Author(s): Stian Reimers and Neil Stewart

Article Title: Using Adobe Flash Lite on mobile phones for psychological research: Reaction time measurement reliability and interdevice variability

Year of publication: 2008

Link to published version: http://dx.doi.org/10.3758/BRM.40.4.1170

Publisher statement: None

Running Head: FLASH LITE ON MOBILES

Using Adobe Flash Lite on Mobile Phones for Psychological Research: Reaction Time Measurement Reliability and Inter-Device Variability

Stian Reimers[1] and Neil Stewart[2]

[1]University College London, United Kingdom

[2]University of Warwick, United Kingdom.

Stian Reimers

Department of Psychology

University College London

26 Bedford Way

London WC1H 0AP

England

Tel: +44 207 679 5317

Fax: +44 207 436 4276

s.reimers@ucl.ac.uk

Reimers, S., & Stewart, N. (2008). Using Adobe Flash Lite on mobile phones for psychological research: Reaction time measurement reliability and inter-device variability. *Behavior Research Methods, 40,* 1170-1176.

Abstract

Mobile telephones have significant potential for use in psychological research, possessing unique characteristics – not least their ubiquity – that may make them useful tools for psychologists. We examine whether it is possible to measure reaction times (RTs) accurately using Adobe Flash Lite on mobile phones. We ran simple and choice RT experiments on two widely available mobile phones, a Sony Ericsson W810i and a Nokia 6110 Navigator, using a wireless application protocol (WAP) connection to access the internet from the devices. RTs were compared within-subjects with those obtained using a linux-based millisecond-accurate measurement system. Results show that measured RTs were significantly longer on mobile devices, and that overall RTs and distribution of RTs varied across devices.

Using Adobe Flash Lite on Mobile Phones for Psychological Research: Reaction Time Measurement Reliability and Inter-Device Variability

In recent years, one of the biggest socio-technological developments has been the widespread adoption of mobile telephones among the majority of the Western World. In the United Kingdom, in 1998 under 30% of households owned a mobile phone (Office for National Statistics, 2008). By 2006 the proportion had risen to 80%, and now in 2008, the number of mobile phones exceeds the population. It is estimated that there are over two billion mobile phones globally (Central Intelligence Agency, 2008).

Although the primary use of a mobile phone is to make and receive calls and text messages, many phones now include other capabilities such as digital cameras, audio players, radios, games and organizational tools like alerts and planners. In the ten years to 2008, it has become increasingly easy and popular to use a Wireless Application Protocol (WAP) connection to access the web from mobile devices. This has made it relatively simple to download games and applications across the internet, check email, and search for information on specially-designed websites – such as train times, traffic conditions, or local facilities like restaurants or petrol stations. Mobile internet access has also been developed to tie in with more traditional PC-based access; for example, it is now possible to place a bid on the auction site eBay (wap.ebay.com) or update one's status on the popular social networking site Facebook (m.facebook.com) from a mobile phone costing under $100, using WAP. Thus, at the time of writing it appears that internet use on mobile phones is becoming mainstream.

Existing Research using Mobile Devices

A WAP-enabled mobile phone is essentially a small, cheap, handheld computer which can send and receive data wirelessly across the internet. Put in these terms, mobile phones may appear attractive as devices for running psychology experiments. However – perhaps unsurprisingly, given the speed with which mobile phone capabilities have developed – there has so far only been limited psychological research using mobile devices. Shephard , Kho, Chen, and Kosslyn (2006) developed a system, MiniCog, for administering cognitive tests primarily to astronauts, but also 'in the field' generally, using a personal digital assistant (PDA). This system was capable of measuring RTs on cognitive tasks, and the authors note that the RTs should be accurate to around 10 ms, although they indicate that they had not conducted any systematic tests of RT accuracy. MiniCog was also designed as a package which would be downloaded, and data uploaded, via a host computer, rather than direct contact with a server across the internet.

