Exploring Non-touchscreen Gestures for Smartwatches

Shaikh Shawon Arefin Shimon, Courtney Lutton, Zichun Xu, Sarah Morrison-Smith, Christina Boucher, Jaime Ruiz
Colorado State University
Fort Collins, Colorado
{skshimon, crlutton, lisaxu, sarahms, cboucher, jgruiz}@cs.colostate.edu

ABSTRACT
Although smartwatches are gaining popularity among mainstream consumers, the input space is limited due to their small form factor. The goal of this work is to explore how to design non-touchscreen gestures to extend the input space of smartwatches. We conducted an elicitation study eliciting gestures for 31 smartwatch tasks. From this study, we demonstrate that a consensus exists among the participants on the mapping of gesture to command and use this consensus to specify a user-defined gesture set. Using gestures collected during our study, we define a taxonomy describing the mapping and physical characteristics of the gestures. Lastly, we provide insights to inform the design of non-touchscreen gestures for smartwatch interaction.

Author Keywords
Smartwatch; Wearables; Interaction; Elicitation Study; Gestures; Think-Aloud.

ACM Classification Keywords
H.5.2. User Interfaces: Input devices and strategies, Interaction styles.

INTRODUCTION
A smartwatch is a wrist-mounted wearable computing device whose capabilities extend beyond just showing time. Some of the present smartwatch capabilities include, but are not limited to: running calculations, acting as a GPS tracking unit, tracking activity (i.e. counting steps, burned calories, monitoring heart rate), translating text, and working as an extension of mobile phones. In addition to these capabilities and attributes, other recent advancements--including mobile form factor, battery life, sensor capabilities, and low processor power—have contributed to the growth of smartwatches, which are now mainstream products in the current marketplace [23].

This recent ubiquity of smartwatches in the marketplace necessitates the exploration of the input and output capabilities of these devices. One of the major constraints in this exploration is the restricted touch space, which is a direct result of the small screen size. Due to this constraint, users are limited to providing input to the device by tapping and swiping the touchscreen. However, the small touchscreen results in *fat finger and occlusion* problems. The *fat finger* problem describes the issue of input errors caused by the relatively large size of a users' finger in contrast to the size of a target on the touchscreen. The *occlusion* problem describes the occlusion of a large portion of the viewable screen because of the relatively wide finger surface. These problems are more acute in smartwatches than smartphones because the screen size is significantly smaller.

To tackle these issues, new input techniques based on voice and non-touchscreen-based gestures have been proposed [23]. Android Wear, which is a voice-command based interaction method, is highly reliable for interacting with smartwatches [36], but becomes unreliable in public spaces because of environmental noise. This has inspired research in gesture-based interaction with smartwatches (e.g., [9,21]). The main focus of these studies is to extend the present touch-display input capability and reduce both the *fat finger* and *occlusion* problems.

These studies have mainly focused on hardware solutions and related gesture recognition algorithms with pre-defined gesture sets. There is limited research focusing on exploring users' non-touchscreen based gesture preferences in smartwatches. This is problematic since—as stated by Morris et al. [16]—gestures conceived by designers, as opposed to end-users, can sometimes fail to meet important design criteria since the small group of interaction designers fail to represent the larger population. Such design criteria include: *discoverability*, *ease-of performance*, *memorability* and *reliability*. In order to address this concern, an *elicitation study* [34] is needed to determine a set of user-defined non-touchscreen gestures. But until now, no elicitation studies have been performed to explore user preferences for possible gestures in smartwatches.

In this paper, we perform such an elicitation study. We provided 25 volunteers with a set of 31 tasks that can be performed on a smartwatch. The effects of these tasks were described to the users, and the users were asked to come up

CHI'16, May 07-12, 2016, San Jose, CA, USA
© 2016 ACM. ISBN 978-1-4503-3362-7/16/05…$15.00
DOI: http://dx.doi.org/10.1145/2858036.2858385

with appropriate gestures for the study, along with a set of user ratings for the gestures. Exploring the mental model of users for such gestures can result in giving smartwatch researchers and gesture designers better guidance to make smartwatch interaction more effective.

Our specific contributions are:

- A quantitative and qualitative characterization of user-defined non-touchscreen gestures for smartwatches;
- A user-defined consensus gesture set;
- Insight into users' mental model and their criteria for creating non-touchscreen gestures in smartwatches; and
- Implications for incorporating non-touchscreen gestures in smartwatches.

RELATED WORK

Smartwatch Systems Utilizing Non-touchscreen & Non-verbal Interactions

In light of both the fat finger and occlusion problems occurring with small touchscreens, researchers have proposed alternate input techniques for smartwatches that do not involve interacting with the touchscreen. *GestureWatch* [9] and *Hoverflow* [12] use mid-air gestures above mobile devices, smartwatches, and other wearable computing devices as input. *Abracadabra* [7] and *zSense* [33] use magnetic sensor-based around-the-device (air) gestures to interact with wearable devices. Xiao et al. [35] developed a watch prototype that supports pan, twist, binary tilt, and click on the watch face of a smartwatch. *Skinput* [8] and *Skinwatch* [21] make use of a user's skin to extend the gesture space for smartwatches and other wearable devices. Knibbe et al. [10] proposed using the back of the user's watch-wearing hand to enable manual and bimanual gestures for smartwatch interaction.

In addition to research that attempts to extend the size of the interface, there is a body of research that explores the use of non-gesture-based interaction. Akkil et al. [3] proposed and studied the use of facial glances and gazes as an alternate way for interacting with smartwatches. Song et al. [27] created and showed that a 2D RGB camera based gesture recognition system for mobile devices can be used for smartwatch interaction. *WatchMe* [32] is another camera-based gesture recognition technique for smartwatches that uses image processing/OCR to recognize input on a drawing canvas composed of everyday objects. *Blowatch* [4] allows users to blow air towards wearable devices to control interaction. Morganti et al. [15] showed that muscle-computer interfaces implemented in the wrist-band of a watch can recognize objects, grasps, and forearm gestures. *Bandsense* [2] combines touchscreen gestures with pressure-sensitive multi-touch gestures on a wristband to enable interaction. *Edgetouch* [20] uses sensor-enabled edges of a smartwatch to recognize touchscreen gestures around the edges of the watch.

