Mo!Games: Evaluating Mobile Gestures in the Wild

Julie R. Williamson and
Stephen Brewster
School of Computing Science
University of Glasgow
Julie.Williamson@glasgow.ac.uk,
Stephen.Brewster@glasgow.ac.uk

Rama Vennelakanti
Hewlett-Packard Labs India
24 Salarpuria Arena, Hosur Main Road,
Bangalore 560030
rama.vennelakanti@hp.com

ABSTRACT

The user experience of performing gesture-based interactions in public spaces is highly dependent on context, where users must decide which gestures they will use and how they will perform them. In order to complete a realistic evaluation of how users make these decisions, the evaluation of such user experiences must be completed "in the wild." Furthermore, studies need to be completed within different cultural contexts in order to understand how users might adopt gesture differently in different cultures. This paper presents such a study using a mobile gesture-based game, where users in the UK and India interacted with this game over the span of 6 days. The results of this study demonstrate similarities between gesture use in these divergent cultural settings, illustrate factors that influence gesture acceptance such as perceived size of movement and perceived accuracy, and provide insights into the interaction design of mobile gestures when gestures are distributed across the body.

Categories and Subject Descriptors

H.5.2 [**User Interfaces**]: Evaluation/Methodology

Keywords

Gesture-Based Interaction; Mobile Interfaces; User Experience; Social Acceptability; In the Wild Evaluation.

1. INTRODUCTION

The evaluation of user experience is necessarily situated within a specific context, where the current location, cultural setting, existing social relationships, and users' own memories and perceptions can all have a significant effect on their current experiences with an interface. For those interfaces that are specifically designed for use in public spaces, these issues are magnified as users become *performers* and passersby become the *spectators*. When users must perform new and possibly embarrassing actions in public spaces, the evaluation of user experience becomes even more important to ensure that new interaction techniques are usable and acceptable. These kinds of interactions are difficult, or arguably impossible, to evaluate in traditional lab based settings since these experiences are so highly coupled to the real world contexts in which they develop.

However, evaluating such interactions "in the wild" presents a variety of difficulties. For example, capturing, monitoring, or observing interaction in the wild can be difficult, especially if these interactions occur naturally and sporadically throughout the day as opposed to interactions within clearly defined and focused sessions. Additionally, the experimenter must give up a certain level of control over the experience in order to evaluate an interaction within a real world setting, where external factors such as the environment, the divergent responses from spectators, or a user's own internal state will influence how the interaction unfolds. However, there are a variety of benefits to overcoming these challenges, where such studies not only collect usage over time and qualitative experiential data but also demonstrate the practical and applied usage of the interactions being evaluated. This transition from lab prototype or imagined usage scenarios to actual real world usage is a significant contribution and represents an important aspect of the design of interactive systems: the possibility of actual deployment and impact.

This paper presents an in the wild user study of a mobile gesture-based game to explore how users might perform a variety of gestures as part of a game within their daily lives, completed in two very different cultural settings. The application, called Mo!Games, allowed users to play games on a mobile device by gesturing with the head, with the wrists and with the device itself. These different gestures were chosen specifically from acceptable and unacceptable gestures to allow users to explore a variety of different input styles. For six days, participants were asked to play the game in their daily lives, spending two days in each gesture condition. Users were encouraged to experiment with the different gesture modes in as many different places as possible. The goal of this study was to give users the opportunity to explore each gesture condition within their own daily lives to better understand when and where users chose to perform gestures and how they experienced these interactions.

2. PREVIOUS WORK

2.1 Gesture-Based Interfaces

In this paper, we explore gestures performed in 3D space using different parts of the body because these are more likely to have acceptance issues due to their highly visible na-

ICMI'13, December 9–13, 2013, Sydney, Australia

http://dx.doi.org/10.1145/2522848.2522874 .

ture. For example, requiring users to gesture with their arms or head is much more visible than performing screen-based gestures on a small touch sensitive surface. The usability of these highly visible gestures in 3D space while on the move has been investigated for variety of gesturing styles with multimodal feedback. For example, one study demonstrated how device tilting combined with vibrotactile feedback could improve usability and experience [9]. Another study looked at how user perceptions about different modalities affected the way those modalities would be used [15]. Head [3] and wrist [4] tilting interactions have been demonstrated for use in mobile contexts. These studies investigated error rates for menu browsing tasks using tilt interaction for the head and the wrist while sitting and walking in a lab setting. Although error rates increased while walking, participants were still able to successfully interact using the head and wrist. Another study explored the usability of âĂIJbody spaceâĂİ style gestures [14] while on the move. This study demonstrated the successful use of âĂIJbody spaceâĂİ gestures for creating shortcuts on mobile phones while walking in a lab setting.