A different approach to using mobile devices for psychological research had been taken by Hogarth and colleagues (e.g., Hogarth, Portell & Cuxart, 2007), who used mobile phone SMS text messages to implement an experience sampling method of risk perception measurement. They sent text messages to participants at random times during the day. When participants received the message they had to make a note what they were doing, and answer questions about the potential risks that the activity they were engaging in entailed. We presume that – as two of the questions had open ended format – that they did not submit their responses using a mobile phone.

Text messaging has also been used in classroom experiments. Cheung (2008) used a sophisticated system for recording, processing and responding to SMS text messages from students in a class. He used the system to run a version of the ultimatum and public goods games. Thus, half the students in the ultimatum game setup sent a text message with the amount they would offer; the other half sent a text message with the minimum amount they would accept. Proposers and responders were randomly paired, and each received a text message informing them whether agreement had been reached or not. Summary statistics from the game were also generated using a macro, and the results displayed for the class to see.

Advantages and Disadvantages of Mobile Phones in Psychological Research

The three examples above show some of the major advantages of using mobile devices for psychological research. Participants can undergo testing ‘in the field’ in more naturalistic situations, and with little notice or effort. The ubiquity of mobile phones means that large numbers of people can participate in an experiment without the need to allocate a device to collect responses.

However, we suggest that there are additional reasons that using mobile devices might be of benefit to psychologists. Much of the literature on the advantages and disadvantages of web-based testing can also be applied to running experiments on mobile devices. Thus, mobile phone testing allows one to test participants who are not physically present in the lab, and to reach a broader demographic than those who normally participate in lab-based research. In particular it may be easier to run experiments with particular subgroups of the population, such as those with limited mobility, those living far from a researcher’s institution, and others who would be unwilling to go to the effort of booking an appointment and traveling to the laboratory for testing. It also gives people more flexibility in when they participate, and allows them to participate from the familiar surroundings of their own home.

As well as the potential advantages that phone-based testing shares with traditional computer-based web-testing, mobile devices have a number of other advantages. As we saw in the examples of Shephard et al. (2006) and Hogarth et al. (2007), they allow testing to be done ‘in the field’, where it would be harder to test using a traditional computer-based set-up. Shephard et al. use the extreme example of testing in space, but also suggest that there may be down-to-earth scenarios that could be explored using mobile devices. Examples might include testing people in extreme situations, such as before and after a sporting event, parachute jump or a presentation to hundreds of people; tracking the effects of, say, fatigue in drivers, people doing shift work, or people with high-stress jobs such as air traffic control or city trading; examining the effects of alcohol, nicotine or illegal drugs on cognition, or examining how people's quotidian activities affect cognitive or emotional states.

Mobile phones also have the advantage that they tend to be switched on and carried by their owner most of the time. It would therefore be possible – as Hogarth et al. did – to contact people or ask them to respond at a specific, random time. Thus, the effect of circadian rhythms could easily be investigated. Alternatively, memory stimuli with inter-item spacings of hours or days could be delivered using mobile phones.

Finally, Cheung's (2008) setup could conceivably be extended, using a Bluetooth or WAP connection instead of SMS messaging to allow lecturers to run experiments in many areas of cognition , presenting visual stimuli and recording response times, showing students how particular 'classic' experiments were implemented, and the type of results obtained. Whole practical classes could be run without the need for a large number of computers.

Clearly, mobile phones are not a panacea, and there are also significant potential drawbacks that might need to be overcome. These include demographics skewed towards the younger and more affluent members of society, particularly for experiments that require WAP, Java, or Flash. There are also problems of limited geographical network coverage – participants might miss a message or request – or may fail to save their data – if they are out of signal range. Responses may also be interrupted by incoming calls or messages. The cost of sending and receiving data using WAP can be very high, largely precluding the use of audio and video stimuli. The screen size is relatively small, keys are also small and fiddly, and free-text responses are, at best, difficult to give.

In summary, although there are several potential problems to overcome for some types of phone-based testing, there appear to be enough potential advantages to take the notion of phone-based testing seriously.