Elicitation Studies for Gesture Design

An elicitation study is a widely used tool in user-centric computing to inform the design of gestures [30]. In an elicitation study, users are given the results/effects of performing a task or action. Participants have to come up with and perform gestures that they feel best match those effects. Elicitation studies have been performed to help guide the design of gestures in surface computing [17,34]. These have also been applied to determine single-handed and bimanual gestures on tabletops [34]; finger, body and remote based gestures to control the TV set [28,29,31]; hand gestures for augmented reality [22]; motion gestures [25], above-device gestures [6], and back-of-device gestures for mobile devices [26]. Each of these studies report the gesture sets for various application domains, as well as qualitative data such as users' evaluation of the ease of execution and the "fit-to-function" of proposed gestures. In addition, they provide insight into users' conceptual ideal about how they would interact with a specific technology or device. Lastly, these studies have shown that the preferences of a specific (elicited) gesture for a given task is influenced by technical expertise [11] and culture [13].

EXPLORING USER-DEFINED NON-TOUCHSCREEN GESTURE IN SMARTWATCH

To explore user-defined gestures, we elicited input from 25 participants. Participants were asked to design and perform a non-touchscreen gesture (*sign*) with a smartwatch device that could be used to execute a specific task (*referent*) on the smartwatch. Thirty-one tasks were presented to the participants during the study (Table 1).

In our study, we were careful not to constrain the users' behavior by the limitations of current gesture recognition technology. Instead, we sought to remove the *gulf of execution* [19] between end users' psychological goal and physical action. The participants were encouraged to focus on gesture design and assume all conceived gestures would be recognized by the smartwatch. Furthermore, they were not constrained to inventing a unique gesture for each of the given tasks, and therefore, could repeat their gesture for different tasks if they chose to do so. The only constraint we imposed upon the participants was that they could not touch the screen of the smartwatch while performing their gestures.

Each participant performed the gestures for each of the tasks given in Table 1. Each session was video recorded and took approximately one hour to complete. For each participant, a transcript of the recorded video was created to extract individual quotes as well as classify and label each gesture designed by the participant. The quotes were then clustered to identify common themes using a bottom-up, inductive analysis approach.

Selection of Tasks

One of our specific aims is to understand users' mental model of interaction with a smartwatch for different tasks. In addition, our study elucidates the following questions:

- Does orientation of tasks have any effect on the gestures people perform?
- How do users interact with tasks with similar but opposite effects?
- Do the users prefer symbolic gestures over simple tap and swipe gestures?
- Do the users repeat the same gestures based on context?

To understand users' mental model of interaction with smartwatches, we decided to test for the most common tasks that can be performed on different smartwatches. The tasks were chosen by analyzing functionalities of different smartwatches currently in the market. Similar to prior work [25,26], the tasks were grouped into two categories: *actions* and *navigation-based* tasks. Within these categories, we created two sub-categories: a task that can either act on the system/smartwatch (e.g., view/set time, opening menu bar of a smartwatch application) or task that acts on a particular application system (navigating a map). After grouping the tasks into these four sub-categories, a scenario representing each task was chosen for inclusion in the study. This method allowed us to create tasks that would be representative of the tasks used on a smartwatch but minimize task duplication.

Since a significant number of smartwatches can act as extensions for the users' smartphones [23], we included tasks traditionally associated with a smartphone (e.g., controlling and receiving notifications about phone calls). For these smartphone-related tasks, we imagined the user wearing headphones synced to the phone while the smartwatch acts as the secondary display for the phone. This scenario has the effect of keeping both of the user's hands free.

Participants

Twenty-five volunteers, ten female, recruited from a local university (12/25) and community (13/25) participated in our study. Participants were aged between 20-42 (Mean = 24.76, SD = 6.06) and all but three wore a watch on their left wrist. All participants owned a smartphone, but none of the participants had any prior smartwatch experience. The participants were compensated with a $20 Amazon gift card.

Procedure

The study began with the researcher explaining the study and providing the participant with a Moto 360 smartwatch to wear during the study. The purpose of the watch was to strictly act as a reference to size and did not provide any visual elements specific to tasks. We presented a total of 31 different tasks, which were broken up into 6 different groups based on the function of the task. We verbally described the action performed by the device and asked the user to create an input gesture that would activate the device action. We instructed participants to think aloud while making the gesture and to repeat their gesture one additional time. Next, the participants were asked some exploratory questions about the posture of the gesture. For example, the required duration of a finger press, or the number of fingers required to perform a zoom gesture. All verbal responses and gestures created by the participants were recorded using a video camera.

Category	Subcategory	Tasks
Navigation	System/ Smartwatch	Previous (Vertical)
		Next (Vertical)
		Previous (Horizontal)
		Next (Horizontal)
		Go to Home Screen
	Application	Pan Left
		Pan Right
		Pan Up
		Pan Down
		Zoom In
		Zoom Out
Action	System/ Smartwatch	Set Hr/Min/AM-PM
		Switch Between Hr/Min/AM/PM
		Confirm Time
		Start Stopwatch
		Stop Stopwatch
		View Time
		Act On Selection
	Application	Answer Call
		Hang up Call
		Ignore Call
		Mute Microphone (Call)
		Unmute Microphone (Call)
		Turn on Speaker (Call)
		Turn off Speaker (Call)
		Open Context Menu
		Switch Application
		Lock Screen
		Copy
		Cut
		Paste

Table 1. The list of tasks presented to participants grouped by category.

After the participants proposed a gesture that was appropriate for the intended task, the participants were asked to rate the gesture using an 11-point Likert scale on each of the following statements:

- The gesture I picked is a good match for its intended use.
- The gesture I picked is easy to perform.
- The gesture I picked is easy to remember.

We were also interested in exploring whether social context had any effect on the gesture preference. In order to accomplish this, participants were asked to rate their comfort tlevel with regards to performing their gesture in different environments and social contexts (shown in Table 2) on an 11-point Likert scale.

Table 2 lists the social contexts used. With 25 participants, (25*31) = 775 gestures were made and collected.

Environment	Social Context
Home	Alone
	With Family
Work	Alone
	Among Colleagues
Public	Among Strangers
	Among Friends

Table 2. The list of environments and social context used to explore social acceptability.

Data Analysis and Coding

Two researchers coded gestures independently using synchronized audio and video. This classified body part(s) used and their motion characteristics. Transcripts of the sessions were analyzed using grounded theory and an affinity diagram to discover themes.

