Usable Gestures for Mobile Interfaces: Evaluating Social Acceptability

Julie Rico and Stephen Brewster
Glasgow Interactive Systems Group, Department of Computing Science
University of Glasgow, Glasgow, G12 8QQ, UK
{julie, stephen} @dcs.gla.ac.uk

ABSTRACT
Gesture-based mobile interfaces require users to change the way they use technology in public settings. Since mobile phones are part of our public appearance, designers must integrate gestures that users perceive as acceptable for public use. This topic has received little attention in the literature so far. The studies described in this paper begin to look at the social acceptability of a set of gestures with respect to location and audience in order to investigate possible ways of measuring social acceptability. The results of the initial survey showed that location and audience had a significant impact on a user's willingness to perform gestures. These results were further examined through a user study where participants were asked to perform gestures in different settings (including a busy street) over repeated trials. The results of this work provide gesture design recommendations as well as social acceptability evaluation guidelines.

Author Keywords
Social acceptability, gesture interfaces, design recommendations, evaluation methodology.

ACM Classification Keywords
H.5.2 [User Interfaces]: *Input devices and strategies.*

General Terms
Human Factors.

INTRODUCTION

While gesture-based mobile interfaces provide promising new interaction techniques for creating more natural interfaces [18], these kinds of interfaces require users to adopt new behaviors that might be embarrassing or disruptive in certain mobile settings. For example, flicking the phone in a whip like motion in order to send a message might be satisfying in some settings and embarrassing in others. Individuals are constantly aware of their surroundings and how they are presenting themselves at any given time, and mobile phone usage is part of this. These kinds of interactions take place in public locations, and must be designed to account for the presence of spectators [14]. As mobile phones become more integrated into daily life, they become an important aspect of one's social appearance. Individuals must decide when and where they will use mobile phones and how they will interact with them. With respect to gesture-based interfaces, users must evaluate if their motivation to use the technology outweighs the risk of looking strange or making a social blunder. In the face of such issues, gesture-based interfaces must be designed with an awareness of social context and social acceptability.

Although social factors have an influence on technology acceptance [13], little work has been done to determine the social factors of gesture acceptance and even less has been done to produce recommendations for the design and evaluation of socially acceptable gestures. In this paper we describe two studies that begin to determine the factors of social acceptability and some methods for evaluating the acceptance of gestures in social situations using video prototypes and field studies. This method allows designers to choose acceptable gestures before investing expensive implementation efforts developing gestures that can be recognized but that users might never accept. Gesture-based interfaces should not be built based only on the state-of-the-art of gesture recognition software, but also consider the limitations of user willingness to perform gestures in the proposed usage contexts.

GESTURES

Previous research involving gestures and gesture-based interfaces has primarily revolved around defining and classifying gestures and recognizing gestures successfully. These two issues form major challenges in the design of gesture-based interfaces [18]. The former because the lack of a widely accepted definition and classification of gestures that is applicable to multimodal interfaces has led to interfaces with arbitrarily chosen gesture sets or gesture sets chosen based solely on available technology. Successful gesture recognition is a major challenge because the ability to use natural movements in gesture interfaces is greatly limited by current technology. For example, the segmentation issue [15] of knowing when gesture actions begin and end still presents a difficult challenge for system designers in creating a successful interface.

CHI 2010, April 10–15, 2010, Atlanta, Georgia, USA.

Gesture definitions vary greatly, leading to gesture-based systems that vary greatly in form and usability. Cassell defines gestures as body movements that occur during speech [4] and gesture based systems that utilize this definition have demonstrated positive effects of using speech and gesture together. For example, Put That There [1] used both speech and gesture to create a richer interaction than just speech or gesture alone. Kendon defines gestures as voluntary and expressive movements of the body [12]. Väänänen and Böhm [17] discuss how gesture definitions that are commonly used in gesture-based interfaces are often limited in order to reflect the limitations of technology. These definitions describe arbitrary sets of hand or body movements that may not have any intrinsic meaning or regular usage in daily life. For example, the lexical gesture interface proposed by Mo and Neumann [20] presents a set of gestures that can be used in a gesture-based system. An open palm and closed fist were used as base positions and a straight, half bent, or fully bent index finger were used as finger states. Although this type of interface provides a clearly defined set of gestures, Cassell argues that individuals have as much natural ability to use these kinds of gestures as DOS commands [4]. For the purpose of describing this research, we will define gestures as any deliberate movement of the body used with the intention to communicate some idea. This movement could be used during communication between two people in natural conversation, or communication between a human and a computer as part of an interface (our focus in this paper).

SOCIAL ACCEPTABILITY

Individuals evaluate social acceptability when the motivations to use technology compete with the restrictions of social settings. Individuals make decisions about the social acceptability of their actions by gathering information about their current surroundings and using their existing knowledge [10]. Appropriate actions are then carried out and feedback is gathered through the reactions of observers. These actions should therefore be viewed as a performance [10], an intentional action executed by an individual with the awareness of spectators. The social acceptability of technology usage is not a simple matter of embarrassment or politeness, but a combination of factors ranging from appearance, social status [10], to culture [3]. The process of performing actions and gathering feedback is circular, with ideas naturally changing over time and each experience helping individuals make better decisions in the future.