RT Experiments on Mobile Devices

Having discussed the use of mobile devices generally, we now focus specifically on running experiments that measure millisecond RTs. We know of three broad ways of running RT experiments on mobile devices. One is to download and install a program that runs on the phone's operating system. For example, there are many applications that have been developed for the Symbian operating system found on many brands of mobile device. This is broadly equivalent to downloading and installing an executable file on a PC. The advantages of this technique are that the program, once installed, has access to a wide range of the device's functionalities, such as establishing web connections, sending text messages and operating the camera. A major disadvantage is that people are often reluctant to install software that is so unconstrained. Further, a piece of software will only run on phones that use the operating system for which it was designed.

A second way would be to install a free-standing application written in Java that has limited access to the phone's data and processes. The majority of mobile phone games are written in Java, and users generally download an application, respond to a prompt asking if they wish to install the program, and the application is saved locally on the user's device. The functionality of Java is sandboxed, so that an application cannot access other information on the device, or use all of its functionality. Thus, Java trades off perceived security threats with decreased functionality.

Third and finally, and the focus of our attention here, it is possible to run experiments using Flash Lite. Flash Lite differs from Java because many mobile devices are set up to run

Flash Lite applications automatically after download without asking the user if they want to run the code. This makes the experience of completing Flash Lite experiments more ergonomic for the user than Java, and also means that on many devices, Flash Lite applications can be embedded in HTML code, allowing developers to use a combination of Flash and HTML in their applications. Potential drawbacks of Flash Lite are the very restricted access programs have to other information stored, and processes running on the phone – Flash Lite 1.1 cannot use Bluetooth, for example; and that running in a browser may lead to a decrease in performance.

Adobe Flash and Adobe Flash Lite

Adobe's range of Flash products have been used extensively to develop and deliver games, advertisements, and other interactive content over the web, and have been used for running RT-based psychology experiments (see Reimers & Stewart, 2007, for an introduction, and Reimers, 2007; Reimers & Maylor, 2005, for an example). In the past few years one of the major developments in the use of Flash has been the release of Flash Lite, a cut-down version of Flash designed to be run on mobile devices. Flash Lite is included in a number of retail mobile devices. Strategy Analytics (Robinson, 2008) reports that between 2003 and 2007, over 300 million Flash Lite enabled devices were shipped, and this number is predicted to rise to 1.4 billion by the end of 2010. A list of current devices supporting Flash Lite can be found at http://www.adobe.com/mobile/supported_devices/handsets.html.

One advantage of Flash Lite is its similarity to standard Flash. This means that researchers who are used to implementing web experiments in Flash will find it relatively easy to write experiments in Flash Lite. It is also the case that an experiment has only to be written once – although appearances may differ slightly on different devices, all are capable of reading and executing Flash Lite files. This is one advantage that Flash Lite has over, say, Java, which uses specific APIs for a particular device, and might therefore require multiple versions.

We therefore set out to investigate whether it would be viable to run RT psychology experiments using Flash Lite, and in particular, the extent to which currently-available devices could record reaction times consistently and accurately.

The Present Study

In the present study we extend the research we conducted on measuring RTs using Flash over the web (Reimers & Stewart, 2007) to measuring RTs using Flash Lite and WAP on two widely-available mobile phones. We have two aims. The first is to examine generally whether RT measurement is accurate enough on currently available mobile devices for RT-based testing to be worth considering at all. For example, if RTs measured on mobile phones were hundreds of milliseconds longer than under accurate measurement, or were quantized to the nearest 500ms, then testing using mobiles would be unviable. The second aim is to examine the extent to which RT measurement accuracy varies across devices. This is important because if there is significant variation, it will be important to ensure that differences in RT due to hardware are accounted for, and are not assumed to be psychological differences.

