Designing Consistent Gestures Across Device Types: Eliciting RSVP Controls for Phone, Watch, and Glasses

Tilman Dingler[1], **Rufat Rzayev**[2], **Alireza Sahami Shirazi**[2*], **Niels Henze**[2]
[1]Osaka Prefecture University, Osaka, Japan, tilman@cs.osakafu-u.ac.jp
[2]University of Stuttgart, Stuttgart, Germany, {firstname.lastname}@vis.uni-stuttgart.de

ABSTRACT
In the era of ubiquitous computing, people expect applications to work across different devices. To provide a seamless user experience it is therefore crucial that interfaces and interactions are consistent across different device types. In this paper, we present a method to create gesture sets that are consistent and easily transferable. Our proposed method entails 1) the gesture elicitation on each device type, 2) the consolidation of a unified gesture set, and 3) a final validation by calculating a transferability score. We tested our approach by eliciting a set of user-defined gestures for reading with Rapid Serial Visual Presentation (RSVP) of text for three device types: phone, watch, and glasses. We present the resulting, unified gesture set for RSVP reading and show the feasibility of our method to elicit gesture sets that are consistent across device types with different form factors.

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation

Author Keywords
Consistency; Gesture Elicitation; Design Methods; RSVP

INTRODUCTION
Digital devices have reached all aspects of people's life. One development that made this pervasive interaction with computers possible is their increasing diversity. Smartphones and smart watches are used to access digital information at almost any time and place. Already today, a typical user owns multiple devices and, depending on the context, uses the same type of application on different devices. Email, messaging, and reading applications, for example, exist for all major platforms and, consequently, users can choose which device they want to use for accessing a particular service.

Interface guidelines help designers to implement consistent interfaces for individual devices. Examples include the design guidelines by Google for developing Android applications and Microsoft's Windows User Experience Interaction Guidelines.

CHI 2018, April 21–26, 2018, Montréal, QC, Canada

DOI: https://doi.org/10.1145/3173574.3173993

Figure 1. Consistent, user-defined gesture sets for controlling reading flow via RSVP on three device types: phone, watch and glasses.

Such guidelines ensure a consistency that helps users who learned to use one application to transfer this knowledge to other applications. This is especially beneficial for gestural interfaces: it is challenging to communicate the available gestures, which makes gestures hard to discover and learn. Using the same type of application on different devices poses challenges to the interaction consistency. Each type of device has its own means for input and output. Despite being developed for similar use cases, smart watches and smart glasses, for example, offer different input and output mechanisms. As device diversity increases, using an application on one particular device must allow users to transfer interaction knowledge to similar applications on other devices.

Previous work intensively studied approaches for developing gesture sets [23, 37, 39]. Guessability studies [38], asking potential users to propose gestures for given actions, have been highly effective for developing intuitive gestures. A body of work applied guessability studies to a wide range of applications and devices [1, 12, 15]. While guessability studies have been widely adopted to develop intuitive gestures for individual devices, much less work investigated its use across device boundaries. Vatavu [36] took a first step towards an understanding of cross-device consistency for gestures by conducting a study to develop gestures for two devices. He reported that only 9 out of 22 gestures were shared across the two devices, hence conducting guessability studies for different devices individually does not ensure cross-device consistency of the resulting gestures.

In this paper, we develop a deeper understanding of creating gesture sets that are consistent across devices. We design gestures sets to control Rapid Serial Visual Presentation (RSVP) of text. RSVP has been proposed as a technique for displaying text on very small screens. It shows text word by word in one focal point and is therefore suitable for being used on

* The work was conducted when he was a researcher at the University of Stuttgart.

small-screen displays, such as phones, watches, and glasses. It can also increase reading speed [9] but often at the cost of text comprehension: Schotter *et al.* [29] found regressions (*i.e.*, the rereading of words) to be crucial for text understanding, which is not directly supported by RSVP reading. The same accounts for pauses: as readers need to be able to mentally digest what has just been read, controls for pausing the text flow should be accessible. Dingler *et al.* [4] report of readers feeling alienated when having no control over the reading process.

Using RSVP to display text is not restricted to a particular device. It is, therefore, important that interactions are consistent across different devices. We aimed to develop gesture sets to control reading flow for three different mobile devices, namely 1) phones, 2) watches, and 3) glasses. We elicit gesture sets for the most fundamental reading controls, such as play, pause, changing reading speed, and text position. Current systems, such as *Spritz*[1], offer controls that are dictated by the application developer or designer. This may result in arbitrary controls and gesture sets. Our approach combines the advantages of gesture consistency and user acceptance [23]. We ensure consistency by conducting a single study to elicit gestures for all three device types, consider proposed gestures for all devices when constructing the final gesture sets, and validate the transferability of the gestures in a validation study.

This paper presents the following three contributions:

1. Three user-defined gesture sets for reading via RSVP that are consistent across three different device types.

2. A system, which implements and validates the elicited gesture sets and allows users to transfer interaction knowledge between the three device types.

3. An in-depth discussion resulting in a method for creating consistent gesture sets across device types.

BACKGROUND

Our work is at the crossroads of reading on mobile devices and gesture elicitation, which we discuss in the following.

Screen Size, Reading, and RSVP

A strand of research assessed the effect of screen size on text presentation and reading behavior. An early investigation by Duchnicky and Kolers [6] explored the effect of window height and line width finding four lines to be an optimal height for smaller window sizes and two lines having a negative impact on reading performance. Shneiderman [33] reported that the number of lines, in which text was presented, did not significantly affect reading speed. Marshall & Ruotolo [19], however, found that screen size influenced the type of reading users do: small displays were used for casual reading, whereas, larger displays were rather used for opportunistic and intensive reading. This suggests that the choice of display size or device type could also be subject to the type of reading the user intends to do. Reading applications should, therefore, support cross-device reading activities.

[1] http://www.spritzinc.com

Forster [8] introduced the term RSVP as an experimental model for examining temporal characteristics of attention. Visual cues are thereby presented at a single focal place. Instead of requiring users to move their focus point to see further content, the visual cue is exchanged, which eliminates the need for saccadic eye movements. The capability of the human visual system, however, is a limiting factor for the application of RSVP. Presentation rate and the visual similarity of stimuli affect the effectiveness of RSVP streams [25]. Raymond *et al.* [26] found that people fail to detect subsequent targets occurring in succession when stimuli are presented in rapid succession ($180-450ms$), a phenomenon they described as *attentional blink*. It is therefore vital to provide user controls and allow speed adjustments as users' susceptibility depends on various contextual factors, such as current attention level, fatigue, and distractions. Spence [35] gives an account of different modes of RSVP. In our work, we focus on the sequential presentation of words in one spot. RSVP has further been investigated on mobile phones, including simple controls, such as speeding up, slowing down and going back to the last punctuation mark [11].

Other RSVP reading studies focused on memory effects finding that RSVP leads to rather coarse recall of text [20] and due to the suppression of regressions (*i.e.*, voluntarily or involuntarily re-reading a piece of text) can hinder text comprehension [29]. For devices with small screens, RSVP allows to trade space for time, which makes this technique so interesting for small devices, such as mobile phones and watches. Georgiev [9] used RSVP to investigate reading speeds on mobile devices and compared it to reading on PC screens and paper. Even though top reading speeds were achieved on computer screens and on paper, RSVP was still rendered feasible as a reading technique on mobile phones. Dingler *et al.* [4] assessed the feasibility of RSVP to explicitly increase reading speed on electronic devices. Their studies report on a substantial learning effect for users after initial alienation by being dictated how to read, one of the main reasons for this being a lack of control over the reading flow. Recently, RSVP has gained popularity due to a commercial application called *spritz* employing a text presentation technique on mobile phones and smart watches. To control the reading flow and speed, interaction techniques are limited to conventional buttons.

