry also rarely happen in the lab.

One reason why mobile text entry is mainly studied with such homogeneous samples in the lab is because HCI researchers try to conduct studies where all contextual factors are controlled to isolate certain aspects. Especially when conducting controlled experiments we aim to ensure that the obtained measurements are due to our manipulation and not due to other factors. Well designed and executed, such experiments lead to studies with high internal validity. Experiments, however, can have a high internal validity and be very reliable without necessarily having much to do with real life behaviour [4]. In particular in mobile HCI the participants' background and contextual factors can have a large influence on the behaviour. Mobile devices' screens are, for example, hardly readable in sunlight, typing performance is affected by walking speed, and interaction is affected by fragmented attention. Therefore, mobile HCI studies often lack external validity. We cannot be sure that the findings can be generalized to the average human being using mobile devices in the real world.

Studying mobile text entry

To get an overview about current mobile text entry research we selected eight researchers from the text entry research community (the eight workshop organizers of course). For each researcher we retrieved the most recent paper regarding mobile text entry that is publicly available. Unfortunately, only six researchers have a recent publication in the mobile text entry domain (sorry Brian and Annalu). Table 1 provides an overview about the participants and the context of the studies described in the six papers.

subjects	females	age	context	paper
8	?	?	lab	Kristensson et al. [10]
14+8	?	?	lab	Clawson et al. [2]
20	11	18-50	lab	Dunlop et al. [3]
6+2	?+0	?	lab	Isokoski et al. [9]
44	?	?	field	Vertanen et al. [12]
6	2	21-28	lab	Wobbrock et al. [13]

Table 1: The participants of the studies described in six selected mobile text entry research papers.

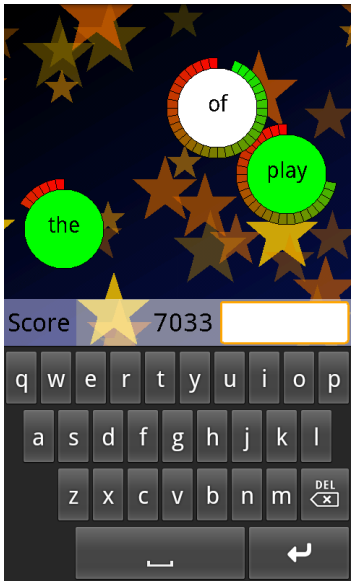

Figure 2: Screenshots of the mobile typing game *Type It!* that is available in the Android Market via http://tiny.cc/type_it.

Two papers detail the subjects age ranges [13, 3] but no paper reports the participants average age. The same papers report the participants gender and another paper reports the participants' gender for one of two studies [9]. While it is not always clear we assume that only one paper reports results from the field [12] (by extracting mails written on a Blackberry from a text corpus). Thus, five of the six studies about mobile text entry have been conducted in the lab. Little is reported about participants' background but one paper highlights that 'Not all were students, and half were *not* technology majors' [13]. We could imagine that for some of the other studies, students that were technology majors might have participated.

The small collection of papers cannot provide a complete picture of current research in the field. There are longitudinal studies, studies in the field, and studies with participants from around the globe in the workshop organizers' publication lists and there are also other researchers in this field. In addition, we are certainly not in the position to point a finger at the field considering the studies we conducted ourselves. Still, the small collection of papers from high class researchers might serve as an indicator that mobile text entry research is often less mobile than desired and that the samples are biased.

Towards high external validity

In contrast to most previous work, in [7] we aimed at observing and manipulating the touch behaviour of a diverse sample that use a large number of devices in various contexts. To collect a large amount of keystrokes on a virtual keyboard we developed a mobile typing game (see Figure 2) and published it in the Android Market. Our approach thus allows studying a large number of users with varying backgrounds in realistic contexts with their own devices (low internal validity due to a high variance but high external validity). This allowed analysing the typing performance of users whose behaviour might have been significantly altered in a very controlled setting.