RESULTS

During our study, the data we collected included transcripts, video recordings, non-touchscreen gestures designed by users, and user ratings of gestures. From this data, we present themes emerging from our interviews, taxonomy for non-touchscreen smartwatch gestures, and a user-defined consensus gesture set for smartwatch interaction.

Gesture Taxonomy

We constructed taxonomy for non-touchscreen gestures using the gestures collected from our elicitation study. Similar to Ruiz et al. [25], our taxonomy consists of two main taxonomy dimensions—*gesture mapping* and *physical characteristics*. Gesture mapping describes how gestures are mapped to different tasks by the participants. These include the *nature*, *context*, and *temporal* dimensions of the gesture. Physical characteristics describe the gesture characteristics themselves and include the *duration*, *size, complexity*, and *modality* dimensions of the gesture. The full gesture taxonomy is listed in Table 3.

Gesture Mapping

The *nature* dimension defines the mapping of the gesture to physical objects. This dimension can be viewed in the following ways:

- *Metaphoric gestures:* The gesture is a metaphor of another physical object. For example, to cut a piece of text on screen, the user makes a two finger scissor gesture above the smartwatch.
- *Physical gestures*: The gesture directly acts on screen content (i.e., direct manipulation).
- *Symbolic gestures:* The gesture depicts a symbol. For example, drawing a '3' in air above smartwatch.
- *Abstract gestures:* The gesture mapping is arbitrary.

Gesture Mapping		
Nature	Metaphor	Gesture is a metaphor of another physical object
	Physical	Gesture acts physically on object
	Symbolic	Gesture visually depicts a symbol
	Abstract	Gesture mapping is arbitrary
Context	In-context	Gesture requires specific context
	No-context	Gesture does not require specific context
Temporal	Discrete	Action occurs after completion of the gesture
	Continuous	Action occurs during the gesture
Physical Characteristics		
Duration	Short	Duration of the gesture is less than 0.5s
	Medium	Duration of the gesture is between 0.5 and 1.5s
	Long	Duration of the gesture is longer than 1.5s
Size	Small	Gesture can be performed in less than 439cm^3 of physical space
	Medium	Performing gesture requires between 439cm^3 and 1467cm^3 of physical space
	Large	Performing gesture requires over 1467cm^3 of physical space
Complexity	Simple	Gesture consist of a single gesture
	Compound	Gestures can be decomposed into simple gestures
Location	Rim	Gestures performed on the rim of the watch
	Band	Gestures performed on the watch band
	Skin	Gestures performed on the user's skin
	Mid-Air	Gestures performed in mid-air
	Multiple	Gestures that are performed in multiple locations.

Table 3. Taxonomy of non-touchscreen gestures for smartwatch interaction based on collected gestures and modeled after [21].

The *context* dimension describes whether a gesture requires a specific context or is performed independently. For example, the swipe right mid-air gesture is context specific (*in-context*). If performed while viewing a list, the content will scroll right, whereas, performing the gesture while the phone is ringing will answer the phone. In contrast, hovering the hand over the watch for a period of 2 seconds will lock the screen regardless of context, and therefore, is a context independent gesture.

Lastly, the *temporal* dimension describes if an action on an object occurs while or after making a gesture. In a discrete gesture, the action occurs after the gesture has been made – for example, a tap on the side of the watch to start/stop the application. In a continuous gesture, the action occurs while the gesture is ongoing–for example a swipe on the rim to scroll.

Physical Characteristics

The physical characteristics dimension of our taxonomy captures the characteristics related to a gesture's duration, size, complexity, and location.

The *gesture duration* describes the temporal requirements of performing a gesture and is divided into 3 categories: short (gestures taking less than 0.5 seconds), medium (gestures taking between 0.5 and 1.5 seconds), and long (gestures taking longer than 1.5 seconds). For example, short gestures include single taps and swipes on the rim of the watch. Double/triple taps and swipes above the device take a longer duration, but take less than 1.5 seconds and, therefore, are categorized as medium in duration. Lastly, an example of a gesture with a long duration would be hovering the hand over the watch face for more than 1.5 seconds.

The *size* dimension of our taxonomy describes the physical space required to perform the gesture, and is divided into the following categories:

- *Small:* The gesture movement can be performed in a region constrained by a 7.6cm cube (i.e., $439cm^3$ area) shown in Figure 1. These gestures involve extremely little physical movement of a small body part—like making a tap or scroll on rim/watch band, on-air tap with a single finger while keeping hand still.
- *Medium*: The gesture movement can be performed in a region constrained 12.7cm x 15.2cm x 7.6cm rectangular cuboid (area equal to $1467cm^3$). Examples include a single twist (rotation) of arm away from body, or making in-air swipes above the smartwatch in the area shown in Figure 1.
- *Large*: All gestures requiring 3D space larger than $1467cm^3$ are considered as large size gestures. These gestures include rotational motions along multiple body joints.

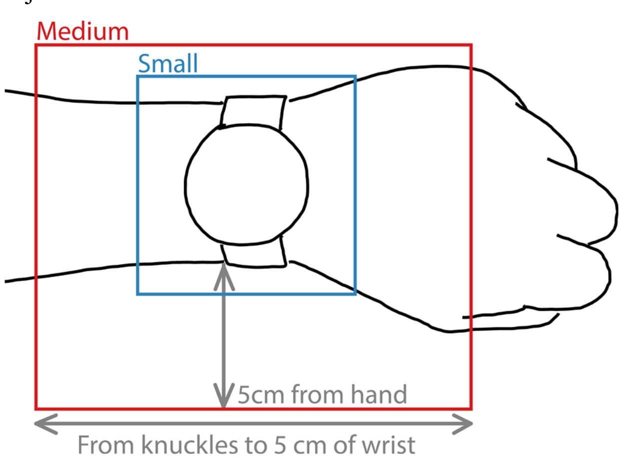


Figure 1: Illustration of the small and medium size dimensions of our taxonomy.

The *complexity* dimension of a gesture describes whether the proposed gesture can be decomposed into constituent gestures (compound gesture) or not. A pinching and pulling gesture used for map panning operations is an example of a compound gesture where the gesture can be divided into two spatial discontinuities–a *pinch* and a *pull.*

Lastly, the *location* dimension captures where, in relation to the user's body, a gesture is performed. Categories in the location dimension include: *rim, band, skin, mid-air,* and *multiple.*

Figure 2 illustrates the breakdown of the 775 gestures collected during the study using our taxonomy. As shown in the figure, gestures tended to be simple continuous mid-air gestures with medium duration.