To investigate this we took two mobile phones, one very modern, one somewhat older, and compared – within-subjects – RTs measured on each of them with an accurate baseline. Our Baseline Condition, which is known to measure RTs with millisecond accuracy, was programmed in C, and used the same set-up described in Stewart (2006a, 2006b). In the other two conditions a single Flash Lite 1.1 swf (Shockwave Flash) file was run on a Sony-Ericsson W810i and a Nokia 6110. The experiment was designed as more of a field test than a calibration: we tested the same participants in the three different conditions and compared

RTs. This does conflate technological variables – like the measurement accuracy of a particular phone – with ergonomic variables – using thumbs on a mobile phone keypad versus index fingers on a button box – and motivation issues – the baseline setup looked much more like a 'proper' psychology experiment than simply sitting in a quiet room with a mobile phone, and participants may therefore have been motivated to try harder. It is also possible that the smaller size of the stimuli on the screen made the mobile condition harder. Thus, the results of this experiment should not be interpreted as just the physical characteristics of the different devices, but rather the overall differences in RTs one might expect if one ran participants in the different conditions.

Method

Participants

Participants were 22 undergraduate psychology students at the University of Warwick. All participated for course credit; there was also a £20 (US$40) prize for the best performance across all conditions.

Design and Procedure

There were three within-subjects conditions, one Baseline – using the Linux setup described below – and two others – Sony and Nokia – that used the same Flash code, designed to appear as similar as possible to the Baseline condition. The order in which participants used the different devices was counterbalanced.

The system used for the desktop-computer-based experiments (Baseline) was a Tranquil T2, a silent and low cost PC. Stimuli were presented on a Viewsonic VX924 19 inch LCD display. Responses were collected using a custom-made parallel-port button box with two buttons, one for the left index finger and one for the right index finger. The Sony Condition used the Sony-Ericsson W810i in standard UK configuration, with a 220x176 pixel screen, connected to the O2 network. The W810i is a mid-market phone, released in early 2006, which currently retails for around £75 (US$150). The Nokia Condition used the 6110 Navigator, with a 320x240 pixel screen, in standard UK configuration, connected to the Vodafone network. The 6110 is – at the time of writing – a high-end 3G phone, with built-in GPS, and retails for around £200 (US$400).

In the Baseline Condition, the C code was executed from the command line, and the experiment window filled the screen. In the Sony Condition, the experiment was downloaded using WAP, and displayed in the device's browser, which was set to full screen; in the Nokia Condition the Flash was downloaded using WAP and run full-screen from a free-standing version of Flash Player, as Flash Player was not installed as a browser plug-in. In both the phone conditions, the Flash Lite code was downloaded in its entirety before the experiment began. Thus, until data were saved at the end of the experiment, the cellular network was not required.

Participants completed two 30-trial tasks, choice reaction time (CRT) and simple reaction time (SRT), always in that order. In all conditions participants sat alone in a quiet room. At the start of each condition printed instructions were given, which were summarized on the monitor or device screen. At the start of a trial a '+' crosshair appeared on the screen, which was followed by a variable delay of between approximately 1500 and 3000 ms (see below). Following the delay, in the CRT task, the crosshair was replaced by a red or green rectangle, which remained on the screen until a response was made. Immediate feedback ('correct!' or 'wrong!' along with the RT in ms) was given, lasting 1000 ms before the next trial began. All trials were independent: on each trial there was an equal (.5) chance of the rectangle being red or green. Response key mapping was counterbalanced: half of participants pressed the left key if the rectangle was green, and the other half pressed if it was red.

Response key mapping was the same for a given participant in all three conditions. In the SRT task, the crosshair was replaced by a blue rectangle, which remained on the screen until a response was given. Participants could respond using either key from the CRT task, and used the same finger and button throughout the SRT task. The RT in ms was displayed as feedback for 1000 ms before the next trial began. In the Baseline condition, participants were instructed to use their left and right index fingers to respond. In the phone conditions, participants were allowed to choose the digits with which they responded (as would be the case with unsupervised, remote, participants), but were informed that cradling the phone and using left and right thumbs to respond might be most comfortable. (In earlier work we found that most participants naturally ended up responding in this manner, and did not want to add noise to our data from people changing hand positions – and hence possibly altering RTs – part way through the experiment.)