Although there have been many studies of the usability and accuracy of gesture-based input in lab studies, in the wild studies of gestures performed using different parts of the body in 3D space (as opposed to on-screen gestures) are rare. One example of an in the wild study of mobile gestures looked at how wrist gestures could be used to support navigation and selection while mobile [16]. Evaluating gestures in the wild presents additional challenges because of the difficulty in collecting reliable input and perceiving output. Gesture-based inputs are difficult to recognise effectively, requiring sophisticated signal processing and recognition techniques and often specialised equipment. If interaction requires users to wear or carry additional sensors or equipment this adds additional barriers to successful interaction in the wild. Users' abilities to perceive multimodal feedback can also be affected by environmental factors, such as noise, vibration, or temporary impairments, such as wearing headphones or shifting focus to navigation in difficult situations. The study presented in this paper addresses some of these issues by designing feedback in three modalities (visual, auditory, and vibrotactile) and supporting simplified gesture input using only a single external sensor. These design decisions and compromises helped to make an in the wild study feasible while maintaining a high-quality interactive experience.

2.2 Social Acceptability

Social acceptability is an important consideration in the design of gesture-based and multimodal mobile interfaces because these interactions often require users to adopt new and possibly embarrassing behaviours in public spaces. In real world settings, social acceptability is not a simple binary issue but a fluid decision process that takes into account users' previous experiences, the feedback they receive from spectators. Previous work in social acceptability has sought to better understand the factors that influence the decision process using a variety of methods. Rico and Brewster [11] completed an on-the-street study that examined how location and audience affected user decisions about social acceptability. Montero et al. [8] ran a focus group study that examined how the visibility of actions and their effects [10] of those actions influenced the perceived acceptability of gesture-based interactions. Ronkainen *et al.* [12] conducted a survey study that looked at scenarios of use and the acceptability or usefulness of gesture-based interaction.

This previous work explore a variety of factors that influence social acceptability, investigating how performance affects acceptance, evaluating the visual aspects of performance, and examining how age might affect evaluations of social acceptability. While these studies clearly demonstrate how these factors influence social acceptability, much of this work is based on imagined scenarios, short lab experiences or highly controlled "on-the-street" settings. The study presented in this paper aims to build on these concepts by bringing them out of the lab into real world situations in different cultural settings.

3. EVALUATION IN THE WILD

When looking at the acceptance and use of gesture-based interfaces, evaluations should investigate how users choose to perform actions, how they feel their actions are perceived by others, and when and where they choose to interact. Often, it is not possible or effective to evaluate these kinds of issues in a lab setting and user studies must be brought out into real world contexts. However, evaluations of such interactions "in the wild" present challenges, especially for mobile interfaces that will be used while on-the-go.

Traditional ethnographic methods of observation are typically not reasonable or practical if interaction occurs naturally and sporadically throughout the day. There are also technical challenges since such a deployment would be subject to a huge variety of contexts, noise levels, and other situational background disturbances that could affect sensor accuracy, usability and experience. For example, recognition accuracy rates for gesture-based interactions can often be affected by unexpected circumstances, such as sensor noise from walking or riding public transport. Additionally, multimodal feedback may be difficult to perceive in different settings where environmental noise impairs perceptions or distractions, such as safely navigating through busy places, limit one's ability to focus on interaction. These challenges make longitudinal or "in the wild" studies difficult for gesture-based or multimodal mobile interfaces. Experimenters must balance the tradeoffs between a consistent and controlled experience where more uniform data can be collected with one that is functional and practical in real world settings. This often means making compromises with respect to the kinds of interactions that can be supported and the kinds of data that can be collected.

The design of the user study presented in this paper approached these issues by simplifying gesture input, gathering data from a variety of sources, and creating a system that could be used flexibly in a variety of contexts. This study combined quantitative metrics (such as accuracy and location of use) and qualitative methods to gain the best possible view into the usage and experience of the application. The user study revolved around a gesture controlled mobile game application that allowed for three different modes of gesture control: gesturing with the wrist, the head, and the device itself. The application, called Mo!Games, was designed to incorporate this gesture-based input with audio, visual, and vibrotactile feedback.

The study was completed with participants from Glasgow, United Kingdom and Bangalore, India. These different cultural settings were chosen in order to compare a typical

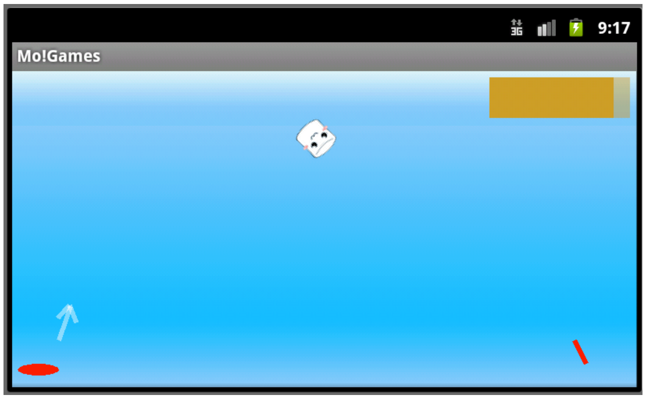

Figure 1: Screenshot of the beginning of a Marshmallow Toss game. Marshmallows are targeted using the white arrow on the left side of the screen, and tossed at that angle when the screen is tapped. The target on the right side of the screen is a red stick, which moves after every toss. The orange bar in the top right is a timer for the game.