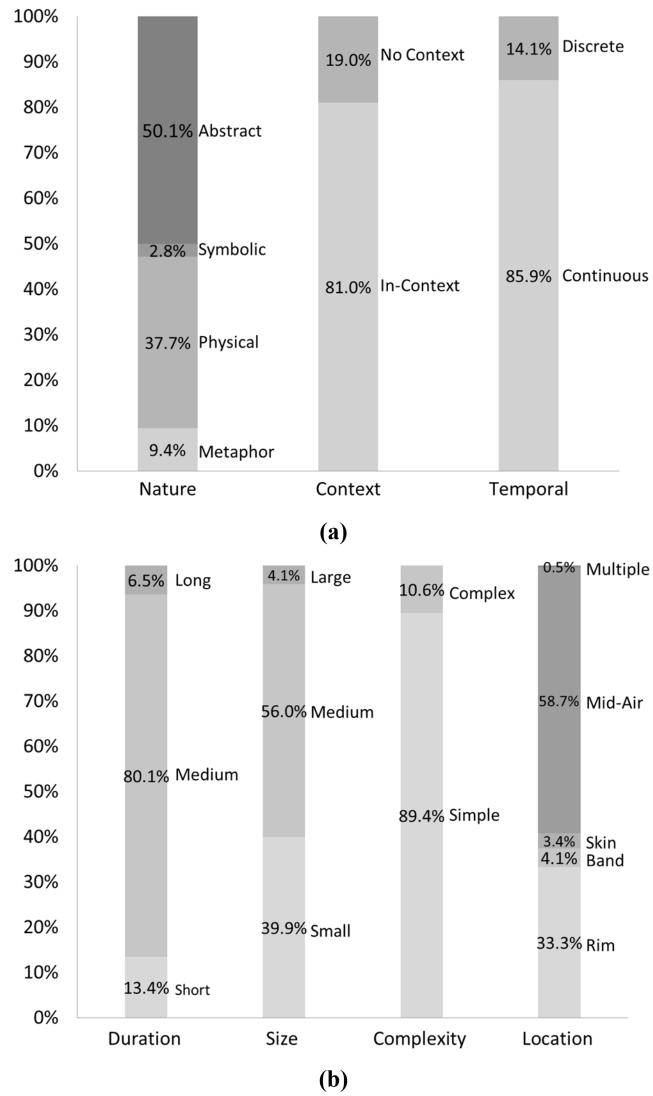


Figure 2. Percentage of gestures in each taxonomy category. (a) Illustrates the categories of the gesture mapping dimension. (b). Illustrates the categories of the physical characteristics dimension.

Designing Non-Touchscreen Smartwatch Gestures

We identified the following common themes using the videos and transcripts from our study: (1) mimicking touchscreen gestures, (2) natural and consistent mapping, (3) preference of simple and subtle gesture, and (4) feedback for non-tactile gestures. Each of these themes is discussed below.

Mimicking Touchscreen Gestures

For many tasks (e.g., scrolling and zooming), the participants designed gestures that mimicked touchscreen gestures. Participants who mimicked a front screen gesture often perceived their gesture as well-suited to the task, easier to perform, and easier to remember. For example, for scrolling tasks (previous/next), 70% of the conceived gestures involved some form of scrolling gesture which mimicked touchscreen scrolling. 58% of the conceived gestures for zooming in or zooming out involved some form of a 2-finger pinch/zoom gesture. When the participants were asked for their reasoning behind their chosen gesture, most of them often made comments describing it as the "*most natural*".

> *"To zoom in, the first gesture that comes to my mind is using my thumb and my index and pull them apart ... it is what I do on the screen of my phone and tablet." [P7]*

The directionality of touchscreen gestures was also mimicked when it was applicable. For example, 88.7% of the gestures conceived for previous and next scrolling mimicked the respective directionality of the corresponding touchscreen gesture. It is interesting to note that a participant's writing hand or watch-wearing hand had no effect on directionality of the gesture.

> *"To move to the next item in horizontal list, I prefer dragging one finger from right to left …when I am touching the screen, I scroll to the left."[P19].*

Natural and Consistent Mapping of Gestures

Two noticeable patterns were observed from the gestures that were elicited for tasks that have similar or opposite effects.

Unary Transition Task Pairs	Binary Transition Task Pairs
Mute Microphone - Unmute Microphone	Pan Left - Pan Right
Turn on Speaker - Turn off speaker	Pan Up - Pan Down
Start Stopwatch - Stop Stopwatch	Zoom In - Zoom Out
Answer Call – Hang up Call	Next (Vertical) - Previous (Vertical)
	Next (Horizontal) - Previous (Horizontal)

Table 4. List of unary and binary transition task pairs.

For *unary transition tasks*—tasks that cannot be performed before a task with the opposite effect is performed (e.g., start or stop a stopwatch)—the participants selected one of two strategies. The first strategy was to perform the same gesture twice (for both opposing tasks). For example, for muting and unmuting a microphone, participants performed the same gesture of a two-finger press on opposite sides of the smartwatch. The second strategy was to perform the same gesture in the opposite direction. Continuing with our muting example, some participants performed gestures with opposite effect (covering the watch face with their opposite hand for muting and uncovering it for unmuting). Overall, 57% of the participants envisioned the unary task pair shown in Table 4 as *toggling a switch* and repeated the same gesture.

Binary transition tasks are tasks that can be performed a consecutive number of times before performing a task with the opposite effect (e.g., pan left and right). For this type of gesture, participants always suggested similar gestures but in the opposite direction. For example, a horizontal swipe to the left was the most common gesture for the next task, and a horizontal swipe to the right was the most common gesture for the previous task.

Current touchscreen interfaces in mobile devices commonly require the user to move the content while the viewpoint remains static. We wanted to determine if users would continue to follow this paradigm. Results from our study showed that the participants preferred moving the content rather than moving the viewpoint. Users stated that this was a result from them being more comfortable mimicking the touchscreen gestures of other technologies.

In addition, we wanted to understand if task orientation or content layout had any effect on the gestures participants proposed. Therefore, we presented two lists in both a horizontal and vertical layout. Our results demonstrated that the orientation almost always affected the participants' choice in gesture, with gesture direction being consistent with the orientation of the task.