At the end of a set of 30 CRT or SRT trials, all the data from all 30 trials were saved to disk (in the Baseline condition) or transmitted using WAP to a perl script on our server, which handled saving the data (in the mobile conditions). Data in the mobile conditions were not saved on the phone itself.

Implementation

Flash (Sony, Nokia). A 220x176 pixel Flash Lite 1.1 movie was created using Adobe Flash CS3. As Flash Lite 1.1 has very restricted capabilities relative to normal Flash and Flash Lite 2.0, the implementation was slightly different from that in Reimers & Stewart (2007). To create a variable stimulus onset delay we created 16 frames, which were played in sequence. With each frame there was a probability of jumping to a later frame in which the stimulus appeared. Thus in Frame 1 there was a 1/16 chance of a jump to the stimulus display frame; in Frame 2 there was a 1/15 chance, up to frame 15 where there was a 1/2 chance of jumping, and Frame 16 where there was a 1/1 (i.e., certain) chance of jumping. This would, in theory, give a flat distribution of 16 different delays of 1500ms, 1600ms, 1700ms … 3000ms, assuming the theoretical frame rate was achieved.

We used pilot data to generate delays in the Baseline Condition that had approximately the same distribution as those actually seen on the Sony Ericsson, with the shortest of the 16 delay groups set at 1526 ms and the longest set at 3291 ms. We also logged the actual delays on the Sony-Ericsson during the experiment, which came out slightly longer than in the pilot, with the shortest delay category being 1563 ms and the longest delay being 3369 ms. The Nokia, although running the same code as the Sony, had shorter delays, with a shortest delay category of 1421 ms and longest of 3061 ms.

In the stimulus display frame, the movie playhead was stopped, one of two stimuli – a red or green rectangle – was set visible, the other invisible, and a timer polled using the `getTimer()` method. An on key down event listener waited for the '1' or '3' button to be pressed, after which the timer was polled again, and the difference recorded.

C (Baseline condition). The X Windows System was used to run a C program which displayed stimuli (Stewart, 2006b). Responses were recorded on a parallel port button box (Stewart, 2006a) Following Finney (2001), the program was scheduled to the highest priority (`SCHED_FIFO`) so it was not interrupted, had its memory locked (with a call to `mlockall()`) prevent swapping to disk and timings were measured (using the `gettimeofday()`) to the nearest microsecond. RTs were measured from the end of the vertical retrace upon which the stimulus was drawn to the first detection of a close of a button on the parallel port button box

Results

During testing, the remote server from which the Flash versions of the experiment were being accessed, and to which data were being saved, crashed temporarily. This meant that data from participants who were part-way through either of the Flash tasks were lost, and the participants in one time-slot were excused the Flash conditions. (We note here that the crash was due to our dubious choice of internet service provider, rather than failures of Flash or the mobile phone network.) We were left with 20 participants for whom we had data for all three SRT conditions.

In the CRT condition we had 15 participants. The server problems were compounded by – we suspect – the fact that in the Nokia condition, where the experiment was run in using the standalone Flash Player, the data occasionally expired before the web was accessed at the end of the SRT stage. For both SRT and CRT conditions we only use participants for whom we have data for all three device cells.

The group means and standard deviations for the SRT and CRT conditions, along with error rates for CRT responses, are given in Tables 1 and 2. Only RTs between 100 and 1000 ms are included in the analysis (SRT: 98.2% of Baseline trials, 98.0% of Sony trials and 98.7% of Nokia trials; CRT: 99.6% of Baseline trials, 98.9% of Sony trials and 98.9% of Nokia trials).

To test the significance of any differences in RT and SD for the SRT data, we ran repeated-measure ANOVAs with condition (Baseline, Sony, Nokia) as a within-subjects variable. (With 24 counterbalancing permutations it was not possible to include counterbalancing as a between-subjects variable.) There was a significant effect condition on mean RTs, $F(2, 38) = 79.1$, $p < .001$, median RTs, $F(2, 38) = 123.8$, $p < .001$, and variable effects on measures of variability: SD, $F(2, 38) = 3.38$, $p = .048$, interquartile range, $F(2, 38) = 6.62$, $p = .003$, and full range, $F(2, 38) = 1.01$.