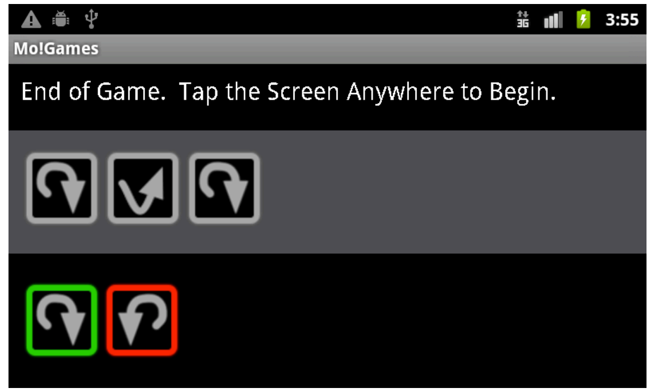

Figure 2: Screenshot from the end of a Simon Says game. The grey list of icons in the top row are presented during the first half of the game. These must be memorised and repeated during the second half once the list is no longer displayed. As the gestures are repeated, the corresponding icon is displayed in the bottom row with a red or green border to indicate correctness. At the end of the game, as shown in this screenshot, both lists are displayed together.

Western user group with an emerging market user group. These cultural groups offered an interesting comparison because both would be familiar with mobile technology but individuals from these cultures might use mobile technology in very different settings with different histories of mobile technology in everyday life. The results of this study explore participants' perceived usability of the interactions, their preferences and appropriations of the different gesture modes, the experience of gesturing as part of a mobile interface in everyday life, and the cultural differences between these groups.

3.1 The Mo!Games Application

The Mo!Games application was designed for use in everyday life in order to explore how users might perform gestures as part of a mobile interface in the wild. The application consisted of two games, one which utilised continuous gesture control and one with discrete gesture control. These two interaction styles support different performances which may influence acceptance and usability. Each game lasted roughly one minute, allowing participants to easily pick up the application for short playing sessions throughout the day in a casual gaming style. The application included an interface for leaving voice notes, where participants could leave feedback, comments or relate stories about their experiences during the course of the study. Additionally, the application included help screens with information about how to use each gesture mode, how to position the external sensor used in the study, and how to troubleshoot issues with the external sensor and the application itself. The application also included an 'achievements' page where participants could keep track of their daily usage and gain achievements based on how many games they had played, how well they had been doing, and other simple aspects of the games. This was included to encourage participants to play the games and be aware of how long they had played thus far.

The first game in the application was a trajectory/toss game, where users had to toss marshmallows onto a target, shown in Figure 1. In the Marshmallow Toss game, the angle of the toss is controlled using gestures with continuous visual feedback. The tilt or rotation corresponds to the angle of the launch, and a marshmallow is tossed after the user taps the touchscreen. Vibrotactile and audio feedback are provided when the marshmallow successfully hits the target, with a different vibration pattern when the target is missed. The goal of this game is to gain as many successful hits within one minute, with achievements for actions like most successful hits, best hit rate, and fastest time to successfully hit five marshmallows. The target moved randomly around the shooting area after each toss.

The second was a 'Simon Says' style game where each game began by presenting a list of icons corresponding to rotate left, rotate right, and flick downwards gestures that must be memorised within 30 seconds, shown in Figure 2. Once 30 seconds had passed, the list disappeared and the user had to perform the gesture list correctly from memory within 30 seconds. After three successful games, the list size increased by one. If users made a mistake, they moved back one level. The interface provided visual, audio and vibrotactile feedback when users were repeating the gesture list based on the correctness of their actions. The goal of this game was to successfully complete as many rounds as possible without making mistakes, with achievements such as passing different levels and best success rates.

Both of these games could be completed using one of three gesturing styles: gestures with the wrist, the head and the device. For the Marshmallow Toss, the angle of the trajectory was continuously controlled by the angle of the wrist, head, or the device itself. For the Simon Says game, rotate left, rotate right, or flick could be performed with the wrist, head, or device. Thus, each game can be completed using the same basic movements performed with different parts of the body. The application was deployed as an Android application on a Google Nexus One. The external sensor used was a SHAKE sensor, which is a Bluetooth enabled sensor pack that includes an accelerometer, gyroscope, and magnetometer. The external sensor pack was only used for the head and wrist conditions with the internal Nexus One sensors used for the device condition. Participants wore the sensor using a black elastic wrist band for the wrist condition and a hook/loop attachment with a hat or headband for the head condition.