Preference of Simple and Subtle Gestures

The participants preferred user-defined gestures that were classified as being *simple*, believing that simple gestures are easier to perform and remember. The swipe, tap, pinch with the alternate hand, and twist of watch-wearing hand gestures are examples of the *simple* gestures that were elicited for our tasks. For example, participant P8 used a single tap on the outer rim of watch near the hand for starting and stopping the stopwatch, a two finger pinch on the top-bottom rim of watch for muting/unmuting, and a pinch on the left-right rim for turning the speaker on / off. These gestures were *simple* and *short* in duration. In addition, natural body movements—such as putting the watch-wearing hand beside the body—were also preferred because they were perceived as subtle and more socially acceptable.

For making air gestures with the non-watch wearing hand, most participants preferred the enclosed area shown in Figure 1, making the gesture simple and subtle.

Feedback for Non-Touchscreen Gestures

There was a strong feeling among the participants that feedback accompanying the gesture is important. While they expected visual feedback to be displayed on the screen, participants also stated a preference for additional feedback through vibration and/or sound. The vibration feedback was stated to be favored sound. Hence, a majority of participants gave similar opinion to P6 who said

> *"I prefer vibration to sounds coming from [the] watch ... I may not hear the sound when I am in public but I can definitely feel the watch vibrating." [P6]*

In addition, feedback was deemed to be more important in some tasks than others, with there being a significant consensus among participants to receive feedback for hang-up call, ignore call, act on selection, and confirm time.

> *"For every air-scroll to move to the next or previous item, a vibration is unnecessary and irritating. But when I am terminating a call, a vibration feedback is useful – this lets me know if I accidentally terminated the call." [P12]*

A User-defined Gesture Set

A user-defined gesture set for our specific tasks was generated using the set of all 775 elicited gestures collected from our participants. For each task, identical gestures were grouped together, and the group with the largest consensus was chosen to be the *representative gesture* for the task. Ties in group size were broken by using the subjective ratings. We refer to this gesture set as both our *consensus set* and our *user-defined gesture set*. The user-defined gesture set is shown in Figure 3.

We used the agreement score standard [30] to extract the consensus among participants for each task. Similar to other elicitation studies [25,26,34], agreement scores range between 0.4 and 0.1. The overall agreement score (A) is 0.16. Similarly to Wobbrock et al. [34], we rated each task's conceptual complexity independently. Referents' conceptual complexities correlated significantly and inversely with the agreement scores ($r = -0.743$, $F_{1,\ 29} = 35.684$, $p < 0.01$). In general, we found that as conceptual complexity of the task increased, participant agreement decreased.

Social acceptability

Analysis of the collected social acceptability ratings for the elicited gestures revealed several findings. Firstly, alternate hand gestures that continued and went beyond the enclosed areas illustrated in Figure 1, were rated to be less socially acceptable. Hence, the participants did not feel comfortable performing such gestures in public and office environments. Next, touch gestures on the watch rim and watch band received high social acceptability ratings. All users felt comfortable performing such gestures in populated social contexts.

Moreover, small and medium sized gestures received nearly perfect social acceptability rating in all cases. The results of this rating showed that participants were comfortable making *small* and *medium* gestures in public and office settings, and *large* gestures were only deemed comfortable in non-public settings (e.g., being alone or being among family). All participants echoed the sentiment of participant P10, who said the following:

> *"I don't want to attract too much attention to myself." [P10]*

This was re-iterated by participant P15:

> *"When I am among people, I don't prefer making gestures that attract attention from people and makes [sic] me look crazy." [P15]*

Our findings support prior work exploring the social acceptability of gestures [14,18,24]. Gestures that can be performed without drawing a lot of attention or cannot easily be interpreted by bystanders are considered socially appropriate. In our case, gestures that can be performed in the interaction volume described by Figure 1 are considered socially appropriate gestures for our context.

DISCUSSIONS

In this section, we discuss the implications of our results to designing for non-touchscreen based gestural interaction on a smartwatch.

Legacy Bias

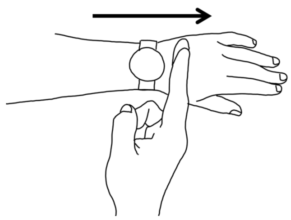
Participants in our study had no prior experience with using a smartwatch. However, participants showed a considerable amount of legacy bias from using touchscreen devices, desktop-based systems, analog dial watches, and stopwatches. For example, during tasks such as starting and stopping the stopwatch or changing the time, some of the users mimicked the actions they performed on analog watches (i.e., *winding a screw using two finger pinch* to set the time) to perform the task. While legacy bias hinders discovering new gestures, Kopsel et al. [11] argue legacy bias eases the transition of interacting with new gesture paradigms.

Preference of Different Touch Gestures

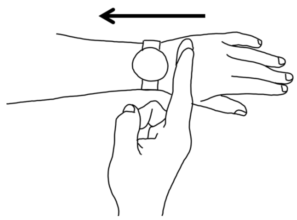
For touch-based gestures, users showed significant preference for touching the watch rim over touching the band or skin. The rim was used gestures for both vertical (scrolling/panning up and down) and horizontal (scrolling/panning side to side) directions, whereas the band was used only in scrolling up and down.

> *"For scrolling left or right, the band is not wide enough." [P2]*

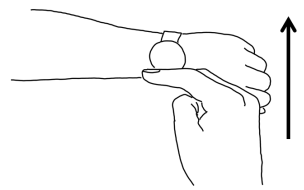
One notable pattern was the preference for using the "outer" half of the rim located towards the hand for vertical



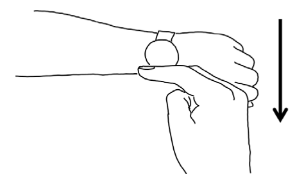
Above device air-swipe from left to right : Pan Left/Previous (Vertical)/Receive Call / Clock hand switch from Hr-Min



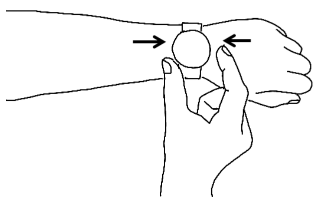
Above device air-swipe from right to left : Pan Right / Next (Vertical)/ Hang-up Call / Ignore Call