We conducted an identical analysis for the CRT data. There was a significant effect condition on mean RT, $F(2, 28) = 28.7, p < .001$, median RT, $F(2, 28) = 43.9, p < .001$, but not on any of the measures of variability: SD, $F(2, 28) = 1.04$, interquartile range, $F(2, 28) = 0.81$, or full range, $F(2, 28) = 0.62$.

We next examined responses for evidence of quantizing or other irregularities in the distribution of RTs. As in Reimers and Stewart (2007), we plotted a cumulative frequency graph of RTs in the three conditions (Figure 1).

Figure 1 about here

The baseline and Nokia RTs follow the expected smooth sigmoid shape reflecting an underlying approximately normal distribution. In contrast, the Sony RTs show approximately 20ms plateaux around 250, 370, and 480 ms, followed by an increased gradient shortly afterwards. This suggests that the Sony delays processing of a response at certain times.

Discussion

Our investigation showed that RTs measured on two mobile phones were, on average, 60-80 ms longer than those using a millisecond-accurate linux system. RTs were also device-specific: Those measured on the Nokia Navigator were 20-30 ms shorter than those on the Sony-Ericsson W810i. In addition, RTs on the Sony-Ericsson appeared to be substantially quantised.

Cause of Differences

There was a clear difference in measured RTs between the two phones, and between each phone and the Baseline Condition. We ran this experiment as a field-test comparison of devices rather than a calibration, and so did not attempt to distinguish between the technical performance of the phones and the ergonomic effects of using thumbs to respond, small display size, posture, and other differences between computer and phones. However, the fact that RTs were so much faster on the – more powerful – Nokia than the – less powerful – Sony suggests that technical performance was an issue.

The quantising effects on the Sony but not the Nokia also suggest performance limitations. The quantising appears to occur at around the same rate as the Flash frame rate (and hence screen update rate), which suggests that they might be the results of the processor being overloaded when running screen updates and other housekeeping activities at the start of each frame. If this is the case, it appears that the Nokia is powerful enough to avoid any quantising of RTs on entering a frame. Whether these effects are due to differences in processor power or running Flash inside a browser on the Sony is not clear. Although the quantising effects are not desirable, averaging across many trials would largely remove any effects. Clearly, the quantising could conceivably be specific to the way in which we programmed the experiment, and there may be other implementations that avoid this problem. However, it is the fact that the quantising only occurred on one of the two phones running the same code that suggests RTs recorded on different devices may exhibit qualitative as well as quantitative differences.

Conclusion and Recommendations for Mobile RT Research

Overall, we think that mobile phones may be used for running RT-based research. Mean RTs were longer, but RT SDs were only slightly increased, suggesting that – on the Nokia at least – it would be relatively easy to detect within-subject changes to RTs, or – controlling for the device used – differences across between-subjects conditions. It is also clear that mobile devices are becoming increasingly powerful, and this may well lead to the attenuation of differences in RTs measured across devices. Finally, some of these effects may be specific to using Flash Lite – embedded in a browser in the case of the Sony condition – to run experiments. It would be interesting to compare RTs measured using Flash with those using Java.

One key implication of these results for running experiments on mobile devices is that there is a risk of confounding device capabilities with variables that are under investigation. For example, if older people tend to own less recent mobile phones, any apparent age-related decline in RT may in fact be caused by the device rather than the person. However, there are several ways round the issue. One is to note the device used to run an experiment – which in Flash Lite 1.1 and above can be done using the `GetDevice` method – and include it as a variable in any analysis. A second is to use a within-subjects design, so that absolute effects disappear. A final – albeit crude – option might be to run an internal calibration, measuring how long it takes to execute a particular piece of looped code, on a given device to gain a measure of the processing resources it has available. If experiments and analyses are designed to take into account differences in device performance and the slight distortions to RT means and SDs, we see no reason why mobile phones should not be used in RT-based psychological research.