3.2 Gesture Design

The Mo!Games application made use of gestures performed with the body in 3D space to explore how these highly visible gestures might be used and accepted in different cultural settings. The head, wrist, and device-based gestures were chosen based on previous work that demonstrates these as relatively unacceptable, moderately acceptable, and highly acceptable, respectively [11]. These gesture also provide flexibility in supporting both continuous and discrete control using the same actions performed across the body. For example, tilt left can by performed as a continuous or discrete tilt using either the head, wrist, or the device itself. By implementing gesture control in this way, this study also explores issues of interaction design when using gesture input with different parts of the body. For example, how does symmetry across the body influence usability? Do different body parts support continuous or discrete interaction more easily? By using the same gesture input in different styles with different body parts this study could explore these issues.

3.2.1 Recognition Design

Because this application was used while mobile in real world settings without an experimenter present, it was important to keep the gesture recognition simple and robust. The gesture recognition was designed using accelerometer-based sensing that calculated the roll and the current energy or total movement of the sensor. Using a thresholding technique combined with a simple state machine, rotation left and right were recognised by rotating a sufficient amount in the appropriate direction and returning to a central position. The flick gesture was recognised by calculating the current change in energy and testing against a threshold. Small adjustments to these thresholds were made for the head mode, so that slightly less movement was required for successful recognition.

3.3 The Study

Participants in the Mo!Games evaluation were asked to use the application for six days, completing three gesture conditions of two days each. Participants were not required to complete a certain amount of usage, but were encouraged to use the system in as many different locations as possible. As a guideline, participants were asked to use the system for roughly twenty minutes per day. The order of the gesture conditions was counterbalanced between the participants and was controlled and changed automatically by the game application. All activities within the application were logged, including any click or touch input, the path taken through the game screens, and all gesture interactions. Logging of the gesture interactions included successful and unsuccessful actions as well as the progress of the current game.

As part of the evaluation, we collected a wide variety of data beyond basic usage logs, including a variety of user-reported data about location and their experiences as well as qualitative interview data. Throughout the course of the study, participants were asked to tag their current location from a list of possible locations consisting of: home, work, pavement or sidewalk, on public transport, on private transport, in a shop, in a restaurant, or other with an optional text input field. These location categories were used in order to provide useful information about the participants' current context without compromising privacy or security. For this reason, the application did not log GPS or any other personal information that would identify participants' current physical location. Additionally, participants were asked to provide feedback about their experiences by recording voice notes throughout the study. At the end of the study, participants were interviewed about their experiences with the game and their preferences for and perceptions of different gesture modes.

Because of the inherent challenges with completing an in the wild investigation for multimodal interactions, the design of this study was based around another successful in the wild evaluation of multimodal feedback [7]. In this user study, participants were asked to answer trivia questions on a mobile phone for three feedback conditions lasting two days each, could leave voice notes as feedback throughout the evaluation, and could tag their current location while interacting with the application. Important aspects of that study were incorporated into the Mo!Games evaluation, such as giving participants specific equipment to use over the period of roughly a week while going through a set of multimodal conditions each lasting two days. The work completed by Hoggan and Brewster also demonstrated the success of asking users to tag their current location whenever using the system, and allowing them to provide feedback throughout the course of the study using voice notes recorded within the application.

20 users took part in this study, with ten each from the UK and India. The participants ranged in age from 20 to 45, 7 were female and 13 were male. 10 of the participants were students, and 10 worked in technical office jobs. The participants were recruited through email and social networks.

4. RESULTS

The results of this user study are based on detailed usage logs generated by participants, voice notes and other feedback provided by participants, and semi-structured interviews completed at the end of the study. This includes over 35.5 participant hours playing games, distributed throughout the participants' everyday lives. Participants each spent an average of 106 minutes each interacting with the application over the span of six days. On average, participants interacted with the application for 40 minutes in the device mode, 37 minutes in the wrist mode, and 34 minutes in the head mode. This section will begin by discussing the cross-cultural comparisons between these two user groups, then go on to discuss overall results for user perceptions, gesture mode preferences, and interaction design from both groups together.

4.1 Cultural Differences in Input Technique Preference

One of the interesting results of this study is in the *similarities* between these highly divergent cultural settings. Even though the meaning and use of gestures can vary widely between different cultures, participants from the UK and

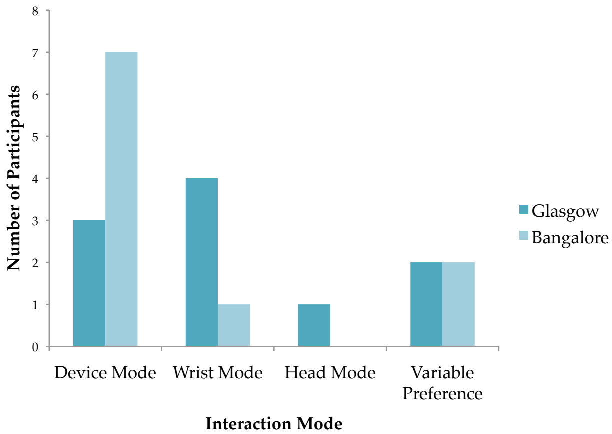

Figure 3: Number of participants that preferred different gesture modes, or described variable preferences based in the game or location, for participants in Glasgow and Bangalore.