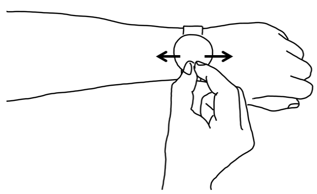
Above-Device air-swipe from down to up: Pan Down / Next (Horizontal)



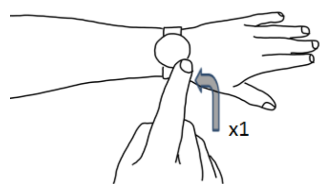
Above-Device air-swipe from up to down: Pan Up / Previous (Horizontal)



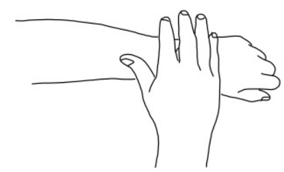
Above device air-pinch: Zoom out/ Turn off speaker



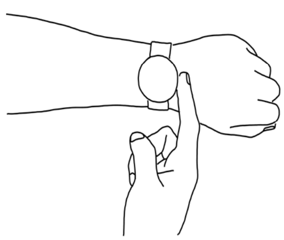
Above device air-zoom: Zoom in / Turn on speaker



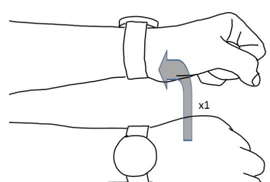
Above device air quick tap (once): Act on selection/



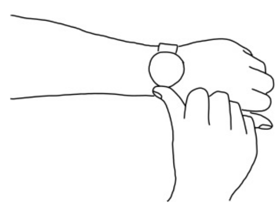
Cover watch face without touching with open palm: Mute Microphone (Call)



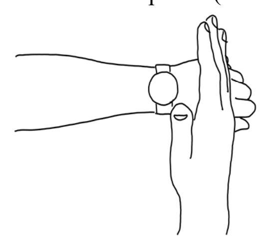
Single tap on outer half rim of watch (half near hand): Start/stop stopwatch/ Confirm time



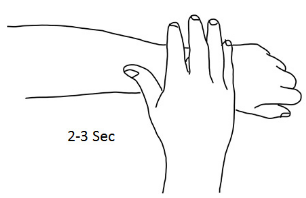
Quick twist wrist away from body once: View time / Home screen



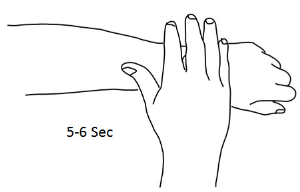
Single tap on bottom half of watch rim: Open Context Menu



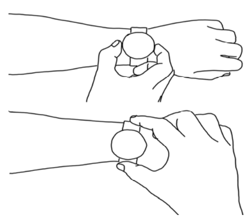
Take watch face covering open palm off: Unmute Microphone (Call)



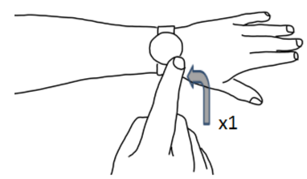
Short hover open palm above watch (2-3 seconds): Switch application



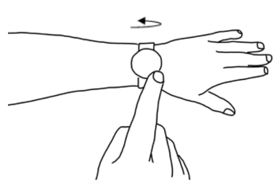
Long hover open palm above watch (5-6 seconds): Lock screen



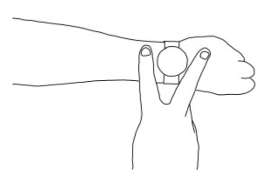
2 finger pinch on opposite side of watch: Copy



Above device air long press (1/2 finger): Paste/



Drawing in air above-device: Setting Hr/Min/AM-PM



Cut

Figure 3: Consensus set of gesture of non-touchscreen gestures for smartphone interaction obtained from participant elicited gestures

swiping and tapping gestures, and the "bottom" half of rim located closer to thumb for horizontal swiping. A user who wore the target watch on his left hand said the following for next/previous item scrolling in vertical direction:

> *"I am swiping on the left half because the right half contains the watch notch. If the watch notch was not there, I would prefer swiping on the right side."[P22]*

The target watch contained a notch on the right half, which fell on the side of the hand when being worn on the left wrist. For touch based tapping gestures, most taps tended to occur on the outer half of the watch (near three o'clock) and mostly with the index finger. The second largest number of physical taps occurred on the bottom half of the rim near 6 o'clock, mostly with the thumb. Users were asked about their preferences for different physical tap gestures. Most participants stated that they prefer single finger taps, adding that the finger preferred for performing taps depends on the location of the gesture.

Several types of physical 2-finger pinch gestures were observed during the study. The most common place to perform a 2-finger pinch was on the watch rim. Two finger presses almost always tended to be on opposite halves of the rim–mostly along the arm axis. 76% of 2-finger pinch observed on the watch rim was along the arm axis.

Map panning and zooming gestures:
Some participants used above-device swipes for map panning and touch-based scrolls for moving to next and previous items. One of the participants who did this said the following:

> *"For scrolling, you go up-down or left-right… for panning, you can pan in x-y axis."[P9]*

Non-touch gestures for map panning and zoom were almost always made directly above the watch. When asked about potential occlusion on screen, one participant said:

> *"There is a gap between the hand and watch – I can see more than touching…"[P6]*

The concept of physical target acquisition seemed prevalent when users were trying to map points on the screen to any alternate gesture space. Participants were more comfortable mapping the screen of the watch with the space directly above the watch compared to back of the hand. As a result, even though the back of the watch wearing hand offers the possibility of making 2D scroll gestures, the participants preferred above-device gestures for map panning operations.

A common theme observed during the mapping of the zoom operations was the association of finger spreading and pinching gestures—a legacy bias from using touchscreens as discussed before.

Another observation was the use of finger rotation for zooming operations. This rotation was observed in the form of finger swipes along the watch rim. For example, participant P7 touched and rotated the rim of the watch clockwise to zoom in and counterclockwise to zoom out and gave the following explanation:

> *"To open a bottle I turn the cap counterclockwise and clockwise to close" [P7]*

In this case, the participant was mapping zooming out with the motion of opening the cap of a bottle and zooming in with closing the cap.

Concern for Accidental Triggering:
Users were concerned about accidental touches, taps and swipes on the smartwatch, and subsequent accidental action triggering. Participants showed a conflict between designing simple gestures—and gestures that were more deliberate and less prone to accidental triggering. For example, for the *act on selection* task, a significant minority of participants made a double air-tap gesture above the screen—considering a double tap gesture more distinct and less likely to be accidental.