Gestures and Eliciting User Input

Touch gestures are an effective interaction technique with mobile devices as they neither require dedicated buttons nor screen space. One of the first investigations into human gestures was conducted by Efron [7]. Later taxonomies, such as McNeill's studies on gestures [21], are based on human discourse and extended this work. Poggi [24] presented a typology of gestures and their relationship to verbal signals, memory, meaning, and motivation. Hence, gestures relate to other signals, cognitive constructions, gesture-meaning relationships and semantic content.

Motivated by the widespread use of gestures for communication between humans, a large body of work proposed the use of gestures to interact with computing systems. Already in the

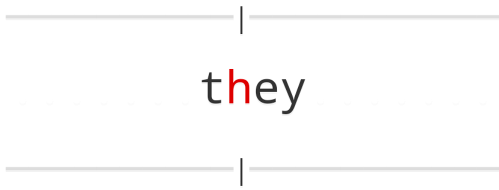

Figure 2. RSVP reading interface used in the application set. A red letter marks the optimal viewing position.

1980s, Bolt [3] proposed the use of pointing gestures to interact with projected objects. Especially early work focused on challenges to recognize gestures [13, 28] often neglecting the human factor. Proposing a human-centered approach to design gesture sets, Nielsen *et al.* [23] concluded that technology-based approaches lead to awkward gestures without intuitive mapping towards functionality and systems which only work under strictly pre-defined conditions. As an alternative, they presented a human-centered approach to designing gestures. They describe what has later been coined as guessability studies [38], in which referents, such as *actions* that a system should perform, are shown to participants to elicit the *symbols*, such as gestures, meant to invoke them. Subsequently, metrics were provided to calculate a guessability score for an existing set of symbols as well as agreement scores to describe the agreement among symbols proposed by study participants [38]. Wobbrock *et al.* [39] later showed how to use this method to elicit new gestures by developing user-defined gestures for surface computing, which was further improved by Vatavu *et al.* [37]. This general approach has been used and adapted by a substantial body of work, including developing gestures for mobile phones [27], head-worn displays [31], smart homes [16], cars [5], flexible displays [17], as well as for visually impaired users [14]

While these guessability studies have been widely adopted, much less research investigated the consistency of gestures across devices. Vatavu [36] conducted the first study that compared user-defined gestures for different capture technologies, in which gestures were elicited for handheld devices. While this work focused on analyzing the two resulting gesture sets, Vatavu also reports that only 9 out of 22 commands were shared between the two sets. Inspired by previous guessability studies, Anthony *et al.* [2] used existing gesture sets and shapes to understand how consistent they are performed with fingers and pens. For given shapes and gestures, they found a high consistency within participants but a lower consistency between-participants. While recent work showed that gestures elicitated by using different devices are not necessarily consistent, previous work did not investigate how to ensure consistency across devices. We, therefore, base our work on previous work describing guessability studies to elicit gestures for individual devices [37, 38, 39] and extend existing methods by considering multiple device types in order to provide guidelines for ensuring consistency and transferability.

As it is becoming more common to use applications across various devices, we set out with an application scenario to design a consistent gesture set. Since RSVP has been proposed as a feasible technique for reading on small screens and bears potential for making reading more efficient, we address its lack of commonly used controls. Because reading activities take place in various situations and on various devices, we focus on the design of an application- rather than device-specific gesture set for controlling reading flow in RSVP interfaces with the goal of using such gestures seamlessly across device types.

Command	*Mean*	*SD*
Play/Resume	1.00	0.00
Pause	1.25	0.50
Stop	1.50	0.57
Speed up	2.50	0.57
Slow down	2.50	0.57
Rewind a sentence	3.50	0.57
Rewind several sentences	4.50	0.57
Forward a sentence	3.25	0.50
Forward several sentences	4.25	0.50
Mean	**2.69**	**0.49**

Table 1. List of commands to control the reading flow and their corresponding conceptual complexities as assessed by four independent researchers (1=simple, 5=complex).

GESTURE ELICITATION STUDY

To explore reading flow controls and to create a gesture set for each device type, we first conducted a gesture elicitation study. Therefore, we identified nine commands most vital to controlling the RSVP reading flow. Since commands possess different conceptual complexities, four independent researchers rated each command's complexity on a scale from 1 (simple) to 5 (complex), similarly to Wobbrock *et al.* [39]. Table 1 lists all nine commands together with their mean complexity ratings.

Method

We followed the approach originally proposed by Wobbrock *et al.* [39] by first portraying the *effect* of a gesture and subsequently asking participants to perform a gesture *cause*. We employed a within-subjects design, in which participants were asked to design a gesture set for each of the respective device type—phone, watch, and glasses. We applied a partial latin-square with three balanced sequences and two additional participants following the sequence: phone, watch, glasses and watch, glasses, phone. We took video recordings of the performed gestures. We also collected subjective feedback from participants on the goodness, ease of performance, and social acceptability of each gesture they proposed.

Participants

We recruited 20 participants (10 female) with an average age of 28.15 ($SD = 5.2$) years using university mailing lists and social networks. Ten wore glasses, and nine had indicated previous experience with RSVP. 18 participants owned a smartphone, one owned a smart watch, no one had used smart glasses before. Six (30%) participants indicated to read occasionally and 14 (70%) indicated to read frequently. 19 (95%) regularly read printed text, 18 (90%) regularly read on a PC, 5 (25%) on tablets, 9 (45%) on smartphones, and 9 (45%) on e-readers.

Apparatus

To develop a multi-platform apparatus, we implemented web application based on the open-source code base provided by *OpenSpritz*[2]. We then used *Apache Cordova* [3] to create a corresponding Android apps optimized for each device type, namely the Samsung Galaxy S5 smartphone (running Android version 5.0), the Samsung Galaxy Gear SM-V700 smartwatch (running Android version 4.2.2), and Google Glass (running Android version 4.4.4). The resulting app collection shows text in the center of a display via RSVP. Words are displayed sequentially and centered around a red colored letter roughly after the first third of the word (see Figure 2). The colored spot marks the *optimal viewing point* and acts as focal point for the reader's eyes to rest on. The display algorithm takes into account word lengths and punctuation characters to determine the duration of presentation of each word. The duration is doubled for words with more than eight characters and words followed by a period, comma, colon, dash or open bracket. The open-source framework provided functions to start and stop the reading flow as well as set the reading speed in *words per minute* (wpm). We modified the source code to add the remaining control commands, such as jump back and forth in the text, pause, and resume. For each of the nine commands, we then prepared a short video using the software we built to show participants the effect and a textual explanation of the commands to be executed.