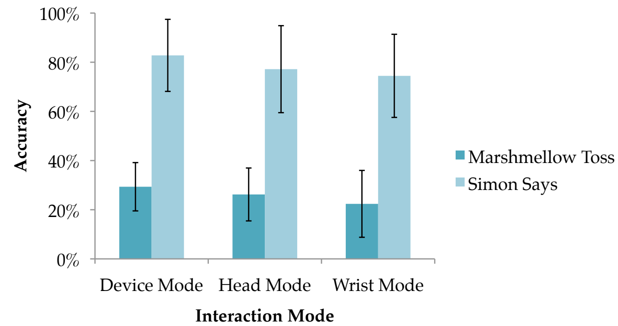

Figure 4: Successful hit rates for each gesture mode. Error bars show one standard deviation.

India did not have significantly different preferences for the different gestures modes. Figure 3 shows the number of participants that favoured each gesture mode, or had no clear preferences. Pair-wise Mann-Whitney tests for independent samples [1] indicate there were no significant differences between the numbers of users preferring each mode in the UK and India. For these participants, there was no clear gesture mode that was favoured in one culture and not the other. It is possible the sample size is not large enough to tease out significant differences between these groups.

However, device, head, and wrist mode were ranked in a similar order for participants from these two cultural groups, and the order of those ranks was significant. Friedman tests [2] for ordinal data showed a significant difference ($p = 0.0017$) between the ranks for different gesture modes for all participants. Pair-wise Wilcoxon Signed-Rank [1] tests showed significant differences between the ranks of gesture modes for all participants. Wrist mode was ranked significantly higher than head mode ($p = 0.05$, medium effect size $r = 0.4$). Device mode was rank significantly higher than wrist mode ($p = 0.04$, medium effect size $r = 0.4$). Device mode was also ranked significantly higher than head mode ($p = 0.0004$, large effect size $r = 0.8$). This demonstrates that although there were no significant differences in gesture mode preference between different locations, previous results indicating that device-based gestures are more socially acceptable than body-based gestures [11] was true for participants in both locations. This similarity between these cultural settings further demonstrates the importance for users to demonstrate the purpose of their interactions in public, which is more easily achieved when a mobile device is visibly involved.

4.2 Perceptions about Accuracy and Usability

Looking at both groups together, the results show some issues with perceptions of accuracy and usability. During the post-study interview, participants described how the accuracy of the gesture modes varied. Even though the successful hit rates were not significantly different, having similar means and standard deviations as shown in Figure 4, participants had strong opinions about which modes had the best control and accuracy. The majority of participants felt that the device mode was the most accurate, closely followed by the wrist mode. In both cultural settings, head mode was often described as the least accurate and usable mode. Given that the recognition techniques were nearly identical for each gesture mode, these different perceptions of accuracy and usability between gesture modes are based on a variety of experiential and physical factors. For example, the flexibility and dexterity of the head and wrist are very different. However, participants discussed a variety of issues that went beyond simple physical constraints.

4.2.1 Perceived Size of Movement

Participants described how some of the gesture modes did not seem to respond as quickly as others, where the perceived size of the movement had a strong influence on how well the system seemed to respond. Even though the recognition techniques required roughly the same amount of displacement for each of the three gesture modes, participants often felt that some modes required more movement than others. For example, when describing the head gestures, one participant stated that “it felt like it was the least accurate, it felt like you had to make big movements, it didn’t feel good to use and I didn’t seem to pick it up as easily” (P1, India). Differences in the physical constraints of the different modes and how the gestures were performed made some modes appear to require more movement. For example, the head gestures often seemed to require more movement because head gestures used a larger proportion of neck flexibility than wrist gesture required of wrist flexibility. In other cases, participants performed extraneous movements while completing gestures and mistakenly thought that those interactions required larger movements. Previous work has investigated how users develop misconceptions about the actions required for successful interactions [6], where these results demonstrate how these misconceptions can have significant effects on the acceptability of interaction.

4.2.2 Isomorphism Errors

Certain gesture conditions seemed more prone to isomorphism errors [17], where there was a disconnect between a user’s perception of what a system is recognising and what a system is actually recognising. Particularly with the wrist condition, participants had difficulty identifying and repeating successful performances of the tilt left and tilt right gestures. For example, one participant stated that "initially I could not get the hang of this, though I could move it around easily the exact direction was difficult so I could not score" (P5, India). In this case, participants often confused speed or holding a certain rotation as being important aspects of performing the gesture.