References

Central Intelligence Agency (2008). *The CIA world factbook, rank order, telephones, mobile cellular.* Retrieved February 2, 2008 from https://www.cia.gov/library/publications/the-world-factbook/rankorder/2151rank.html

Cheung, S. L. (2008). Using mobile phone messaging as a response medium in classroom experiments. *Journal of Economic Education, 39,* 51-67.

Finney, S. A. (2001). Real-time data collection in Linux: A case study. *Behavior Research Methods, Instruments, and Computers, 33,* 167-173.

Hogarth, R. M., Portell, M., & Cuxart, A. (2007) What risks do people perceive in everyday life? A perspective gained from the experience sampling method (ESM). *Risk Analysis, 27,* 1427-1439.

Office of National Statistics (2008). Consumer Durables: Consumer durables ownership increases. Retrieved February 2, 2008 from http://www.statistics.gov.uk/CCI/nugget.asp?ID=868&Pos=5&ColRank=2&Rank=480

Reimers, S. (2007). The BBC Internet Study: General methodology. *Archives of Sexual Behavior, 36,* 147-161.

Reimers, S., & Maylor, E. A. (2005). Task switching across the life span: Effects of age on general and specific switch costs. *Developmental Psychology, 41,* 661-671.

Reimers, S., & Stewart, N. (2007). Adobe Flash as a medium for online experimentation: A test of RT measurement capabilities. *Behavior Research Methods, 39,* 365-370.

Robinson, S. (2008). *Flash-enabled handset forecast: 91% global growth in 2007 as Flash Lite gains momentum outside Japan.* Abstract Retrieved February 2, 2008 from http://www.strategyanalytics.net/default.aspx?mod=ReportAbstractViewer&a0=3727.

Shephard, J. M., Kho, S., Chen, J., & Kosslyn, S. M. (2006). MiniCog: A method for administering psychological tests and experiments on a handheld personal digital assistant. *Behavior Research Methods, 38,* 648-655

Stewart, N. (2006a). A PC parallel port button box provides millisecond response time accuracy under Linux. *Behavior Research Methods, 38,* 170-173.

Stewart, N. (2006b). Millisecond accuracy video display using OpenGL under Linux. *Behavior Research Methods, 38,* 142-145.

Author Note

Stian Reimers, University College London, and Neil Stewart, Department of Psychology, University of Warwick, Coventry, UK.

We are grateful to Sarah Reimers for trusting us with the Nokia.

This research and Stian Reimers were supported by a fellowship from the Centre for Economic Learning and Social Evolution (ELSE). Neil Stewart was supported by ESRC grants ESRC RES-062-23-0952, ESRC RES-000-22-2459, and ESRC RES-000-23-1372.
Correspondence concerning this article should be addressed to Stian Reimers, Department of Psychology, University College London, WC1H 0AP, UK. E-mail may be sent to s.reimers@ucl.ac.uk.

Table 1

Differences in Average and Variability of SRT Across Testing Condition (All in ms).

Condition	Mean of participant RT means	Mean of participant RT SDs	Mean of participant RT medians	Mean of participant interquartile ranges	Mean of participant ranges
Baseline	214.4	53.2	202.6	34.8	268.0
Sony	286.0	56.9	276.0	45.3	272.7
Nokia	266.9	75.3	242.0	60.6	330.3