Gesture-based interfaces face adoption issues as they require users to evaluate a whole new set of actions and define new standards of social acceptability. While devices using screen-based gestures such as the Apple iPhone have seen widespread use, gesture-based interfaces that utilize gestures away from the screen have not. Screen based gestures do not change the way the device is held or used so, for the most part, do not conflict with the existing perceptions of mobile phone usage in social settings. Devices using gestures such as wrist rotation, head pointing, or device squeezing have had little or no widespread usage. However, this does not mean that asking people to begin using unfamiliar or strange actions prevents technology adoption. For example, Bluetooth headsets for hands-free mobile phone calls have had widespread social acceptance. Even though these devices require users to act outside of their normal behavior, namely by appearing to talk to themselves, they are widely popular. The acceptance of headsets has grown out of experiences seeing others using them or using them personally with some benefit. This is a good example of how socially unacceptable actions can become acceptable through continued exposure.

While significant work has been completed on the technical side of gesture-based interfaces with gesture recognition algorithms and sensing, the usage of gestures in the real world has had little attention. One of the few pieces of work to discuss social acceptability was done by Ronkainen *et al.* [15]. They chose gestures for use in a usability study based on a survey of social acceptance. This survey asked users to watch a video of a gesture usage scenario and decide if they would be willing to use the presented feature. Possible responses were: *Yes, it's fun / Yes, it's useful / Yes (other reason) / No, it looks silly / No, it's not useful / No (other reason)* [15]. Examining user opinions of gestures before implementation allowed for unacceptable gestures to be ruled out early, resulting in a more effective usability study. This study was limited, however, because gestures were coupled to locations and tasks rather than evaluated on their own. This study identifies that social factors influence user willingness to use gestures, but does not provide insight into what those factors are or how they might influence user opinions. Social factors have been identified as an important aspect of technology adoption [13], and its absence from the Technology Acceptance Model (TAM) [7] is one of TAMS's biggest weaknesses [13]. TAM describes the adoption behaviors of information system users based on the Theory of Reasoned Action [8], which describes behaviors based on behavioral intention, subjective attitude, and subjective norm. Malhotra and Galletta [13] describe their method for integrating social factors into TAM by adding the element of psychological attachment, or the influence of social factors. However, this work does not address what those factors might be or how they might influence technology adoption.

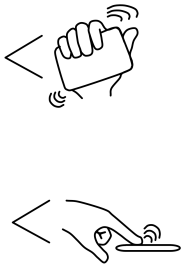
In the rest of the paper we present two studies that begin to investigate some factors of social acceptability while also testing our evaluation methodology. These studies look at two different ways of assessing acceptability so that we can understand what makes a gesture acceptable and what are the best methods to analyze acceptability. Although there are many possible factors that should be examined in social acceptability, such as task pairing, cultural differences, gesture performance with respect to energy, and personality traits, we chose to look first at the factors of location and audience, influenced by Goffman's arguments [10] that these factors play a large role in social behavior. First, a

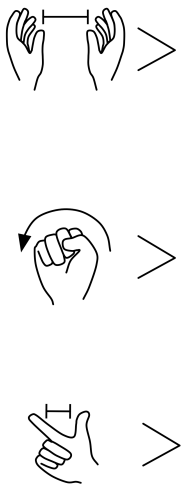
gesture survey was completed in which respondents could assess the acceptability of gestures by the locations where they might be used and the audiences they might be used in front of. Second was an on-street study to see how people responded to doing gestures in real settings. This study attempted to determine what factors influenced gesture acceptance and served as a comparison to the survey method.

THE GESTURE SURVEY

As an initial step towards understanding the social acceptability of gestures, a Web survey was conducted to examine how location and audience affect user willingness to perform gestures. Both of these aspects of social situations affect how we behave and make decisions about appropriate actions [10] and therefore provide a useful base from which to begin exploring social acceptability.

Device Based Gestures	
Whip	Whip-like movement of the device in the hand.
Device Shake	Shaking the device in the hand, side to side.
Pocket Tap	Tapping the device with the hand in a pocket.
Device Squeeze	Squeezing the device with the hand.
Table Tap	Tapping the device with the hand setting on a table.
Body Tap	Tapping the shoulder with the device.
Rhythm	Shaking the device in the hand up and down on an even rhythm.





Body Based Gestures	
Shoulder	Rotation of the shoulder, forwards and backwards.
Hand Proximity	Measuring the space between the hands.
Nose Tap	Single tap on the nose with the hand.
Head	Nodding of the head, side to side.
Balance	Shifting the balance while sitting.
Wrist	Rotation of the wrist, side to side.
Clap	Clapping of the hands.
Foot Tap	Tapping the foot, sitting or standing.
Belt Tap	Tapping of the belt with the hand.
Finger Proximity	Measuring the space between two fingers on one hand.
Arm Squeeze	Squeezing of the forearm with the opposite hand.

Figure 1. Names and descriptions of all the gestures used in the survey.

This survey examined a set of eighteen gestures that included both device-based and body-based gestures. A body-based gesture refers to any gesture that directly involves movements of the body without the use of a typical mobile device. For example, head nodding and foot tapping are both gestures that are performed without directly touching a typical mobile device. A device-based gesture refers to any gesture that directly involves touching or moving a typical mobile device, in this case a mobile phone, for example, device shaking and squeezing. The gestures were chosen based on usage in existing gesture-based interfaces [2, 6, 16, 19] and their potential to be used in future interfaces. The names and descriptions are given in Figure 1.