4.2.3 Lack of Familiarity

Participants found actions that controlled the games in an unusual or unfamiliar way to be less accurate and more difficult to use. One participant stated that "I'm not used to using head or wrist to activate anything with the device. So with the device, I found it easier. I found device mode easy and comfortable" (P9, India). The strangeness and unfamiliarity of an interaction gave participants much less confidence when performing interactions and led to perceptions of decreased accuracy. In some cases, participants recognised that the "most accurate" mode was not their favourite. For example, one participant stated "my best score was in the wrist mode, but actually that's my least preferred mode" (P9, India). Another participant stated that "I felt the most control with the head mode, but I think it would be the one I would be least likely to use" (P15, UK). These situations highlight the need to explore how desirability and usability affect actual usage behaviours.

4.3 Gesturing in Everyday Life

Because this application was used in participants' everyday lives, they had the opportunity to experience performance in a variety of different settings. Figure 5 shows some of the most commonly tagged locations by participants throughout the study. Although the majority of games were played either at work or home, some were also played on the pavement, on public transport, and at restaurants. In these scenarios, participants provided detailed feedback about the effect of spectators and their feelings about how their performance appeared in public and semi-public locations.

Participants described how spectators could make them feel uncomfortable performing to the point that they would stop using the application altogether. One participant stated that "I was sitting on the steps in the mall and some people who were passing by stopped and looked at me. And then I didn't continue that game. I stopped it, and took off my cap and put it in my bag. So obviously I'm not comfortable playing it in head mode in public" (P5, India). Even the anticipated effect of unwanted attention from spectators was enough to prevent interaction in public places. One participant stated that "I made a very conscious decision that the only one I would ever play outside of work or home would be the device mode" (P11, UK). Another participant also described how only certain modes were acceptable in public locations: "I played the marshmallow game even in the restaurant, but I would not have felt comfortable playing it in the head mode or the wrist mode. Since I was in the device mode, it was more comfortable to be playing it in public places" (P7, India). For some participants, these initial barriers to interaction meant that they were not able or willing to experience the application in certain settings. Our evaluation specifically included unacceptable interaction techniques identified from earlier research in the UK [11], so rejection of the application in certain locations was

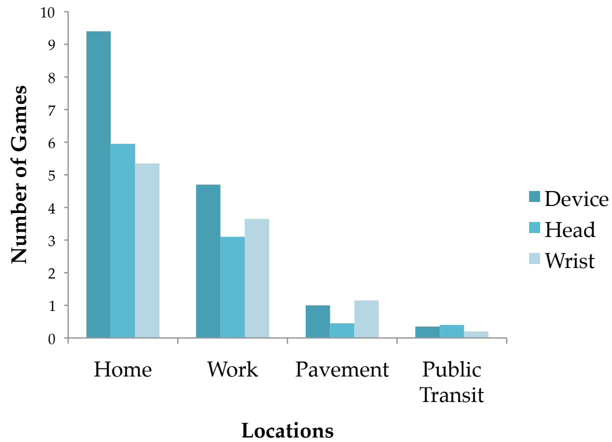

Figure 5: Average number of games played by locations for each gesture mode.

not surprising. However, these 'missed opportunities' for interaction would be detrimental to a large scale deployment of a real game if they were unexpected. This demonstrates the importance of completing early and low-cost evaluations of social acceptability to identify which interaction techniques are simply unacceptable for usage in everyday life.

Another factor that participants discussed was the greater social acceptance of interactions, with consideration of how this kind of acceptance might change their own perceptions a acceptability in the future. For example, one participant described how the ubiquity of smart phones meant that moving a device in a strange way is not as unacceptable as it used to be. This participant stated that "I was quite aware of moving my head around in an unnatural way, as where no one would say anything if you are doing that with a device" (P11, UK). Evaluating how greater social change affects new interaction technique acceptance remains an open issue in evaluations of social acceptability.

4.4 Gesture Mode Preferences

When looking at the overall gesture mode preferences for both user groups, the device mode was the most commonly favoured with ten of twenty participants ranking this as the best. However, those participants that preferred other gesture modes provided interesting insights into the reasons why body-based control might be more desirable. Five of the twenty participants favoured wrist mode, one favoured head mode, and four had variable preferences depending on the game or the context. During the post-evaluation interviews, participants described different aspects of the body-based interactions that made them more desirable than the device-based ones.

4.4.1 Better Control

Participants described how body-based gestures gave them a better sense of control and greater confidence while performing interactions. One participant stated that "I found that in wrist mode, it was responding much better. Also, when you are in the device mode, you have a fear that the device might fall while you are doing the gestures. That's not a problem with the wrist mode, once it [the sensor] is tight on your wrist you can move your wrist around as much as you want" (P12, UK). This sense of confidence and control allowed participants to explore the experience of gesturing more thoroughly, leading to a more enjoyable experience.