Procedure

After participants had signed a consent form we explained the purpose of the study and asked them to fill in an initial survey on demographics, general reading habits, and their technology usage. We then handed out the first device and explained its general input capabilities. To allow participants to familiarize themselves with RSVP, we handed them the respective device first displaying a short text without any control functions. When participants felt comfortable reading with it, we walked them through the list of nine commands (see Table 1). We then subsequently presented each video showing the respective command to design a gesture for. To make sure each command was fully understood, we showed participants the corresponding video, which included the written command name and clearly demonstrated the effect. Participants could watch the video multiple times and while the commands were generally easy to understand, the experimenter explicitly asked for confirmation before moving on. The sequence of videos, and therefore the sequence of the commands presented, was randomized. While we asked participants to think-aloud, we video-recorded the entire session, which we transcribed later to obtain qualitative data to analyze users' mental models. Participants rated each of their gestures on a 7-point Likert scale with regard to the following aspects: 1) "The gesture I picked is a good match for its intended purpose", 2) "The gesture I picked is easy to perform", and 3) "The gesture I picked can be carried out safely in public". Each scale ranged from 1 ('strongly agree') to 7 ('strongly disagree'). After having come up with nine distinct gestures for each command, we handed participants the next device. This was done until a gesture was allocated for each command on all three device types. When handing out each device, we explained its particular input capabilities but pointed out that participants should not feel constrained by them. The study took about an hour, for which participants were compensated with 10 EUR.

[2]https://github.com/Miserlou/Glance-Bookmarklet

[3]https://cordova.apache.org

Taxonomy of gestures for reading control		
Nature	Symbolic	Gesture visually depicts a symbol
	Metaphorical	Gesture indicates a metaphor
	Physical	Gesture acts physically on object
	Abstract	Gesture mapping is arbitrary
Flow	Discrete	Action occurs after gesture
	Continuous	Action occurs during gesture
Dimension	Single-Axis	Motion occurs around a single axis
	Tri-Axis	Motion involves either translational or rotational motion, not both
	Six-Axis	Motion occurs around both rotational and translational axes
Complexity	Simple	Gesture consists of a single gesture
	Compound	Gesture can be decomposed into simple gestures
Interaction	One-Finger	Gesture can be performed with one finger
	Multi-Finger	Gesture can be performed with two or more fingers
	Without Touch	Gesture can be performed without touching the input field
Location	Dependent	Action is dependent on the location of gesture
	Independent	Action is location independent

Table 2. Taxonomy of gestures for reading control based on collected gestures.

Results

With 20 participants, nine commands, and three device types, we collected a total of $20*9*3 = 540$ gestures. Our results include a gesture taxonomy, a user-defined gesture set for controlling RSVP of text for each of the three device types, subjective ratings for each set of gestures, and qualitative feedback. Finally, we report the results of our feedback assessment as participants were asked for appropriate ways to confirm successful command execution.

In contrast to Wobbrock *et al.* [39], designing gestures for reading control is not a device-specific task. The goal is rather the elicitation of a set of application-specific gestures that can be applied across device platforms with a focus on their transferability, *i.e.*, how easily they can be transferred from one device to another. To understand the user-defined gestures for the reading control, we first constructed a taxonomy of these gestures.

Taxonomy of Reading Control Gestures

Participants created 44 unique gestures for the phone, 50 for the watch, and 43 for the glasses. We first classified each of these gestures along six dimensions: *nature*, *flow*, *dimension*, *complexity*, *interaction*, and *location*. Each dimension has multiple categories as listed in Table 2. We adopted the dimensions from Wobbrock *et al.* [39] and Ruiz *et al.* [27] for comparability and further extended the taxonomy by the two dimensions *interaction* and *location*.

Taxonomy Breakdown				
		Phone	**Watch**	**Glasses**
Nature	Symbolic	0.00	0.00	0.00
	Metaphorical	0.55	0.46	0.37
	Physical	0.02	0.04	0.00
	Abstract	0.43	0.50	0.63
Flow	Discrete	0.70	0.76	0.77
	Continuous	0.30	0.24	0.23
Dimension	Single-Axis	0.86	0.66	0.65
	Tri-Axis	0.11	0.24	0.33
	Six-Axis	0.02	0.10	0.02
Complexity	Simple	0.66	0.72	0.65
	Compound	0.34	0.28	0.25
Interaction	One-Finger	0.61	0.56	0.56
	Multi-Finger	0.27	0.16	0.19
	Without Touch	0.11	0.28	0.26
Location	Dependent	0.20	0.18	0.09
	Independent	0.80	0.82	0.90

Table 3. Gestures and their ratios for each taxonomy category.

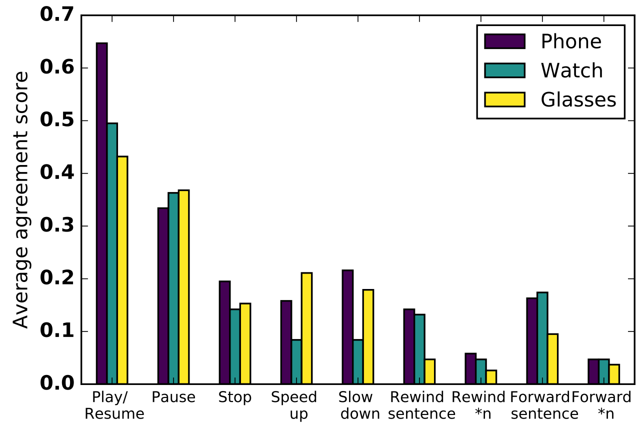

Figure 3. Agreement scores for each command on phone, watch and glasses (*n indicates several sentences).

The *nature* dimension comprises *symbolic* gestures which represent symbols. An example of this kind of gestures is drawing a letter "P" on the phone's display to pause the reading flow. *Metaphorical* gestures have figurative characteristics. For example, a *metaphorical* gesture can be looking at the watch in order to start the reading flow or a shake of the user's head to stop the reading flow on glasses. In the *physical* category, gestures are applied directly to the content, in our case, to the text. For example, a user defined a gesture where she flicked left over a text on the phone in order to jump to the beginning of the current sentence. Finally, gestures that did not fit into any of the previously described categories were defined as *abstract*.

The *flow* dimension describes whether the effect occurs during or after the gesture is performed. An example for the former one is a single tap gesture to stop a reading flow. In the latter category, an effect occurs while the gesture is being performed and finishes as soon as the user stops acting it. For example, a "tap and hold" gesture on the left part of the phone screen in order to jump back several sentences.

Dimension categorizes the user-defined gestures based on the amount of axis needed to perform them. Participants performed most of the gestures on the input field of a device, and therefore only a single axis was needed. An example of a *tri-axis* gesture includes tilting the phone to the right side in order to accelerate the reading flow. This gesture demands a rotational motion. Moving the watch-wearing arm down to the relaxing position to stop the reading flow is an example of a *six-axis* gesture.

Complexity describes if an effect is caused by a single or a composition of two or more gestures. An example of a single gesture would be the single tap for pausing, while a complex gesture would be multiple finger flicks to the right in order to skip several sentences.

Interaction classifies a gesture based on the number of fingers needed to perform it. A single finger tap to start the reading flow *vs.* a 2-finger flick to the front of the glasses in order to skip a sentence are examples for *one-finger* and *multi-finger* gestures. The *without touch* category contains all gestures that do not require physical contact with the input field of the device. For example, a user tilts the phone to the right in order to speed up the reading flow.

Location categorizes gestures based on their dependence on the location they are acted on. During the elicitation, study participants described gestures that differently acted when performed in different locations of the input field. For example, a user taps on the front side of the touchpad of the glasses in order to skip a sentence and on the back side to jump to the beginning of the current sentence.

Table 3 shows the percentage of gestures proposed within each taxonomy category for all device types. Participants did not suggest any *symbolic* gestures. We found significant positive correlations among taxonomy categories of gestures for each device type ($r = -.946$, $n = 16$, $p < .01$ between gestures for phone and watch, $r = -.911$, $n = 16$, $p < .01$ between gestures for phone and glasses, and $r = -.975$, $n = 16$, $p < .01$ between gestures for watch and glasses). The correlations show a consistence among proposed gestures for each device type, and thus the potential for constructing a transferable gesture set for all.