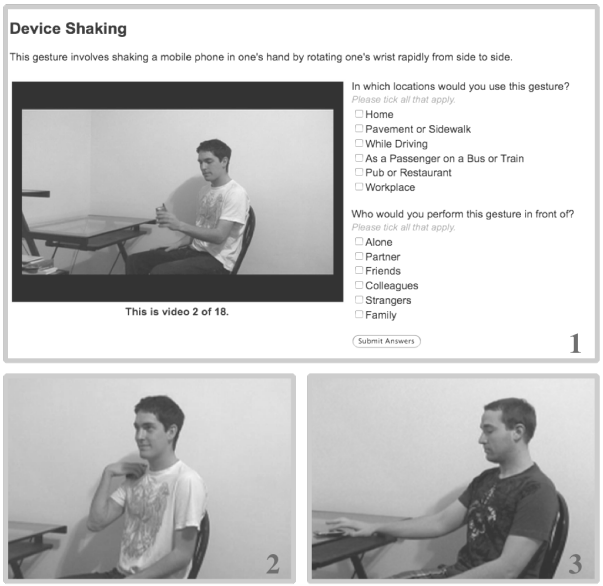


Figure 2. Screenshots and video frames from the gesture survey. 1. Survey Questionnaire Layout. 2. Frame from Body Tapping video. 3. Frame from Table Tapping video.

Before beginning the survey, respondents were told the aim of the study was to assess the social acceptability of these gestures for use as part of a mobile interface. For each gesture, survey respondents watched a video of the gesture being performed and answered multiple-choice questions. Each video was displayed with the gesture name and a short description of the gesture portrayed. The videos lasted from one to four seconds and were played on a loop while survey respondents answered the associated questions. These videos did not include any audio. Each video portrayed a single gesture being performed by a male actor sitting at a desk in front of a plain background. Figure 2 shows screenshots of the survey questions as well as frames from two of the gesture videos. Because survey respondents were asked to imagine the locations and audiences where they might perform these gestures, the videos were designed to focus solely on the gesture itself. The videos used in this survey intentionally portrayed a plain scene without a defined context so that the setting would not distract viewers from evaluating the gesture. Since an important role of the videos was to make gestures easy to observe, they were performed using clearly visible movements. For example, wrist rotation was performed with the arm held in front of the performer rather than at the side of the body. While these constraints did not always reflect

real-world usage, this was necessary to give respondents a clear view of the gesture being performed.

After watching each video, respondents were asked to select from a list all of the locations where they would be willing to perform the given gesture as part of a mobile interface. Users were then asked to select from a list all of the types of audiences they would be willing to perform the gesture in front of (locations and audiences are given in Figure 3). These options were selected to address a broad cross-section of common situations without providing an overwhelming number of options for survey respondents. These responses intentionally left out the 3rd person point of view, asking participants to imagine themselves in these settings in order to focus on the decisions made about personal actions rather than one's opinions of others' actions.

Locations	Audiences
Home	Alone
Pavement or Sidewalk	Partner
Driving	Friends
Passenger on a Bus or Train	Colleagues
Pub or Restaurant	Strangers
Workplace	Family

Figure 3. Locations and Audiences used in survey questions.

Results

From the survey responses, a level of "acceptability" was determined. The acceptability was calculated using the ratio of positive responses to negative responses for a given category. A positive response was a "yes" answer for willingness to perform the gesture with the given location or audience. Significance levels were determined using the Cochran [5] and the Chi-square [11] tests. The Cochran Test was chosen for its applicability to non-parametric data with multiple treatments and related samples. The Chi-square test was used for calculating the differences between frequency results. Significance levels were adjusted using Bonferroni's correction when multiple tests were taken.

Location	p value	Audience	p value
Home	.276	Alone	.152
Pavement	**<.0001**	Partner	**<.0001**
Driving	**<.0001**	Friends	**<.0001**
Passenger	**<.0001**	Colleagues	**<.0001**
Pub	**<.0001**	Strangers	**<.0001**
Workplace	**<.0001**	Family	**<.0001**

Figure 4. p values calculated using the Cochran Test for significance of gesture types for locations and audiences. $p < 0.003$ are significant, adjusting for Bonferroni's correction.

There were 55 survey respondents ranging from age 22 to 55. With respect to current geographic location, 43% of respondents were living in the United Kingdom, 45% in the United States, and 12% declined to state or were the only respondent in a given country. 29% of the respondents were female and 71% were male. Respondents were recruited through university email lists over the span of six weeks. Location and audience played a significant role in determining the social acceptability of gesture usage. There were significant differences in acceptability rates for all locations except home and for all audiences except alone, as shown in Figure 4. These two factors are important for individuals when determining acceptable actions.

Where? Locations and Gesture Usage

The awareness of one's surroundings and the ability to take advantage of the physical setting are important aspects of a performance. By comparing average acceptability rates between locations, we can determine which settings were the best facilitators of gesture usage and which settings were the most controversial. Figure 5 shows the average acceptance rate and standard deviation for each location.

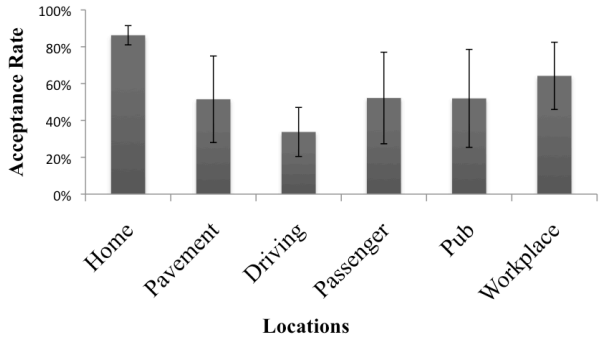


Figure 5. Average percentage of gesture acceptability by location. Error bars show one standard deviation.