4.4.2 Hiding Interaction

In some cases, participants appropriated body-based actions by disguising or hiding their performance. Although the application was designed to support performative actions, the ability to perform more subtle interactions or hide interactions altogether made body-based interactions more comfortable in public places. One participant described how the body-based gesture modes could be performed with subtlety. This participant stated that "for the marshmallow game, I could just tilt my head slightly and tap, and no one could know what I was doing" (P1, India). Another participant described how using the device-based gestures to play the game was more obviously playful and thus inappropriate for a work place than the body-based modes, which could be performed surreptitiously. Although the device-based interactions made it clear that actions were directed towards an interface, the obvious use of an interface was not always desirable. In those situations where technology use might not be socially acceptable, the ability to hide interactions made these gestures favourable.

4.4.3 Novel Experience

Some users preferred body-based interactions simply because this was an enjoyable and new way to interact with a mobile phone. One participant stated that "I quite like the [wrist] gesture because it was novel, and I thought the device mode was less interesting" (P19, UK). Participants enjoyed experimenting with new interaction techniques simply for the sake of getting a new experience. Although this effect might wear off over time, other factors might encourage sustained use. For example, an interaction might be initially accepted for its novelty then continue to be used because it increased ease of use and allowance for better control. This may be a way for an initially less socially acceptable gesture to become accepted into normal life.

4.4.4 Continuous versus Discrete Interaction

Some participants described how their preference for a gesture mode was based on the game they were playing, with body-based interactions favoured for continuous interaction and device based interaction favoured for discrete interaction. One participant stated that "it felt odd doing the movements so jerkily, at least with the marshmallow game it was a smooth moving of the head as opposed to a twitch" (P11, UK). For discrete control, body-based interaction often had a much more rigid and abrupt appearance, making these interactions less acceptable and less comfortable than their continuous counterparts.

4.5 Discussion

The purpose of this study was explore how different gesture-based interaction styles would be accepted in the daily lives of participants from two different cultures. This occurred in the variety of ways that participants took a simple interaction and adapted it to make it work personally and socially for them. For example, participants discussed how head-based interaction could be easily hidden in everyday life by performing subtle actions and still appearing to use their mobile phone traditionally. Alternatively, participants described how they included friends and family in their game play in order to explain their actions to others and make their interactions more acceptable and even enjoyable in front of these spectators. Although the game was designed for individual use, many participants described how they used the game socially to "show off" novel technology as well as make their interaction a better experience. This balance between the ability to dynamically hide or perform interactions allowed for more possibilities for interaction in everyday life. When participants were alone in public, subtle and hidden actions worked best. However, the same interface could be used in a performative manner when with a group of friends. This kind of flexibility is an important aspect of acceptable gesture-based interaction, where a variety of performances should be supported.

There was also a variety of issues that made successful or acceptable interaction difficult in these real world settings. In some cases, participants were simply unwilling to experiment with different gesture modes in public places, whatever the cultural context. Head-based interaction was intentionally included in this study as an unacceptable interaction as demonstrated by previous works [11], and many participants were simply unwilling to perform these gestures in public. While this validates previous work (which used low-cost ways of gathering data about the acceptability of gesture-based input), it also demonstrates interesting situations in which users were willing to perform these actions. For example, participants described ways that even the least acceptable interactions could be incorporated into everyday life. By exploiting some of the benefits of unusual or traditionally unacceptable interactions, users of gesture-based input could be gradually encouraged to adopt these new interactions using the "foot in the door" technique [5]. This technique, named for the successful practice of door-to-door salesmen, describes how compliance or acceptance can be much more easily achieved through incremental steps rather than large changes.

One of the interesting aspects of this work was that these studies were completed in two very different cultural settings. Although one might expect to find significant differences between these cultural groups, it is interesting in and of itself that we found more *similarities* than differences in this study. Completing this cross-cultural study of user experience and social acceptability was motivated by a common criticism of research completed in specific contexts that cultural factors are not taken into account. While these results do not indicate that cultural factors do not have a significant influence on acceptance, it does indicate that these might not be as significant an effect as anecdotally indicated and these factors can be accounted for in design. For example, the gestures chosen in this study were specifically designed to be arbitrary in nature, with meanings that had to be learned in relationship to the device. In a system where gestures were designed based on existing cultural meanings, then cultural factors might have a much greater influence on usability and acceptance. In the case of Mo!Games, cultural differences simply did not have a significant effect on which modes users preferred or how they chose to integrate these gestures into their daily lives. The gestures chosen for this study were based on previous work demonstrating the relative acceptability/unacceptability of these actions. As expected, participants found some gestures, notably the head gestures, as unacceptable in public places and often were not willing to perform these gestures at all. This highlights the importance of completing early evaluations of acceptance in order to determine baseline preferences.