User-defined Gesture Sets

We collected a total of 180 gestures per device type, which we used to generate one respective gesture set. We, therefore, grouped identical gestures for each command. The group with the largest number of gestures was then chosen to be the representative gesture for the corresponding command, resulting in one *consensus set* for each of the three device types. Each numeric agreement score reflects the degree of consensus among participants regarding a corresponding control gesture. To evaluate the degree of consensus among participants, we computed the agreement score, as proposed by Vatavu and Wobbrock [37], for each command on each device type as follows:

$$A_c = \frac{|P_c|}{|P_c|-1} \sum_{P_i \subseteq P_c} \left(\frac{|P_i|}{|P_c|}\right)^2 - \frac{1}{|P_c|-1} \quad (1)$$

Figure 4. Our RSVP implementation provides Visual feedback for *position*, *reading speed*, *stop* and *pause* commands. During pauses, a progress bar is shown displaying the percentage of text read and the current reading speed. Visual cues disappear after one second.

We determined an agreement score A_c per command using the equation 1 where P_c is the set of user-defined gestures for command c in the set of all commands C, and P_i is the subset of identical gestures for that command. As an example, consider the calculation of an agreement score for *play/resume* using a phone. The command has three groups of identical gestures with a size of 16, 3, and 1. Thus, the agreement score for *play/resume* using a phone is calculated as follows:

$$A = \frac{20}{19}\left(\left(\frac{16}{20}\right)^2 + \left(\frac{3}{20}\right)^2 + \left(\frac{1}{20}\right)^2\right) - \frac{1}{19} = 0.647 \quad (2)$$

Figure 3 summarizes the agreement scores per command and device type. Agreement scores from our gesture set are similar to those from other elicitation studies [39, 32, 27]. Participants had the least agreement on gestures for *skip several sentences* and *jump back several sentences*. This can be explained by the complexity of these commands. Thus, both commands have higher conceptual complexity than the rest of the commands, namely 4.5 and 4.25. Agreement scores of our gesture sets for each device are inversely correlated with their conceptual complexities ($r = -.805$, $n = 9$, $p < .01$ for phone, $r = -.763$, $n = 9$, $p < .01$ for watch, and $r = -.892$, $n = 9$, $p < .01$ for glasses), *i.e.*, the more complex the gesture, the lower the agreement on a single gesture by participants.

Consensus Set

After classifying the gestures, we defined the *consensus set* for controlling the RSVP reading flow. We, therefore, grouped gestures for each command proposed for each device type and selected the most common gesture as a representative of the command. Solely using this method for developing a consensus set, however, worked only for the gesture set for the watch, but caused conflicts between competing gestures on the other two device types, where the same gesture was selected for issuing another command, for example. To resolve this conflict, we took into account the consistency of interactions: therefore, we made sure that all gestures were **unique** for each device type, **reversible**, *e.g. speed up* and *slow down*, and repeated gestures were **consistent**, *e.g. jump to beginning of current sentence* and *jump back several sentences*. If the same gesture was the most frequently selected one for more than one command, we assigned the gesture to the command with the higher agreement score and took the second most common gesture for the other command. If inconsistent, we selected the next most frequently submitted gesture that was compatible with other related gestures. Figure 5 depicts the resulting gesture set for each device type. The overall gesture set is without conflicts and covers 77.77% of gestures that

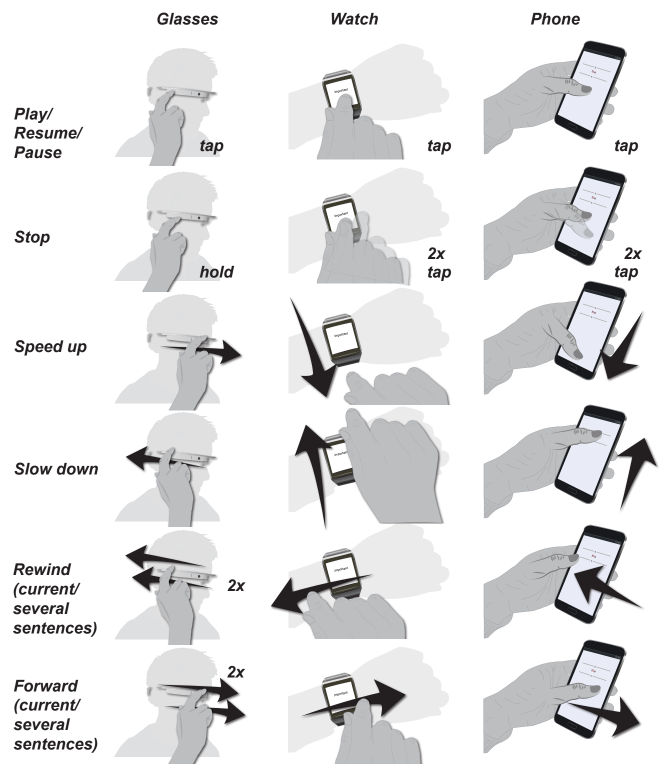

Figure 5. Consensus gesture set for RSVP interaction across all three device types: phone, watch, and glasses.

were most frequently submitted. In the study, participants also assessed fitness, ease and social acceptance of the gestures they proposed. Comparing the gestures from the selected consensus set and the rest of the suggested gestures, we did not find significant differences in fitness, ease and social acceptance between them. Table 4 contains the mean scores for gesture ease and social acceptability of the consensus set.

Eliciting System Feedback

System feedback is vital to signal users that a command has been successfully issued [34]. A system can provide feedback to a user in different modalities, such as visually, through audio or vibrations. Therefore, we asked participants after the elicitation study, what kind of feedback they would prefer for reading control verification. All participants were familiar with possible feedback modalities for each device type. They could select several modalities for the control verification. 16 (80%) participant preferred visual feedback while audio and haptic feedback was chosen by five (25%) and nine (45%) participants respectively. However, one participant reported that feedback is not essential for the controlling the reading flow on these devices and did not select any modality. Participants mostly preferred haptic feedback on watches because of the wrist's sensitivity. They further reported that audio feedback might be a disturbing factor during reading as well as inappropriate in public spaces. 62.5% of participants preferred an iconic visualization of applying commands.

During reading with RSVP, users see text word by word, hence they are not explicitly aware of the current text position. We, therefore, asked participants what kind of information they would prefer in order to know the current text position, and at which point this information should be displayed on the device. We provided them a list of text position cues, including: *percentage of text read*, *remaining number of words*, *remaining time to finish the text at current reading speed*, and *a visual progress bar*. The *visual progress bar* was chosen 17 times while *remaining time at current reading speed*, *a remaining number of words* and *percentage read* were selected ten, two, and nine times, respectively. After choosing a cue, participants decided whether it should be displayed permanently or only when the reading flow was paused. Participants preferred all cues except *remaining number of words* to be displayed while reading was paused.