	Home	Pavement	Driving	Passenger	Pub	Workplace
Home		**<.0001**	**<.0001**	**<.0001**	**<.0001**	**<.0001**
Pavement			**<.0001**	.629	.862	**<.0001**
Driving				**<.0001**	**<.0001**	**<.0001**
Passenger					.758	**<.0001**
Pub						**<.0001**
Workplace						

Figure 6. p values for pair-wise comparisons of locations. $p < 0.003$ are significant, adjusted for Bonferroni's correction.

Figure 6 gives the details of pair-wise comparisons for statistical significance of locations. The majority of locations were significantly different from every other location, with the exceptions of passenger and pavement, pub and pavement, and passenger and pub. Home, workplace, and driving were each significantly different from every other location. These differences show that location plays an important role in how users determine acceptability of a given gesture. For example, the significant difference between pavement and workplace shows that these places are socially different for survey respondents. As compared to the pavement setting, gestures were 12% more likely to be used in the workplace than on the pavement and 34% more likely to be used at home. The pavement, as a public location, provides few opportunities for privacy and exhibits a restricted set of social norms due to compromises of using a shared space with strangers. Home, on the other hand, offers many opportunities for privacy and clearly defined social norms as a familiar place shared with family members, with the workplace being somewhere in between the two. In each location, individuals use the social norms and stan-

dards as well as privacy affordances to determine if they will use a given gesture there.

An important aspect of the data is the differing levels of variance between the gestures' acceptability for each location. Higher levels of variance indicate that survey respondents had more varied opinions about which gestures were acceptable and which were not. Comparing home, with a standard deviation of 5%, to pub, with a standard deviation of 27%, shows a major difference in the way these locations were perceived. While home was relatively uniform in its perception, pub was very divergent. Acceptability rates varied from as low as 14% for shoulder rotation to as high as 95% for table tapping. These differences show how important gesture design is in order to create systems that are usable in a variety of locations.

Who's There? Audience and Gesture Acceptability

Each audience type was significantly different from every other type except friends and family as shown in Figure 7. The average acceptance rates and standard deviations for each audience are shown in Figure 8.

	Alone	Partner	Friends	Colleagues	Strangers	Family
Alone		**<.0001**	**<.0001**	**<.0001**	**<.0001**	**<.0001**
Partner			**<.0001**	**<.0001**	**<.0001**	**<.0001**
Friends				**<.0001**	**<.0001**	.008
Colleagues					**<.0001**	**<.0001**
Strangers						**<.0001**
Family						

Figure 7. Significant difference using Cochran test in pair-wise comparison by audience. p < 0.003 are significant, adjusting for Bonferroni's correction.

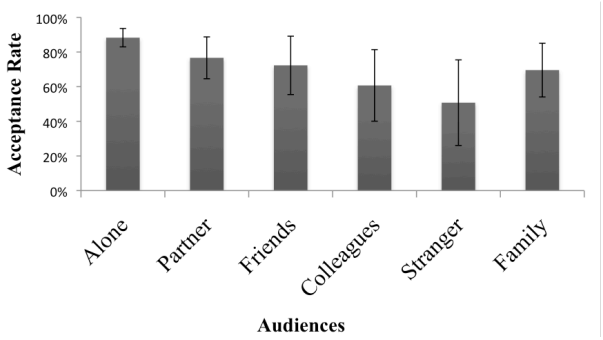


Figure 8. Average percentage of gesture acceptability by audience. Error bars show one standard deviation.

Familiarity with the audience played a significant role in gesture acceptability. Strangers, the least familiar audience with an average acceptability of 51%, were significantly lower than partners, with an average acceptability of 77% and alone with an average acceptability of 88%. This shows that more familiar audiences gave performers more confidence in using gestures than less familiar audiences. For example, 60% of respondents were willing to use the shoulder rotation in front of their partner, 54% were willing to use it in front of their friends, and only 20% were willing to use this in front of strangers. Familiar audiences provide performers with more freedom to try new things and more opportunities to explain their behavior. Friends or family members may have previous experiences witnessing an individual's usage of a particular gesture while a stranger may only be present during one isolated event.

Device and Body-Based Gestures

The survey data demonstrates a significant difference between the social acceptability of device-based gestures versus body-based ones. Overall, the p value for device based versus body-based gestures was < 0.001, determined using the Chi-square Test. Pair-wise comparisons, shown in Figure 9, demonstrate the device-based gestures were significantly more acceptable at every location except home and driving and every audience type except alone. The average acceptance rates and standard deviations of the device and body-based gestures are shown in Figure 10.

Location	p value	Audience	p value
Home	.038	Alone	.021
Pavement	**<.0005**	Partner	**<.0005**
Driving	.086	Friends	**<.0005**
Passenger	**<.0005**	Colleagues	**<.0005**
Pub	**<.0005**	Strangers	**<.0005**
Workplace	**<.0005**	Family	**<.0005**

Figure 9. p values calculated using Chi-square test for device versus body based gestures. p < 0.004 are significant (adjusting for Bonferroni's correction)

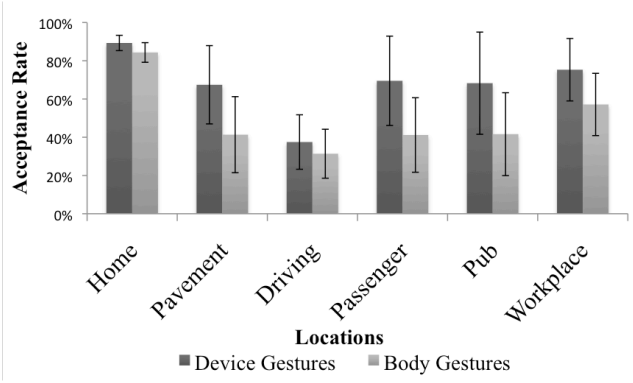


Figure 10. Average acceptability of device and body-based gestures by location. Error bars show one standard deviation of acceptability for each gesture type at the given location.