5. CONCLUSIONS

The results of this 'in the wild' user study demonstrate how users might actually use gesture-based interactions in their everyday lives. This study expands on previous work by bringing these interaction out of the lab into real world scenarios and evaluating usability and user experience in the wild. This represents a significant step forward from the previous lab-based studies and provides qualitative feedback about how these interactions might be woven into the fabric of everyday life. Although this study looked at a gaming application, these interaction techniques could also be beneficial in a variety of typical mobile phone tasks, such dealing with phone calls [13] or basic "eyes free" menu interaction [18]. Augmenting these common mobile phone tasks with gesture input represents another interesting opportunity to integrate these interactions into everyday life. The results of the cross-cultural comparisons in this study provide a useful insight into the similarities between people from different cultures with regards to how they integrate new technologies into their daily lives. The presence of these similarities can be more interesting than the differences because it represents an opportunity to engage an even wider audience into the evaluations of new technologies. While some significant differences in the social acceptability of new interactions may exist between users in different cultures, this study demonstrates that these differences can be accounted for in design. Cross-cultural evaluation of the user experience of performing gestures and multimodal interactions in public places represents a exciting challenge for the future of multimodal interaction research.

6. ACKNOWLEDGMENTS

This research was supported by a National Science Foundation Graduate Research Fellowship and generous support from HP Labs in Bangalore. We would also like to thank all of our participants for performing in so many different places with our system and sharing their experiences with us.

7. REFERENCES

[1] G. Corder and D. Foreman. *Nonparametric statistics for non-statisticians: a step-by-step approach.* Wiley, 2009.

[2] D. Cramer. *Introducing statistics for social research: step-by-step calculations and computer techniques using SPSS.* Routledge, 1994.

[3] A. Crossan, M. McGill, S. Brewster, and R. Murray-Smith. Head tilting for interaction in mobile contexts. In *Proc. of MobileHCI '09*, pages 6:1–6:10, New York, NY, USA, 2009. ACM.

[4] A. Crossan, J. Williamson, S. Brewster, and R. Murray-Smith. Wrist rotation for interaction in mobile contexts. In *Proc. of MobileHCI '08*, pages 435–438, New York, NY, USA, 2008. ACM.

[5] J. Freedman and S. Fraser. Compliance withouth pressure: The foot-in-the-door technique. *Journal of Personality and Social Psychology*, 4:195–202, 1996.

[6] D. Freeman, H. Benko, M. R. Morris, and D. Wigdor. Shadowguides: visualizations for in-situ learning of multi-touch and whole-hand gestures. In *Proc. of ITS '09*, pages 165–172, New York, NY, USA, 2009. ACM.

[7] E. Hoggan and S. A. Brewster. Crosstrainer: testing the use of multimodal interfaces in situ. In *Proc. of CHI '10*, pages 333–342, New York, NY, USA, 2010. ACM.

[8] C. S. Montero, J. Alexander, M. T. Marshall, and S. Subramanian. Would you do that?: understanding social acceptance of gestural interfaces. In *Proc. of MobileHCI '10*, pages 275–278, New York, NY, USA, 2010. ACM.

[9] I. Oakley and S. O'Modhrain. Tilt to scroll: Evaluating a motion based vibrotactile mobile interface. In *Proc. of WHC '05*, pages 40–49, Washington, DC, USA, 2005. IEEE Computer Society.

[10] S. Reeves, S. Benford, C. O'Malley, and M. Fraser. Designing the spectator experience. In *Proc. of CHI '05*, pages 741–750, New York, NY, USA, 2005. ACM.

[11] J. Rico and S. Brewster. Usable gestures for mobile interfaces: evaluating social acceptability. In *Proc. of CHI '10*, pages 887–896, New York, NY, USA, 2010. ACM.

[12] S. Ronkainen, J. Häkkilä, S. Kaleva, A. Colley, and J. Linjama. Tap input as an embedded interaction method for mobile devices. In *Proc. of TEI '07*, pages 263–270, New York, NY, USA, 2007. ACM.

[13] J. Ruiz, Y. Li, and E. Lank. User-defined motion gestures for mobile interaction. In *Proc. of CHI '11*, pages 197–206, New York, NY, USA, 2011. ACM.

[14] S. Strachan, R. Murray-Smith, and S. O'Modhrain. Bodyspace: inferring body pose for natural control of a music player. In *In CHI EA '07*, pages 2001–2006, New York, NY, USA, 2007. ACM.

[15] M. Turunen, A. Melto, J. Hella, T. Heimonen, J. Hakulinen, E. Mäkinen, T. Laivo, and H. Soronen. User expectations and user experience with different modalities in a mobile phone controlled home entertainment system. In *Proc. of MobileHCI '09*, pages 31:1–31:4, New York, NY, USA, 2009. ACM.

[16] J. R. Wiliamson, A. Crossan, and S. Brewster. Multimodal mobile interactions: usability studies in real world settings. In *Proc of ICMI 2011*, pages 361–368, New York, NY, USA, 2011. ACM.

[17] J. Williamson. *Continuous Uncertain Interaction.* PhD thesis, University of Glasgow, 2006.

[18] J. Williamson, S. Robinson, C. Stewart, R. Murray-Smith, M. Jones, and S. Brewster. Social gravity: a virtual elastic tether for casual, privacy-preserving pedestrian rendezvous. In *Proc. of CHI '10*, pages 1485–1494, New York, NY, USA, 2010. ACM.