Preference of Non-touchscreen Gestures over Touchscreen Gestures:
Participants were asked when they would prefer using gestures over touching the screen. Users stated that they would employ a combination of touch and non-touch gestures, depending on the context of use. Some of the scenarios in which users stated non-touchscreen gestures would be more appropriate than other interaction methods included performing tasks where fingers are dirty or interacting with the screen would soil the smartwatch (e.g., cooking or cleaning); situations where on screen items are difficult to acquire or interacting with the screen would cause occlusion (e.g., users specified that using gestures to scroll a list to find an item is easier and more efficient than using the touchscreen); and when gestures provide a "shortcut" to actions that are not readily available on the touchscreen (e.g., muting the microphone or turning on/off the speaker).

Comparison to Prior Proposed Non-touchscreen Gestures
As stated in the related work section, researchers have proposed several non-touchscreen gesture sets motivated by preventing users from occluding the screen. Statements by our participants supported the need for interaction techniques that did not obstruct view of the screen (i.e., non-touchscreen gestures). However, very few participants suggested the previously proposed gestures for the same tasks. We attribute this to the fact that most prior work focused on a specific type of non-touchscreen gestures (e.g., the band [2], watch face [35], mid-air [7], or hand [8]), whereas, we were more open-ended about the types of gestures that participants could perform. The small number of times where our participants mimicked prior designer based gestures occurred when that gesture could be seen as having high legacy bias. For example, *Bandsense* [2] suggested a tap on the band and Knibbe et al. [10] a single

tap on the back of the watch. Some of our participants proposed similar gestures, with the majority of the users proposing mid-air tap above the watch face. All of which can easily be attributed to the legacy bias associated with interacting with touchscreens.

Implications for Gesture Recognition of Non-touchscreen gestures

Similar to other elicitation studies, we did not consider how to track gestures. The goal of our study is to understand users' mental models, regardless of technology. However, results from our work present several implications regarding what technology would be needed to recognize the gestures in our user-defined gesture set. More specifically, our results show that users prefer short (between 0.5-1.5 seconds), simple gestures constrained by the medium region shown in Figure 1. This suggests that very little hardware may be required to enable this type of interaction on smartwaches and that approaches such as using infrared hardware similar to HoverFlow [12] may be more appropriate than those that use complex depth cameras (e.g., that of Air+Touch [5]).

FUTURE WORK

Our study was limited by cultural and social demographics since all our participants were educated adults who lived in a Western culture. The cultural and social norms observed by our participants were more likely to be homogeneous and we could not ascertain that these same gestures would have the same acceptance in a different culture. For example, the scissor gesture (see Figure 4) means "two" in the western culture but "go to hell" in Greek culture [1]. One important extension of our work is to continue this elicitation study across different cultures with the help of online tools, and determine a user-defined gesture set that is appropriate for other cultures as well.

CONCLUSION

We presented the results of an elicitation study for making non-touchscreen gestures in smartwatches. Based on the elicitation study, we present users' mental model of visualizing and mapping gestures. From the observed patterns in the gestures, we present taxonomy and a set of heuristics which inform non-touchscreen gesture designing in smartwatches. We believe this research can specifically complement present hardware and algorithmic research on gesture recognition in smartwatches. We also believe common themes identified in this research can help to tackle input space problems with small screen wearables by suggesting design guidelines for alternate gestural interactions.

REFERENCES

1. Hajar Ghadirian Najaf Abadi, Lim Yan Peng, and Ali Mohammad Hossein Zadeh. 2012. Guessability Study on Considering Cultural Values in Gesture Design for Different User Interfaces. *International Proceedings of Economics Development & Research* 37.
2. Youngseok Ahn, Sungjae Hwang, HyunGook Yoon, Junghyeon Gim, and Jung-hee Ryu. 2015. BandSense: Pressure-sensitive Multi-touch Interaction on a Wristband. *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, ACM, 251–254. http://doi.org/10.1145/2702613.2725441
3. Deepak Akkil, Jari Kangas, Jussi Rantala, Poika Isokoski, Oleg Spakov, and Roope Raisamo. 2015. Glance Awareness and Gaze Interaction in Smartwatches. *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, ACM, 1271–1276. http://doi.org/10.1145/2702613.2732816
4. Wei-Hung Chen. 2015. Blowatch: Blowable and Hands-free Interaction for Smartwatches. *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, ACM, 103–108. http://doi.org/10.1145/2702613.2726961
5. Xiang "Anthony" Chen, Julia Schwarz, Chris Harrison, Jennifer Mankoff, and Scott E. Hudson. 2014. Air+touch: interweaving touch & in-air gestures. ACM Press, 519–525. http://doi.org/10.1145/2642918.2647392
6. Euan Freeman, Stephen Brewster, and Vuokko Lantz. 2014. Towards Usable and Acceptable Above-device Interactions. *Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices & Services*, ACM, 459–464. http://doi.org/10.1145/2628363.2634215
7. Chris Harrison and Scott E. Hudson. 2009. Abracadabra: Wireless, High-precision, and Unpowered Finger Input for Very Small Mobile Devices. *Proceedings of the 22Nd Annual ACM Symposium on User Interface Software and Technology*, ACM, 121–124. http://doi.org/10.1145/1622176.1622199
8. Chris Harrison, Desney Tan, and Dan Morris. 2010. Skinput: Appropriating the Body As an Input Surface. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 453–462. http://doi.org/10.1145/1753326.1753394
9. Jungsoo Kim, Jiasheng He, Kent Lyons, and Thad Starner. 2007. The Gesture Watch: A Wireless Contact-free Gesture Based Wrist Interface. *Proceedings of the 2007 11th IEEE International Symposium on Wearable Computers*, IEEE Computer Society, 1–8. http://doi.org/10.1109/ISWC.2007.4373770
10. Jarrod Knibbe, Diego Martinez Plasencia, Christopher Bainbridge, et al. 2014. Extending Interaction for Smart Watches: Enabling Bimanual Around Device Control. *CHI '14 Extended Abstracts on Human Factors in Computing Systems*, ACM, 1891–1896. http://doi.org/10.1145/2559206.2581315