While gesture type had no significant difference for the home or driving settings, gesture type had significant differences for every other location. Device-based gestures were significantly more acceptable than body-based ones. Reeves *et al.* [15] describe the importance of the visibility of manipulations and effects in how spectators perceive interactions. With respect to gesture-based interfaces, the visibility of the gesture with a device plays an important role in gesture acceptance. The presence of a device in performers' hands changes their appearance and how they feel others will perceive them. Gesture-based interface users are more willing to use a gesture if it provides visual cues that explain their behavior. Device-based gestures provide clear indicators for audience members that a performer's actions are directed towards the mobile interface. The performer's

appearance gives them greater confidence that observers will perceive gestures as part of this interaction. There were some body-based gestures with high levels of acceptability. Foot tapping, for example, had an acceptance rate of 88% based on audience. This gesture utilized subtle, everyday movements. The ability to disguise gestures as everyday activities appears to make them more acceptable.

Discussion

The survey demonstrated that individuals take into account audience and location when deciding to interact with a gesture-based interface. These factors, which have already been used to describe behavior in public places [10], also influence gesture usage. Locations that provided users with more privacy were more likely to have higher acceptability rates than those that did not. With respect to audience, users were more likely to perform gestures in front of familiar audiences. Gesture type also influenced user willingness to perform it. Device-based gestures, which provided visual cues to audience members about the performers' actions, were more likely to be used than body-based gestures. The ability to demonstrate that an action is part of a mobile device plays in important role in gesture acceptability.

This survey demonstrated the fact that the social acceptability of gesturing in certain locations is sometimes heavily coupled to the expected audiences. People expect the pavement to be full of strangers and expect partner or family members in the home. Because of this, we will describe settings using both the location and audience. For example, a private setting, such as the home, is typically comprised of familiar audiences, such as a partner or family members. A public setting, such as the pavement, is comprised mainly of strangers and represents a more restricted social setting. The semi-public setting is somewhere in between, with a restricted but not necessarily familiar audience. For example, an office workplace is not usually open to the public, but each colleague is not necessarily familiar and social norms and standards influence behaviors. Unusual combinations also exist, for example when one is hosting a new acquaintance or love interest in his or her home. Although the home is usually considered private, an unfamiliar audience changes the dynamics of this situation.

While these results provide an interesting insight into the social acceptability of gestures, they are based on imagined situations rather than real ones. The following study further examines these results using real world settings. Rather than focusing on specific locations and audiences, this study uses private and public settings to evaluate gesture usage. A select set of gestures from the survey was examined in the user study in order to compare the results and evaluate the effectiveness of the survey method.

GESTURES IN THE WILD

Eight gestures from the previous gesture survey were investigated in this study. They were chosen to include four body-based gestures and four device-based gestures. For both types, two acceptable and two unacceptable gestures were chosen. Acceptability was determined using the acceptance rate for each gesture from the gesture survey. Unacceptable gestures were those with comparatively low acceptance rates as compared to other gestures of the same type. These are shown in Figure 11.

Body Based Gestures		Device Based Gestures	
Wrist Rotation	Acceptable	Shaking	Acceptable
Foot Tapping	Acceptable	Screen Tapping	Acceptable
Head Nodding	Unacceptable	Rhythm	Unacceptable
Nose Tapping	Unacceptable	Shoulder Tap	Unacceptable

Figure 11. Gestures used in the second user study.

Study participants attended three gesture sessions on an individual basis. These sessions were spaced one week apart and lasted roughly one hour. These repeated sessions were completed in order to better understand how gesture usage over time affected participants' opinions with respect to gesture preferences, emotional comfort and interest. During each session, participants were asked to perform repetitions of each gesture in both an indoor and an outdoor setting. These settings, private and public, were chosen to reflect two highly divergent locations as seen in the gesture survey. The indoor setting consisted of a private room in the University with a desk and chairs. For the outdoor setting, participants stood in a public location on the pavement near a bus stop and an underground station on a busy city street as shown in Figure 12. After completing the experiment in each setting, the participant was given a semi-structured interview in the indoor setting. Gesture order and setting order were randomized for each session.

Figure 12. Outdoor setting for user study. Experimenter on left, participant on right on a busy city street.

During each session, participants used an application that prompted gesture usage and logged accelerometer data. The accelerometer data was used to determine the amount of energy, in the form of acceleration, used during each gesture repetition. This information was collected to better

understand how subjective measures of energy compared to user preferences. The application used for this study ran on a Nokia N95, a standard mobile phone. It utilized the internal accelerometer to log device-based gesture movements and an external SHAKE sensor pack to log accelerometer data for the body-based gestures. The external sensor pack, roughly the size of a matchbox, contains an accelerometer, magnetometer and gyroscope that can all be accessed through a Bluetooth connection (Figure 13, left). The SHAKE was attached to the wrist for the wrist rotation and nose tapping gestures (Figure 13, right), attached to the shoe for the foot tapping gesture, and attached to headphones worn during the head nodding gesture.

Figure 13. Left: SHAKE device shown with a fifty pence piece. Right: SHAKE worn on the wrist.