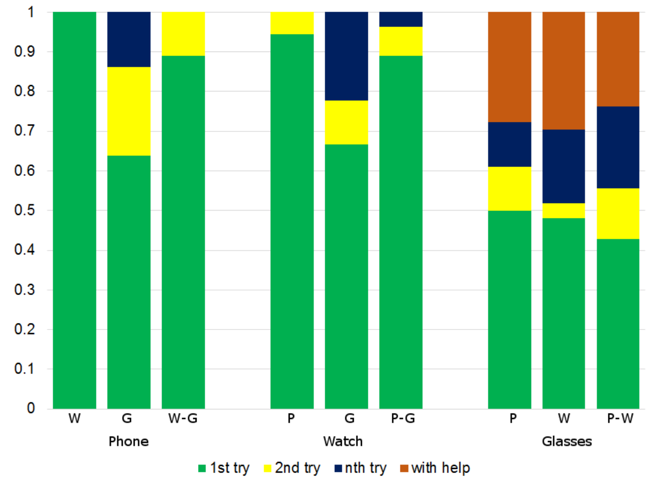

Figure 6. Objective assessment of transferability of gestures between the three device types. The diagram shows 1) the transferability of gestures to the phone (left) after using them on the watch (W), glasses (G) or on both of them (W-G), 2) the transferability of gestures to the watch (middle) after using them on the phone (P), glasses (G) or on both of them (P-G), and 3) the transferability of gestures to glasses (right) after using them on the phone (P), watch (W) or on both of them (P-W).

VALIDATION STUDY

To verify the applicability of the resulting user-defined gesture sets, we implemented the corresponding sets of controls for each device type and conducted a second user study to validate them. We were especially interested in an assessment of each set with regard to learnability and consistency in mapping, *i.e.* how transferable a gesture set is from one device to another.

Method

To assess the transferability of a gesture set from a particular device to the remaining ones, we assigned the variable *device type* between-subjects. Hence, we had three conditions, namely: phone, watch, and glasses. Six participants were recruited for each condition resulting in a total of 18 study participants. We assessed the ease of transferring each gesture set by recording the number of trials needed to find the gesture corresponding to a command on the remaining devices. Before being asked to transfer the command set, participants learned and practiced each command on the primary device. We also recorded subjective feedback on how well each gesture fits its corresponding command.

Apparatus

For this study apparatus, we extended our existing prototype by implementing all nine control gestures for each device type (see Figure 5). During pilot tests, we found $20wpm$ to be a feasible speed change for each *speed up / slow down* gesture. Jumping back to the beginning of the sentence and jumping back several sentences resulted in the same gesture being applied several times within a 1-second timeout. We used the information collected in the previous study to provide feedback for each gesture: since participants mostly preferred visuals to communicate command execution, we added iconic visual feedback whenever a command has been recognized and applied by the device. To give an idea about the current text position and reading speed, the system display contains a progress bar with current reading speed while on pause (see Figure 4).

Participants

We recruited 18 participants (5 females) through university mailing lists and social networks, who had not taken part in the previous elicitation study. Their average age was 28.66 ($SD = 10.12$) and their background was mainly academic. Seven participants wore glasses, three were already familiar with RSVP, 17 owned a smartphone, one a smart watch, and none were using smart glasses.

Procedure

After signing the consent form, participants filled in a short survey comprising questions about demography and technology usage. We then handed them one device type according to their experimental group assignment. Participants received a paper sheet that listed each command and the corresponding gesture in text form for the device used. After participants had enough time to familiarize themselves with the command set, we asked them to read a practice text and try out each command until they felt comfortable enough to continue. The initial reading speed was set to $180wpm$. To become familiar with each command, we had a list of instructions prepared for the first part of the study: first, we told participants to speed up to $300wpm$. Once they confirmed having succeeded doing so, we asked them to slow down to $200wpm$. The following tasks included re-reading of passages, finding specific paragraphs in the text, and skimming the text until the end, thereby making sure that all commands needed to be executed. After this familiarization part, we gave participants another text instructing them to read it in full detail while using the controls as they pleased. To make sure participants took this reading exercise seriously, we announced a pending text comprehension test.

		Play/Resume/Pause	Stop	Speed Control	Flow Control
Phone Gestures	**Fit**	6.75 (+/-0.4)	6.5 (+/-0.8)	6.4 (+/-0.7)	6 (+/-1)
	Ease	6.9 (+/-0.4)	6.7 (+/-0.8)	6.7 (+/-0.6)	6.3 (+/- 1.2)
	Social Acceptance	6.9 (+/-0.3)	6.8 (+/- 0.4)	7 (+/-0)	6.8 (+/-0.4)
	Transferable to Watch	6.9 (+/-0.2)	6.7 (+/-0.5)	6.6 (+/- 0.4)	6.6 (+/-0.5)
	Transferable to Glasses	6.3 (+/-1)	2.5 (+/-1.4)	5.2 (+/-1.4)	2.9 (+/-1.5)
Watch Gestures	**Fit**	6.6. (+/-0.4)	6 (+/-0.9)	6.7 (+/-0.4)	6.4 (+/-0.8)
	Ease	6.8 (+/-0.5)	6.4 (+/-0.9)	6.9 (+/-0.4)	5.9 (+/-1.9)
	Social Acceptance	6.8 (+/-0.5)	6.4 (+/-0.9)	6.9 (+/-0.4)	6.5 (+/-0.9)
	Transferable to Phone	6.9 (+/-0.2)	6.8 (+/-0.4)	6.8 (+/-0.4)	6 (+/-2)
	Transferable to Glasses	6.7 (+/-0.5)	4.3 (+/-1.9)	3.9 (+/-1.1)	3.4 (+/-2.3)
Glass Gestures	**Fit**	6.7 (+/-0.4)	5.7 (+/-1.4)	5.9 (+/-0.7)	5 (+/-1)
	Ease	6.8 (+/-0.5)	7 (+/-0)	6.7 (+/-0.5)	6.5 (+/-0.6)
	Social Acceptance	6.6 (+/-0.6)	6.3 (+/-1)	6.8 (+/-0.4)	7 (+/-0)
	Transferable to Phone	6.7 (+/-0.5)	5.7 (+/-0.8)	4.8 (+/-1.8)	5 (+/-1.2)
	Transferable to Watch	6.8 (+/-0.4)	5.7 (+/-0.8)	4.7 (+/-1.7)	5 (+/-1.2)

Table 4. Study participants' subjective assessments of gesture fit and transferability to other devices collected during the validation study. Gesture ease and social acceptance scores represent the ratings of the consensus set collected during the elicitation study. Values depict the means (SDs) of 7-point Likert-style ratings from 1 (*strongly disagree*) to 7 (*strongly agree*).

Meanwhile, we recorded the commands that were executed during reading. After the reading task, we handed participants one of the other two device types. For each of the nine commands, we asked them to find the corresponding gesture on the new device. Thereby, we measured participants' ability to transfer the gesture set learned on one device to the next by counting the trials until they found the correct gesture. The sequence of the commands to be executed was randomized. After transferring the gesture set to the second device type, we handed participants the third one with the same instructions to perform all commands. The sequence of devices to transfer the command sets was counterbalanced. In the end, participants filled in a final questionnaire where each command was listed again with the corresponding gesture on the device they had handled first. For each gesture, we asked participants to answer the following questions on a 7-point Likert-scale: 1) "How well do you think this gesture fits the control?", 2) "How well was the gesture transferable to the second device type?", and 3) "How well was the gesture transferable to the third device type?". The study took about 40 minutes per session, for which each participant was rewarded with 10 EUR.

Results

In the following, we report participants' ability to transfer a gesture set, which was learned on one device, to subsequent ones. We did not analyze the comprehension scores since the tests' sole purpose was to make sure participants took the reading exercise seriously.