Table 2

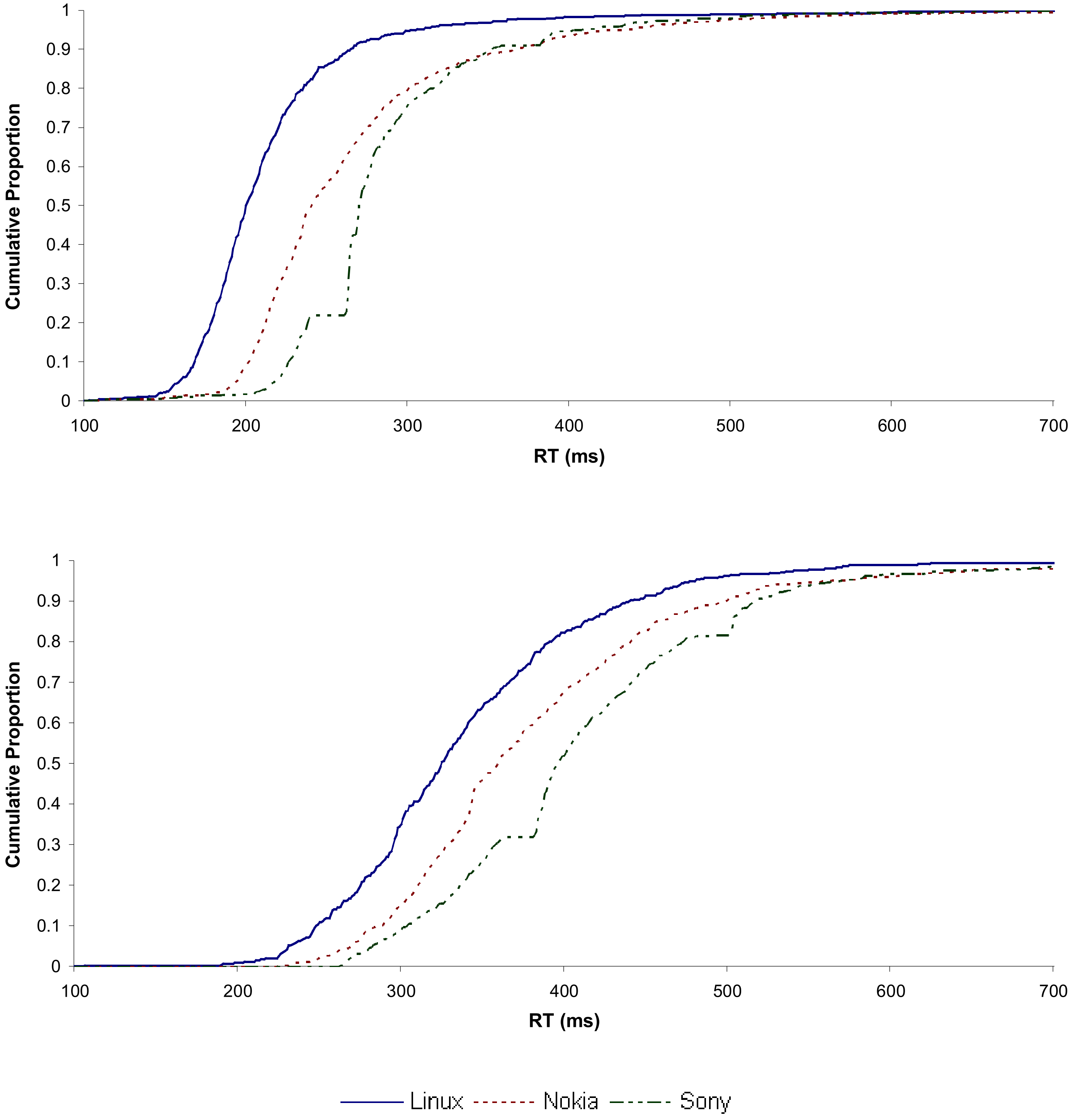
Differences in Average and Variability of CRT Across Testing Condition (All in ms), and Error Rate.

Condition	Mean of participant RT means	Mean of participant RT SDs	Mean of participant RT medians	Mean interquartile range	Mean range	Error rate
Baseline	339.7	74.7	329.2	78.3	342.1	4.0%
Sony	411.8	73.1	405.2	83.2	307.1	5.6%
Nokia	383.4	86.6	366.5	88.9	366.6	4.7%

Figure Caption

Figure 1. Cumulative frequency of trials as a function of RT (ms) for millisecond-accurate baseline, and the two mobile devices. Upper panel shows SRT; lower panel shows CRT.

Figure 1


Cumulative Proportion
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
100
200
300
400
500
600
700
RT (ms)
Cumulative Proportion
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
100
200
300
400
500
600
700
RT (ms)
Linux
Nokia
Sony