For each gesture, the phone application played a short video showing the current gesture to be performed. Once the video was complete, the name and description of the current gesture appeared on screen. Then the application would scroll white text on a black background across the screen. The scrolling text was used to provide users with a short reading task, roughly 10-15 seconds, between gesture repetitions. Participants were told this was meant to appear as though they were reading text messages or checking email. The application would prompt participants to perform the next instance of the current gesture by flashing a white screen with the gesture name. After completing five repetitions of the gesture, participants would continue on to the next gesture until all 8 were completed. During the study, participants were told to perform the gesture based on their interpretation of the video and textual descriptions. The experimenter offered no advice on "correct" performance, always advising the participants to view the video or read the description if they had questions about a gesture.

Results

Throughout the course of the study, data were collected in the form of accelerometer log files and semi-structured interview responses. Eleven participants completed this study for a total of 33 participant hours. Participants were aged from 21 to 28, comprised of 4 females and 7 males.

Why Gesture? Preferences and Rationales

The gestures in the wild study presented a variety of ways that gestures could be described as acceptable or unacceptable as provided by the users themselves. By asking users to provide the reasons for liking or disliking gesture, we can begin to develop a set of guidelines to evaluate the acceptability of new gestures for use in mobile interfaces. After each gesture session, participants discussed which gestures they liked and disliked, listed from most to least frequent.

- **Subtle Movement** – Gestures that were small or unobtrusive such as tapping or foot tapping, were more comfortable for survey participants. For example, one participant stated that "considering what we were doing, the most important part was that it wasn't drawing unnecessary attention to what we were doing, some of the other gestures were."
- **Similar to Existing Technology** – Some of the gestures, such as tapping and rhythm, were described as similar to touchscreen interfaces and tilt-based interfaces, and this made them more acceptable to use. For example, one participant stated that "tapping the phone is quite comfortable because we use touch phones anyway, it's more understandable." Even without experience using touch screen technology, participants still reported this as a reason for finding certain gestures more comfortable. One participant stated that "I haven't used them [iPhone applications], I have seen people using them. I suppose for both the iPhone and things like the [Nintendo] Wii, they are not necessarily used on the phone, they're kind of, part of public consciousness of ways you can interact with something."
- **Looks or Feels Similar to Everyday Actions** –Those gestures that were familiar in feeling or appearance, for example tapping or shaking, were more accepted by study participants. For example, one participant stated that "foot tapping looks very similar to what you do normally anyway. You'd probably be tapping your foot if you were listening to music or something." Another example was a gesture that felt like an everyday action, such as the shaking or rhythm gestures, which were described as "natural movements, they were things you'd do anyway, like shaking a juice bottle to mix it up."
- **Enjoyable Movement** – Gestures that provided a high amount of satisfaction, whether from the actual feeling or appearance of the gesture itself, were more acceptable and even desirable to perform. One participant described their enjoyment of the rhythm gesture as "I don't know, I just fell for it." This participant went on later to describe gesture interfaces as "it's quite funky, and it's quite cool, to dance or something, move your hand up and down… The business men and well off people, I don't know if they would be able to use all of those gestures, but cool people, they would, and they would be happy to use it." This result shows that an element of enjoyment or playfulness in a gesture can make it more acceptable.

Participants also described gestures that they disliked and provided reasons for their opinions. The following reasons

were reported for feeling uncomfortable about gestures, listed in order from most frequent first to least frequent.

- **Looks Weird or Attention Seeking** – Gestures that required the participant to perform large or noticeable actions were the most commonly disliked gestures. For example, one participant stated that "the shoulder tap, nose tap, and head nodding were all quite attention seeking. I was actually aware that someone looked at me when I was tapping my nose, I thought that must seem a bit strange."
- **Physically Uncomfortable** – Some gestures, including head nodding, foot tapping, and wrist rotation, required the participant to move their body in a way that was described as uncomfortable. For example, one participant stated that "the shoulder tap I did not like that one at all. That was uncomfortable. It was like the foot tapping as well, it was uncomfortable."
- **Interferes with Communication** – Participants disliked the head nodding gesture because it would be distracting or confusing to use during a conversation. One participant stated that "I think the head nodding is particularly weird because the head is the primary means of communicating with other people so people look at your head first of all."
- **Uncommon Movement** – Gestures that do not naturally occur in daily life, such as the shoulder tapping, were disliked. For example, one participant stated that "it [shoulder tap] was just a very unusual gesture, I can't imagine a context in which the shoulder would make any sense at all."

There is a significant overlap between these categories, even with some gestures in both the positive and negative lists. For example, the need for enjoyable movements is in direct opposition to the need for subtle movements. With the enjoyable movements, there is an element of taking pleasure in the display of the gesture and a desire for others to notice. Similarly, the need for gestures to be familiar opposes the need for gestures not to be emblematic [12], a gesture with a previously associated meaning. For example, an emblematic gesture such as a thumb up might be familiar but could interfere with communication because this already has a meaning for observers.

Public and Private: How Audience Affects Gesture Usage

To assess how settings affected gesture acceptability, participants were questioned about their opinions of the indoor compared to the outdoor setting. While three participants reported that they felt indifferent, the remaining eight said the settings were different. Of these eight, seven specifically reported that the outdoor setting was less comfortable than the indoor one. Participants discussed their awareness of the audience around them in the outdoor setting, sometimes mentioning it as an overall feeling while outside, or mentioning specific instances were they became aware of others watching them. For example, one participant stated that "being inside is more comfortable, you don't have people wondering what you're doing." Another participant stated that "there was one time when somebody turned around and looked at me when I was tapping my nose."

Participants discussed how this awareness affected the way they chose to perform the gestures. One participant stated that "the nose tapping, when I was outside, I kind of disguised it as a scratch or pushing my glasses up." While some participants took advantage of the ability to disguise movements while outside, others felt they "would tend to them a bit more quickly, just to get it done. Not to look like an idiot too much." Some participants reported that the outside location was uncomfortable for specific gestures. For example, "it felt a bit more conspicuous outside, especially when you are doing the head nodding gesture."