Transferability of Gestures

To ensure gesture consistency across device boundaries we were interested in validating the transferability of these gestures, *i.e.*, how well a set of commands once learned on one device can be ported to another. Therefore, we made sure participants were familiar reading on one device, before handing them the other two devices without further instructions to see how they performed in figuring out the respective reading flow controls. In random order, the experimenter requested each of the nine reading controls to be performed and recorded whether the corresponding gesture was correctly found and applied on the first, second try or in more than 2 trials (see Figure 6). Because gestures for commands that have a logical reverse action (*e.g.*, speed up, slow down) were designed with consistency in mind, finding the reverse gesture is an easy task once a command has been found. Hence, we classified the gestures into four groups: 1) *Play/Resume/Pause*, 2) *Stop*, 3) *Speed Control* (speed up, slow down), and 4) *Flow Control* (skipping sentences forward or backward). Because gestures for *play*, *resume* and *pause* commands were identical, we grouped them together. *Stop* was used as a quit function, while speed controls contain the acceleration or deceleration of the reading flow and flow controls cover any commands causing a change in the current text position.

For each command, we recorded the number of trials participants needed to find the corresponding gesture on the secondary or tertiary device. For calculating a transferability score we define the following four bins:

- **1**: if a participant found the command on the first attempt, we assigned the category score 0.
- **2**: if a participant found the command after a second attempt, we assigned the category score 1.
- **N**: if a participant eventually found the command after more than 2 attempts, we assigned the category score 2.
- **With Help**: if a participant was unable to find the command without explicit help by the experimenter, we assigned the category score 3.

Because the total number of attempts is highly dependent on the individual participant's levels of frustration or inclination to meet the experiment goals, we discounted a limited number of attempts as 3. Hence, we came up with the four bins described, which at the same time create a weight for the following transferability score assessment:

$$T_{AB} = \frac{\sum_{p \subseteq P} \sum_{c \subseteq C} X_{c_p}}{N_P * N_C * w} \tag{3}$$

We define the **transferability score** T from one device A to another B by summing up all category scores $X \varepsilon \{0, 1, 2, 3\}$ for each command c out of the command set C given by each participant p divided by the number of participants N_p multiplied with the total number of commands N_c multiplied with the worst possible category score weight w. A perfectly transferable gesture set between two devices, therefore, results in a transferability score of 0, whereas the worst case (*i.e.*, every gesture needs to be explicitly taught) results in a score of 1.

The overall **consistency score** of the gesture set is defined as the mean of all transferability scores between all device types:

$$\overline{T} = \frac{\sum_{t \subseteq T} t}{N_T} \tag{4}$$

Hence, each transferability score i is summed up and divided by the total number of device transfers N_T. For our gesture set, we calculated the following transferability scores with the worst possible category score weight $w = 3$ and the number of device transfers $N_T = 6$. Table 5 contains the transferability scores for each gesture transfer. The easiest transfer took place from watch to phone ($T_{wp} = 0.0104$) and

	Phone	Watch	Glasses
Phone		0.0313	0.375
Watch	0.0104		0.4062
Glasses	0.1875	0.1354	

Table 5. Transferability scores with primary device type in the left and transfer devices in the subsequent columns. A perfect, *i.e.*, immediate command transfer results in a transferability score of 0, a worst case transfer (i.e., every gesture needs to be explicitly taught) results in a score of 1.

reverse ($T_{pw} = 0.0313$), whereas the transfer to glasses was less straight forward ($T_{pg} = 0.375$ and $T_{wg} = 0.4062$) for participants as indicated by the high transferability score. The overall consistency score of our final gesture set is calculated as $\overline{T} = 0.191$.

Subjective Assessment

During the validation study, we asked participants to assess how well gestures on the primary device fit the commands to be executed. We further collected subjective feedback on the ease of transferability to the remaining devices. Table 4 depicts the mean subjective assessment scores. Ratings were given on a 7-point Likert-style scale from 1 (*strongly disagree*) to 7 (*strongly agree*). Gestures on phone and watch were overwhelmingly rated above 6 ($M = 6.4$), less so on glasses ($M = 5.8$). Gesture transfer between phone and watch was perceived as predominantly easy ($M = 6.7$). Transfers from glasses to phone or watch less so ($M = 5.6$), but seem to be easier than a transfer from other devices to glasses ($M = 4.5$).

FRAMEWORK FOR CONTROL TRANSFERABILITY

When designing gesture sets, a user-centered design approach has been shown to produce easy to use, fitting, and socially acceptable gestures [27, 39]. For applications that are used across multiple devices, however, the design process for user-defined gesture elicitation can produce inconsistencies. Hence, our goal was to finalize a gesture set not solely on agreement scores, but also taking into account gesture consistency and ease of transferability across different device types. In the following, we formalize our process so it can be further validated and re-used in future research and designs. The resulting framework can thus be applied to create application-specific gesture sets with regard to consistency and transferability across device types. It consists of three phases:

1. **Elicitation**: eliciting an independent gesture set for each device type.
2. **Consolidation**: merging gestures into a proposed set based on agreement scores as well as consistency constraints.
3. **Validation**: assessing the ease of porting a learned gesture set to the remaining devices using a transferability score between devices and a consistency score for describing the goodness of the overall gesture set across devices.

During the **elicitation** phase, we follow the processes widely applied for deriving user-defined gesture sets [38, 39] by conducting elicitation studies for each device, during which we first portray the effect of a gesture and subsequently ask study participants to perform the gesture's cause. Based on the agreement scores, this process leaves us with individual gesture sets optimized for a single device type.

In the **consolidation** phase, we take these individual scores and construct a unified gesture set that compromises between agreement scores, device constraints, and both command and device consistency. Command consistency ensures that gestures that have a natural counterpart (reverse actions) are assigned a consistent gesture group. Device consistency entails the similarity between a specific command being issued on one device and its counterpart on another device. The outcome of this second phase is a unified gesture set, which can now be implemented for each device type.

Once the gesture set is implemented on all respective devices, gesture consistency and ease of transfer is assessed in the final **validation**. Therefore, a second study follows an in-between subjects design, in which participants are grouped by the primary device type. The gesture set with its corresponding commands is then explicitly explained on the respective device, after which study participants transfer each gesture to the remaining devices. Records are kept on the number of attempts made to successfully execute the corresponding command on each device and scores are assigned depending on the severity of the transfer barrier. These scores are used to calculate the *transferability score* from one device to another (see Equation 3) as a quantification of the ease of transfer. An overall *consistency score* of the unified gesture set is then calculated across all transferability scores based on every possible transfer direction between the devices of interest (see Equation 4). The resulting consistency score is a metric between 0 (perfect transfer) and 1 (transfer with instructions) for the overall goodness of the gesture set across devices with regard to its consistency and transferability and eventually allows comparison between different gesture set candidates.

DISCUSSION

In the following, we discuss the process of eliciting an application-specific gesture set for RSVP reading and the implications of our proposed method to the design process with regard to consistency and transferability of gesture sets.

RSVP Gesture Set

Reading with RSVP has been the subject of several systems and investigations [9, 11], but its alienating effect often stems from a perceived lack of user control [4]. With the study described in this paper, we address this challenge and present a user-defined gesture set that is applicable to most ubiquitous reading devices. In contrast to other guessability studies [27, 39], we implemented the resulting gesture set, which allowed us to validate our design. During the elicitation process, we observed certain user biases depending on well-established gesture types and device input capabilities. First of all, most of our participants had extensive experiences with ubiquitous computing devices due to being exposed to them for several years now. Hence, the tab or swipe gesture, for example, is quite common and triggers certain expectations towards the interaction. Second, the form factors of the devices themselves influence the user-driven design process: in our study, devices similar in nature, such as phone and watch, lead users to gesture designs that were quite similar as well. Devices with different form factors, however, lead to gestures quite different from the one designed for another device, such as was the case for the rather unfamiliar glasses. Instead of reducing such biases and expectations as proposed by Morris *et al.* [22] we can take them into consideration when unifying a gesture set. With the two metrics described in this paper—transferability and consistency score—we formalized a way to quantify the gap between user expectation and command function.