Using the game we collected 47,770,625 keystrokes that have been produced by 72,945 installations of the game. We analysed the collected data and identified approaches that might influence the typing behaviour positively. We developed a function that compensates a systematic skews found in the touch distribution, shift the keys' labels to the upper part of the keys, and show the position where the user's finger lift-off the screen using a simple dot. To compare the three approaches we conducted an experiment by publishing an update of the game. We collected additional data from 13,013 installations that contributed 6,603,659 keystrokes. We found that showing the users where they touch using a dot improves the error rate but decreases the typing speed. We further show that the adapted shift function that we derived from our observation improves the performance by 2.2% and decreases the error rate by 9.1% compared to a standard Android keyboard.

Our approach has obvious and not so obvious limitations. While not all related work reports participants' background, we simply do not know much about our participants. We are sure that the data has not been produced in a lab but we do not know where it has been produced. The participants are certainly diverse but the sample is biased due to self-selectiveness. We studied users that like to play a particular game and this might not be a good sample of the population. Therefore, the conducted studies have a low internal validity but, compared to common lab studies, we assume that they have a very high external validity. Our approach has the advantage that the results are based on a large number of users, the partici-

pants are likely representative for Android users, and the data has been collected in real life contexts.

Discussion

The perfect study would have a high external validity and a high internal validity. A large and perfect sample of the population would perform tasks in a set of representative contexts over a long period. The experimenter would exactly know the participants background and would assign the contexts. Such a study, however, certainly requires resources that are not available in the average HCI group. Thus, we have to cope with studies that have limitations.

Studies with a high external validity typically have a low internal validity and vice a versa. Studies with a high external validity might tell about users' realistic behaviour but they tell little about what causes this behaviour. We believe that studies with high internal validity that control all factors are required to explain the behaviour observed in studies with high external validity. Thus, both types are absolutely essential but it seems to us that the mobile text entry community might focus too much on one type. For our described studies [7] we use one particular approach. We and others (e.g. [5, 6, 1]) also used this approach for studies in other domains. This approach is, however, obviously not the answer to all research questions. We therefore believe that other approaches for studies with high external validity need to be explored.

References

[1] M. Böhmer, B. Hecht, J. Schöning, A. Krüger, and G. Bauer. Falling asleep with angry birds, facebook and kindle: a large scale study on mobile application usage. In *Proc. MobileHCI*, 2011.

[2] J. Clawson, K. Lyons, A. Rudnick, R. Iannucci Jr, and T. Starner. Automatic whiteout++: correcting mini-qwerty typing errors using keypress timing. In *Proc. CHI*, 2008.

[3] M. Dunlop and F. Taylor. Tactile feedback for predictive text entry. In *Proc. CHI*, 2009.

[4] A. Field and G. Hole. *How to design and report experiments*. Sage publications London, 2003.

[5] N. Henze, M. Pielot, B. Poppinga, T. Schinke, and S. Boll. My App is an Experiment: Experience from User Studies in Mobile App Stores. *IJMHCI*, 2011.

[6] N. Henze, E. Rukzio, and S. Boll. 100,000,000 taps: analysis and improvement of touch performance in the large. In *Proc. MobileHCI*, 2011.

[7] N. Henze, E. Rukzio, and S. Boll. Observational and experimental investigation of typing behaviour using virtual keyboards on mobile devices. In *Proc. CHI*, 2012.

[8] International Telecommunication Union. The world in 2010. 2010.

[9] P. Isokoski, B. Martin, P. Gandouly, and T. Stephanov. Motor efficiency of text entry in a combination of a soft keyboard and unistrokes. In *Proc. NordiCHI*, 2010.

[10] P. O. Kristensson and K. Vertanen. Asynchronous multimodal text entry using speech and gesture keyboards. In *Proc. INTERSPEECH*, 2011.

[11] Portio Research. *Mobile Messaging Futures 2011-2015 – Analysis and Growth Forecasts for Mobile Messaging Markets Worldwide: 5th Edition*. 2011.

[12] K. Vertanen and P. Kristensson. A versatile dataset for text entry evaluations based on genuine mobile emails. In *Proc. MobileHCI*, 2011.

[13] J. Wobbrock, D. Chau, and B. Myers. An alternative to push, press, and tap-tap-tap: gesturing on an isometric joystick for mobile phone text entry. In *Proc. CHI*, 2007.