Energy and Gesture Ranking

An objective measure of energy was calculated using the accelerometer data for each gesture. Each gesture was also given two rankings: one based on the survey results and another on the user study, ranging from 1 to 8. The survey ranking was determined using the average acceptance of the eight gestures based on audience. Rankings for the user study were determined by subtracting the number of negative mentions from the number of positive mentions for each gesture from interview transcripts. The energy measurements and rankings are shown in Figure 14. These results show there is no positive correlation between low energy gestures and highly accepted gestures. The subtlety of a gesture was not necessarily a matter of energy required to complete it. For example, screen tapping was the most popular gesture and had the lowest energy level, but foot tapping, the third most popular gesture, had the second highest energy. It is not the energy required to perform a gesture that makes it acceptable, but the perceived appearance of that gesture.

Gesture Name	Survey Ranking	Study Ranking	Energy Measured
Screen Tapping	**1**	**1**	24.3
Shaking	**2**	**2**	93.5
Foot Tapping	**3**	5	196.0
Wrist Rotation	4	**3**	201.2
Rhythm	5	4	111.2
Shoulder Tap	**6**	**7**	67.5
Nose Tapping	**7**	**6**	74.0
Head Nodding	**8**	**8**	87.5

Figure 14. Gesture Rankings from both the survey and the study and energy measures.

In order to compare the results of the survey with the user study, we will examine the extremes within the rankings. By comparing the top three and bottom three ranked gestures from each study, we can see these are nearly identical, with the exception of rank 3, which was held by foot tapping in the survey and wrist rotation in the user study.

These similarities show that the survey was effective in predicting highly acceptable and unacceptable gestures. Even though these were imagined situations, respondents were able to provide accurate responses when shown videos and questions, as verified by the reported preferences when users performed gestures in real settings.

Changes Over Time: Developing Gesture Preference

Multiple trials were completed in order to better understand how user opinions of the gestures changed over time and which experiences influenced those opinions. After each session, participants were asked to compare their experiences to previous sessions. After the second session, 9 of 11 participants reported positive benefits of repeating the gestures for a second time. 8 participants reported greater levels of comfort, and 1 reported greater confidence during the second session. Interestingly, participants reported these greater levels of comfort when objective measures of energy indicated that gestures were performed in a similar manner across repeated trials. One participant reported that "I was more comfortable, now I'm used to doing gestures. I know what to expect." Participants felt more comfortable during the second session because they remembered positive experiences from the previous session. These experiences helped participants feel more confident that they would not become embarrassed during the subsequent trials. For example, one participant stated that they had "a greater familiarity with them [gestures], more comfortable doing them as well. It's not so new and unknown from last time. I know what it's like to touch my nose in public surroundings." When asked why the second session was different, another participant stated that "I think having the experience of actually performing the gestures. I knew it wouldn't cause me any undue concern, it was just a more comfortable experience overall." Participants performed an action, gathered feedback from the responses of others, and therefore had more information to work with in subsequent performances. After just one experience, participants were more comfortable using the gestures on a busy city street.

Gesture preferences also changed in the second session compared to the first session. One participant stated that they had a "clearer distinction between which [gestures] I liked best and which I disliked." Other participants stated that specific gestures became more acceptable. For example, one participant stated that "some of them, like the shoulder tapping or nose tapping, felt less awkward this time than they did last time. Maybe if I kept using them they wouldn't feel as strange." The continued exposure to these gestures allowed participants to develop clear preferences and provide high quality feedback. After the third session, 10 of 11 participants reported positive benefits of repeating the gestures, with all of these participants specifically reporting greater comfort in performing the gestures.

Discussion

Based on the reported reasons for liking or disliking gestures, we suggest that gestures for use in a mobile interface should mimic the style of gestures encountered in everyday life. While this may be metaphorical, for example device shaking as shaking a bottle, literal, for example foot tapping, or even imitating existing technology, for example tapping a device even though it is not a touch screen, gestures that are familiar will be more socially acceptable than those that are not. However, emblematic gestures with previously associated meanings, such as head nodding, should be avoided as they conflict with and confuse other interactions. It is the appearance of a gesture, not the energy used in its performance, that influences users towards accepting gestures. High-energy gestures can be highly acceptable, while low energy gestures can be distinctly unacceptable.

A survey that uses video prototypes and asks users to imagine usage contexts, with respect to audience and location, is an effective tool to identify acceptable or unacceptable gestures. The survey can rule out unacceptable gestures before time is spent during implementation because it utilizes video prototypes that can be produced without any gesture recognition algorithms needed. However, this survey is best used as a preliminary evaluation because it can only provide limited ranking information. A possible improvement to this study would be to include free form response space, allowing users to explain their answers. This kind of survey could be used to eliminate unacceptable gestures before implementation, saving valuable time and energy for developers and resulting in more acceptable interfaces.

Based on the results of the user study, we suggest two design considerations when completing user studies for social acceptability. Because we observed differences between the first and second trials, we would recommend completing at least two trials when evaluating for social acceptability because users will develop preferences and change their acceptance rates after multiple trials. Additionally, we suggest that such trials be completed in real world settings, rather than lab ones. Study participants based the vast majority of their opinions and feelings about the gestures based on the outdoor setting and the audience there. Gathering the experiences of performing gestures in real settings allowed participants to report more in-depth opinions of the gestures, using real social experiences to explain and understand those opinions. These suggestions were important and influential factors that greatly improved the quality of our user study and could be applied to future studies.