Gesture Transferability

Following the traditional model of elicitation studies, we reached its limitations at the point where we needed to compromise between choosing gestures based on agreement scores and designing for consistent interactions across device boundaries. Because some of the choices we made in the consolidation phase, as can be argued, were rather designer- than user-driven, there was a need for a final user-centered validation. In two out of the three steps of our proposed process—elicitation and validation—the user, therefore, remains to be at the center of the design process as it has been found beneficial when designing input systems [10, 30]. Transferring gestures from phone to watch resulted in more consistent interactions due to their similar form factors, which suggests that other touch devices, such as tablets or touchpads, could be easily added. As the number of commands and devices increase, consistency becomes more important (to support learnability and memorability) while agreement scores for single device gesture sets are likely to be lower (due to an increasing amount of commands to be designed for). Our proposed method supports adding new device types or gesture sets post-hoc, where only the transferability study part needs to be revisited.

Limitations and Next Steps

When designing gesture sets for devices with different form factors, compromises need to be made between what is considered "intuitive" design choices and their transferability, which is limited by the input (or output) capabilities of the secondary devices. We made these compromises in what we called the "consolidation phase" and, as previously mentioned, they eventually are subject to the designer's conception. Therefore, there was a need for additional user validation. Ideally, we would create several gesture sets during that phase, which could be compared against each other during validation. We could also introduce iterative design cycles, in which insights from the validation (or an unfavorable consistency score) would lead to another round of elicitation or consolidation of gesture set options. Eventually, multiple iterations are likely to converge around a certain transferability score. In contrast to allowing users to create customized gesture sets, as proposed by Malloch *et al.* [18], our proposed method focuses on gesture consistency across devices, which eventually reduces users' memory load as the effort to memorize device-specific gesture sets is reduced. Future work could look into how customizable gesture sets could be supported at the time of creation while transferability is ensured or gesture equivalents for other devices are automatically derived.

Further, in order to validate the gesture set, we had to actually implement the unified gesture set. When transferability scores point towards a rather unintuitive gesture transfer, the design needs to be reconsidered and implementation revisited. Several validation cycles would, therefore, be limited by the costs of implementation. Hence, it is worth looking into whether validation would be sufficient through a Wizard-of-Oz approach, where command execution is triggered manually based on observation.

Finally, finding command equivalents on additional devices should get subsequently easier. This learning effect stems from the transitivity of gestures that have been correctly identified on previous devices. We tried to accommodate this effect by counterbalancing the sequence, in which we handed out the devices in our study. However, for devices with similar form factors, this effect can get disproportional and will require adjustments via an additional weight in the transferability score equation, for example. Assigning the weights in the validation part played a crucial role in discounting the number of participants' attempts to successfully execute a respective command. In future work, metrics should be validated through further cross-device control elicitations, but also through extending their applicability by including devices with inherently different form factors and weighing schemes. The proposed transferability and consistency scores allow us to benchmark cross-device gesture sets and will give other designers and researchers a process to quantify and compare the goodness of transferable gestures.

CONCLUSION

Each type of device has its own means for input and output, which poses challenges to the consistency of interactions. Many applications are expected to work across different device types, which is why interactions need to be increasingly designed so that they are consistent and therefore easily transferable. In this paper, we demonstrate the creation of a unified gesture set that follows a user-defined elicitation approach and takes into account the transferability of gesture sets across different device types. By eliciting gesture sets for controlling reading via RSVP on wearable devices, we present the resulting gesture set for phone, watch, and glasses along with a method to design for and assess consistency. While the proposed transferability score describes the ease of transferring a gesture set from one device to another, the consistency score is a metric for a the overall goodness of a cross-device gesture set. By applying these metrics and following the three steps of 1) elicitation, 2) consolidation, and 3) validation, application designers are given a method to design, evaluate, and compare user-defined gesture sets that are consistent and easily transferable across devices with different form factors.

ACKNOWLEDGMENTS

We acknowledge the financial support by JST CREST under Grant No: JPMJCR16E1, by the graduate program *Digital Media* of the Universities of Stuttgart and Tübingen, and the Stuttgart Media University, as well as by the German Research Foundation (DFG) through projects C02 and C04 of the SFB/Transregio 161.

REFERENCES

1. David Akers. 2006. Wizard of Oz for Participatory Design: Inventing a Gestural Interface for 3D Selection of Neural Pathway Estimates. In *CHI '06 Extended Abstracts on Human Factors in Computing Systems (CHI EA '06)*. ACM, New York, NY, USA, 454–459. DOI: http://dx.doi.org/10.1145/1125451.1125552
2. Lisa Anthony, Radu-Daniel Vatavu, and Jacob O. Wobbrock. 2013. Understanding the Consistency of Users' Pen and Finger Stroke Gesture Articulation. In *Proceedings of Graphics Interface 2013 (GI '13)*. Canadian Information Processing Society, Toronto, Ont., Canada, Canada, 87–94. http://dl.acm.org/citation.cfm?id=2532129.2532145
3. Richard A. Bolt. 1980. "Put-that-there": Voice and Gesture at the Graphics Interface. In *Proceedings of the 7th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '80)*. ACM, New York, NY, USA, 262–270. DOI: http://dx.doi.org/10.1145/800250.807503
4. Tilman Dingler, Alireza Sahami Shirazi, Kai Kunze, and Albrecht Schmidt. 2015. Assessment of Stimuli for Supporting Speed Reading on Electronic Devices. In *Proceedings of the 6th Augmented Human International Conference (AH '15)*. ACM, New York, NY, USA, 117–124. DOI: http://dx.doi.org/10.1145/2735711.2735796
5. Tanja Döring, Dagmar Kern, Paul Marshall, Max Pfeiffer, Johannes Schöning, Volker Gruhn, and Albrecht Schmidt. 2011. Gestural Interaction on the Steering Wheel: Reducing the Visual Demand. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, 483–492. DOI:http://dx.doi.org/10.1145/1978942.1979010
6. Robert L Duchnicky and Paul A Kolers. 1983. Readability of text scrolled on visual display terminals as a function of window size. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 25, 6 (1983), 683–692. DOI: http://dx.doi.org/10.1177/001872088302500605
7. David Efron. 1941. Gesture and environment. (1941).
8. Kenneth I Forster. 1970. Visual perception of rapidly presented word sequences of varying complexity. 8, 4 (01 Jul 1970), 215–221. DOI: http://dx.doi.org/10.3758/BF03210208
9. Tsvetozar Georgiev. 2012. Investigation of the User's Text Reading Speed on Mobile Devices. In *Proceedings of the 13th International Conference on Computer Systems and Technologies (CompSysTech '12)*. ACM, New York, NY, USA, 329–336. DOI: http://dx.doi.org/10.1145/2383276.2383324
10. Michael D. Good, John A. Whiteside, Dennis R. Wixon, and Sandra J. Jones. 1984. Building a User-derived Interface. *Commun. ACM* 27, 10 (Oct. 1984), 1032–1043. DOI:http://dx.doi.org/10.1145/358274.358284
11. Björn Hedin and Erik Lindgren. 2007. A Comparison of Presentation Methods for Reading on Mobile Phones. *Distributed Systems Online, IEEE* 8, 6 (June 2007), 2–2. DOI:http://dx.doi.org/10.1109/MDSO.2007.34
12. Niels Henze, Andreas Löcken, Susanne Boll, Tobias Hesselmann, and Martin Pielot. 2010. Free-hand Gestures for Music Playback: Deriving Gestures with a User-centred Process. In *Proceedings of the 9th International Conference on Mobile and Ubiquitous Multimedia (MUM '10)*. ACM, New York, NY, USA, Article 16, 10 pages. DOI: http://dx.doi.org/10.1145/1899475.1899491
13. Frank G. Hofmann, Peter Heyer, and Günter Hommel. 1998. *Velocity profile based recognition of dynamic gestures with discrete Hidden Markov Models*. Springer Berlin Heidelberg, Berlin, Heidelberg, 81–95. DOI: http://dx.doi.org/10.1007/BFb0052991
14. Shaun K. Kane, Jacob O. Wobbrock, and Richard E. Ladner. 2011. Usable Gestures for Blind People: Understanding Preference and Performance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, 413–422. DOI: http://dx.doi.org/10.1145/1978942.1979001
15. Christian Kray, Daniel Nesbitt, John Dawson, and Michael Rohs. 2010. User-defined Gestures for Connecting Mobile Phones, Public Displays, and Tabletops. In *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '10)*. ACM, New York, NY, USA, 239–248. DOI: http://dx.doi.org/10.1145/1851600.1851640
16. Christine Kühnel, Tilo Westermann, Fabian Hemmert, Sven Kratz, Alexander Müller, and Sebastian Möller. 2011. I'm home: Defining and evaluating a gesture set for smart-home control. *International Journal of Human-Computer Studies* 69, 11 (2011), 693 – 704. DOI: http://dx.doi.org/10.1016/j.ijhcs.2011.04.005
17. Byron Lahey, Audrey Girouard, Winslow Burleson, and Roel Vertegaal. 2011. PaperPhone: Understanding the Use of Bend Gestures in Mobile Devices with Flexible Electronic Paper Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, 1303–1312. DOI: http://dx.doi.org/10.1145/1978942.1979136
18. Joseph Malloch, Carla F. Griggio, Joanna McGrenere, and Wendy E. Mackay. 2017. Fieldward and Pathward: Dynamic Guides for Defining Your Own Gestures. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 4266–4277. DOI: http://dx.doi.org/10.1145/3025453.3025764