11. Anne Köpsel and Nikola Bubalo. 2015. Benefiting from Legacy Bias. *interactions* 22, 5: 44–47. http://doi.org/10.1145/2803169
12. Sven Kratz and Michael Rohs. 2009. HoverFlow: Expanding the Design Space of Around-device Interaction. *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services*, ACM, 4:1–4:8. http://doi.org/10.1145/1613858.1613864
13. Dan Mauney, Jonathan Howarth, Andrew Wirtanen, and Miranda Capra. 2010. Cultural Similarities and Differences in User-defined Gestures for Touchscreen User Interfaces. *CHI '10 Extended Abstracts on Human Factors in Computing Systems*, ACM, 4015–4020. http://doi.org/10.1145/1753846.1754095
14. Calkin S. Montero, Jason Alexander, Mark T. Marshall, and Sriram Subramanian. 2010. Would you do that? *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services - MobileHCI '10*, 275. http://doi.org/10.1145/1851600.1851647
15. Elisa Morganti, Leonardo Angelini, Andrea Adami, Denis Lalanne, Leandro Lorenzelli, and Elena Mugellini. 2012. A smart watch with embedded sensors to recognize objects, grasps and forearm gestures. *Procedia Engineering* 41: 1169–1175.
16. Meredith Ringel Morris, Andreea Danielescu, Steven Drucker, et al. 2014. Reducing Legacy Bias in Gesture Elicitation Studies. *interactions* 21, 3: 40–45. http://doi.org/10.1145/2591689
17. Meredith Ringel Morris, Jacob O. Wobbrock, and Andrew D. Wilson. 2010. Understanding Users' Preferences for Surface Gestures. *Proceedings of Graphics Interface 2010*, Canadian Information Processing Society, 261–268. Retrieved from http://dl.acm.org/citation.cfm?id=1839214.1839260
18. Michael J. Muller and Sandra Kogan. 2012. Grounded Theory Method in Human-Computer Interaction and Computer Supported Cooperative Work. In *The Human-Computer Interaction Handboook*. CRC Press, Boca Raton [u.a.], 1003–1023.
19. Donald A. Norman and Stephen W. Draper. 1986. *User Centered System Design; New Perspectives on Human-Computer Interaction*. L. Erlbaum Associates Inc., Hillsdale, NJ, USA.
20. Ian Oakley and Doyoung Lee. 2014. Interaction on the Edge: Offset Sensing for Small Devices. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 169–178. http://doi.org/10.1145/2556288.2557138
21. Masa Ogata and Michita Imai. 2015. SkinWatch: Skin Gesture Interaction for Smart Watch. *Proceedings of the 6th Augmented Human International Conference*, ACM, 21–24. http://doi.org/10.1145/2735711.2735830
22. Thammathip Piumsomboon, Adrian Clark, Mark Billinghurst, and Andy Cockburn. 2013. User-defined Gestures for Augmented Reality. *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, ACM, 955–960. http://doi.org/10.1145/2468356.2468527
23. Reza Rawassizadeh, Blaine A. Price, and Marian Petre. 2014. Wearables: Has the Age of Smartwatches Finally Arrived? *Commun. ACM* 58, 1: 45–47. http://doi.org/10.1145/2629633
24. Julie Rico and Stephen Brewster. 2010. Usable gestures for mobile interfaces: evaluating social acceptability. *Proceedings of the 28th international conference on Human factors in computing systems*, ACM, 887–896. http://doi.org/http://doi.acm.org/10.1145/1753326.1753458
25. Jaime Ruiz, Yang Li, and Edward Lank. 2011. User-defined Motion Gestures for Mobile Interaction. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 197–206. http://doi.org/10.1145/1978942.1978971
26. Shaikh Shawon Arefin Shimon, Sarah Morrison-Smith, Noah John, Ghazal Fahimi, and Jaime Ruiz. 2015. Exploring User-Defined Back-Of-Device Gestures for Mobile Devices. *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services*, ACM, 227–232. http://doi.org/10.1145/2785830.2785890
27. Jie Song, Gábor Sörös, Fabrizio Pece, et al. 2014. In-air Gestures Around Unmodified Mobile Devices. *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*, ACM, 319–329. http://doi.org/10.1145/2642918.2647373
28. Radu-Daniel Vatavu. 2012. User-defined Gestures for Free-hand TV Control. *Proceedings of the 10th European Conference on Interactive Tv and Video*, ACM, 45–48. http://doi.org/10.1145/2325616.2325626
29. Radu-Daniel Vatavu. 2013. A Comparative Study of User-defined Handheld vs. Freehand Gestures for Home Entertainment Environments. *J. Ambient Intell. Smart Environ.* 5, 2: 187–211.
30. Radu-Daniel Vatavu and Jacob O. Wobbrock. 2015. Formalizing Agreement Analysis for Elicitation Studies: New Measures, Significance Test, and Toolkit. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, ACM, 1325–1334. http://doi.org/10.1145/2702123.2702223
31. Radu-Daniel Vatavu and Ionut-Alexandru Zaiti. 2014. Leap Gestures for TV: Insights from an Elicitation Study. *Proceedings of the 2014 ACM International Conference on Interactive Experiences for TV and Online Video*, ACM, 131–138. http://doi.org/10.1145/2602299.2602316
32. Wouter Van Vlaenderen, Jens Brulmans, Jo Vermeulen, and Johannes Schöning. 2015. WatchMe: A Novel Input Method Combining a Smartwatch and Bimanual Interaction. *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human*

Factors in Computing Systems, ACM, 2091–2095. http://doi.org/10.1145/2702613.2732789

33. Anusha Withana, Roshan Peiris, Nipuna Samarasekara, and Suranga Nanayakkara. 2015. zSense: Enabling Shallow Depth Gesture Recognition for Greater Input Expressivity on Smart Wearables. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, ACM, 3661–3670. http://doi.org/10.1145/2702123.2702371
34. Jacob O. Wobbrock, Meredith Ringel Morris, and Andrew D. Wilson. 2009. User-defined Gestures for Surface Computing. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 1083–1092. http://doi.org/10.1145/1518701.1518866
35. Robert Xiao, Gierad Laput, and Chris Harrison. 2014. Expanding the Input Expressivity of Smartwatches with Mechanical Pan, Twist, Tilt and Click. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 193–196. http://doi.org/10.1145/2556288.2557017
36. 2015. Android Wear. *Wikipedia, the free encyclopedia*. Retrieved September 14, 2015 from https://en.wikipedia.org/w/index.php?title=Android_Wear&oldid=680738338