While these recommendations aim to improve gesture design, they do not address the technical issues of implementing the gestures. The desirable gestures described in the results, for example subtle actions, present challenges for implementers of gesture recognition code, especially with respect to the segmentation issue of knowing when intentional gestures towards the interface begin and end [16]. Our work aims to drive the development of gesture recognition from an early stage in order to avoid spending effort on recognizing gestures that simply will not be used. Because of this, we have intentionally looked past the immediate technical restrictions in order to focus on gesture design.

FUTURE WORK

While the studies described in this paper answered some of our questions about social acceptability, they also demonstrated the need to explore other factors of social acceptability. The ability to demonstrate actions as part of an interface could be incorporated into the design of gesture-based interfaces by exploring the visibility of gestures beyond simple visibility of the device. The broader implications of using these kinds of gestures in daily life should also be investigated. For example, the ways in which gesture associations and meanings are developed over time and how this affects existing meanings should be investigated.

CONCLUSION

Social acceptability is an important part of designing usable multimodal interfaces. This research has shown that there are a wide variety of reasons that affect the social acceptability of different gestures. The online survey, which utilized video prototypes, was validated by an on-the-street user study that produced highly correlated gesture preference rankings. The on-the-street user study demonstrated that user acceptance of gestures is increased after even one positive experience. The survey demonstrated the important role that observers play in social acceptability, with highly acceptable gestures including subtle imitations of everyday gestures and gestures with highly visible cues demonstrating their role as an interaction with an interface. These results provide researchers with concrete tools that can be used to assess the social acceptability of multimodal interaction techniques at an early stage of development.

ACKNOWLEDGEMENTS

This work was funded by a National Science Foundation Graduate Research Fellowship and the EPSRC funded GAIME Project (EP/F023405). Equipment and additional funding was provided by Nokia.

REFERENCES

1. Bolt, R. A. 1980. “Put-that-there”: Voice and gesture at the graphics interface. In *Proc. SIGGRAPH 1980,* ACM Press (1980), 262-270.
2. Brown, L. M., Williamson, J. Shake-to-talk: Multimodal Messaging for Interpersonal Communication. Proceedings of the 2nd International Workshop on Haptic and Audio Interaction Design (Seoul, Korea), 2007.
3. Campbell, S. Perceptions of Mobile Phone Use in Public Settings: A Cross-Cultural Comparison. International Journal of Communication 1 (2007), 738-757.
4. Cassell, J. (1998). A framework for gesture generation and interpretation. In R. Cipolla & A. Pentland (Eds.), *Computer vision in human-machine interaction* (pp. 191-215). Cambridge, UK: Cambridge University Press.
5. Conover, W. J. Practical Non-parametric Statistics. John Wiley & Sons, Inc., New York pp 199-208 (1980)
6. Crossan, A., Williamson, J., Brewster, S., and Murray-Smith, R. Wrist rotation for interaction in mobile contexts. In *Proc. MobileHCI 2008,* ACM Press (2008), 435-438.
7. Davis, F. D. A Technology Acceptance Model for Empirically Testing New End-User Information Systems: Theory and Results. In MIT Sloan School of Management. Cambridge, MA (1986)
8. Fishbein, M., and Ajzen, I. Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research. Reading, MA. Addison-Wesley (1975)
9. Fiske, S., Taylor, S. Social Cognition: from brains to culture. McGraw-Hill, New York (2008)
10. Goffman, Erving. The Presentation of Self in Everyday Life. Penguin Books, London (1990)
11. Howitt, D., Cramer, D. An Introduction to Statistics in Psychology: A Complete Guide for Students. Pearson Education Limited, Essex (2000)
12. Kendon, A. GESTURE. Annual Review of Anthropology (1997)
13. Malhotra, Y. and Galletta, D. F. Extending the Technology Acceptance Model to Account for Social Influence: Theoretical Bases and Empirical Validation. In *Proc. of HICSS 1999,* IEEE Computer Society (1999).
14. Reeves, S., Benford, S., O'Malley, C., and Fraser, M. Designing the spectator experience. In *Proc. CHI 2005,* ACM Press (2005), 741-750.
15. Ronkainen, S., Häkkilä, J., Kaleva, S., Colley, A., and Linjama, J. Tap input as an embedded interaction method for mobile devices. In *Proc. TEI 2007*, ACM Press (2007), 263-270.
16. Strachan, S., Murray-Smith, R., and O'Modhrain, S. BodySpace: inferring body pose for natural control of a music player. *Ext. Abstracts CHI 2007,* ACM Press (2007), 2001-2006.
17. Väänänen, K., and Böhm, K. Gesture Driven Interaction as a Human Factor in Virtual Environments – An Approach to Neural Networks. Virtual Reality Systems. R.A. Earnshaw, M.A. Gigante, and H. Jones. London, Academic Press Limited (1993).
18. Wexelblat, A. Research Challenges in Gesture: Open Issues and Unsolved Problems. In *Proc. of the International Gesture Workshop* (September 17 - 19, 1997). I. Wachsmuth and M. Fröhlich, Eds. Lecture Notes In Computer Science, vol. 1371. Springer, London, 1-11.
19. Williamson, J., Murray-Smith, R., and Hughes, S. Shoogle: excitatory multimodal interaction on mobile devices. In *Proc. CHI ‘07,* ACM Press (2007), 121-124.
20. Mo, Z. and Neumann, U. Lexical Gesture Interface. In *Proc. ICVS 2006.* IEEE Computer Society, Washington, DC, 7.