19. Catherine C. Marshall and Christine Ruotolo. 2002. Reading-in-the-small: A Study of Reading on Small Form Factor Devices. In *Proceedings of the 2Nd ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL '02)*. ACM, New York, NY, USA, 56–64. DOI: http://dx.doi.org/10.1145/544220.544230

20. Michael E. J. Masson. 1983. Conceptual processing of text during skimming and rapid sequential reading. *Memory & Cognition* 11, 3 (01 May 1983), 262–274. DOI:http://dx.doi.org/10.3758/BF03196973

21. David McNeill. 1992. *Hand and mind: What gestures reveal about thought*. University of Chicago Press.

22. Meredith Ringel Morris, Andreea Danielescu, Steven Drucker, Danyel Fisher, Bongshin Lee, m. c. schraefel, and Jacob O. Wobbrock. 2014. Reducing Legacy Bias in Gesture Elicitation Studies. *interactions* 21, 3 (May 2014), 40–45. DOI:http://dx.doi.org/10.1145/2591689

23. Michael Nielsen, Moritz Störring, Thomas B. Moeslund, and Erik Granum. 2004. *A Procedure for Developing Intuitive and Ergonomic Gesture Interfaces for HCI*. Springer Berlin Heidelberg, Berlin, Heidelberg, 409–420. DOI:http://dx.doi.org/10.1007/978-3-540-24598-8_38

24. Isabella Poggi. 2002. *From a Typology of Gestures to a Procedure for Gesture Production*. Springer Berlin Heidelberg, Berlin, Heidelberg, 158–168. DOI: http://dx.doi.org/10.1007/3-540-47873-6_16

25. Mary C Potter and Ellen I Levy. 1969. Recognition memory for a rapid sequence of pictures. *Journal of experimental psychology* 81, 1 (1969), 10.

26. Jane E Raymond, Kimron L Shapiro, and Karen M Arnell. 1992. Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human perception and performance* 18, 3 (1992), 849.

27. Jaime Ruiz, Yang Li, and Edward Lank. 2011. User-defined Motion Gestures for Mobile Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, 197–206. DOI: http://dx.doi.org/10.1145/1978942.1978971

28. Thomas Schlömer, Benjamin Poppinga, Niels Henze, and Susanne Boll. 2008. Gesture Recognition with a Wii Controller. In *Proceedings of the 2Nd International Conference on Tangible and Embedded Interaction (TEI '08)*. ACM, New York, NY, USA, 11–14. DOI: http://dx.doi.org/10.1145/1347390.1347395

29. Elizabeth R Schotter, Randy Tran, and Keith Rayner. 2014. Don't Believe What You Read (Only Once) Comprehension Is Supported by Regressions During Reading. *Psychological science* (2014), 0956797614531148. DOI: http://dx.doi.org/10.1177/0956797614531148

30. Douglas Schuler and Aki Namioka. 1993. *Participatory design: Principles and practices*. CRC Press.

31. Marcos Serrano, Barrett M. Ens, and Pourang P. Irani. 2014. Exploring the Use of Hand-to-face Input for Interacting with Head-worn Displays. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 3181–3190. DOI: http://dx.doi.org/10.1145/2556288.2556984

32. Shaikh Shawon Arefin Shimon, Courtney Lutton, Zichun Xu, Sarah Morrison-Smith, Christina Boucher, and Jaime Ruiz. 2016. Exploring Non-touchscreen Gestures for Smartwatches. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 3822–3833. DOI: http://dx.doi.org/10.1145/2858036.2858385

33. Ben Shneiderman. 1987. *User interface design and evaluation for an electronic encyclopedia*. University of Maryland.

34. Ben Shneiderman and Catherine Plaisant. 2004. *Designing the User Interface: Strategies for Effective Human-Computer Interaction (4th Edition)*. Pearson Addison Wesley.

35. Robert Spence. 2002. Rapid, serial and visual: a presentation technique with potential. *Information visualization* 1, 1 (2002), 13–19. DOI: http://dx.doi.org/10.1057/palgrave.ivs.9500008

36. 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 (March 2013), 187–211. http://dl.acm.org/citation.cfm?id=2594684.2594688

37. Radu-Daniel Vatavu and Jacob O. Wobbrock. 2015. Formalizing Agreement Analysis for Elicitation Studies: New Measures, Significance Test, and Toolkit. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 1325–1334. DOI: http://dx.doi.org/10.1145/2702123.2702223

38. Jacob O. Wobbrock, Htet Htet Aung, Brandon Rothrock, and Brad A. Myers. 2005. Maximizing the Guessability of Symbolic Input. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems (CHI EA '05)*. ACM, New York, NY, USA, 1869–1872. DOI: http://dx.doi.org/10.1145/1056808.1057043

39. Jacob O. Wobbrock, Meredith Ringel Morris, and Andrew D. Wilson. 2009. User-defined Gestures for Surface Computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. ACM, New York, NY, USA, 1083–1092. DOI: http://dx.doi.org/10.1145/1518701.1518866