Investigating the Effects of Legacy Bias: User Elicited Gestures from the End Users Perspective

Ceylan Beşevli
Koç University-Arçelik Research Center for Creative Industries (KUAR)
Istanbul/Turkey
cbesevli17@ku.edu.tr

Merve Erkaya
Koç University-Arçelik Research Center for Creative Industries (KUAR)
Istanbul/Turkey
meerkaya@ku.edu.tr

Oğuz Turan Buruk
Koç University-Arçelik Research Center for Creative Industries (KUAR)
Istanbul/Turkey
oburuk@ku.edu.tr
& Gamification Group, Laboratory of Pervasive Computing,
Tampere University of Technology,
Tampere/Finland

Oğuzhan Özcan
Koç University-Arçelik Research Center for Creative Industries (KUAR)
Istanbul/Turkey
oozcan@ku.edu.tr

DIS'18 Companion, *June 9–13, 2018, Hong Kong*

ACM ISBN 978-1-4503-5631-2/18/06.
https://doi.org/10.1145/3197391.3205449

Abstract
User elicitation studies are commonly used for designing gestures by putting the users in the designers' seat. One of the most encountered phenomenon during these studies is legacy bias. It refers to users' tendency to transfer gestures from the existing technologies to their designs. The literature presents varying views on the topic; some studies asserted that legacy bias should be diminished, whereas other stated that it should be preserved. Yet, to the best of our knowledge, none of the elicitation studies tested their designs with the end users. In our study, 36 participants compared two gesture sets with and without legacy. Initial findings showed that legacy gesture set had higher scores. However, the interviews uncovered that some non-legacy gestures were also favored due to their practicality and affordances. We contribute to the legacy bias literature by providing new insights from the end users' perspective.

Author Keywords
User Elicitation; Gesture Control; Legacy bias; Embodied Interaction

ACM Classification Keywords
H.5.2. Empirical studies in Interaction design; User centered design

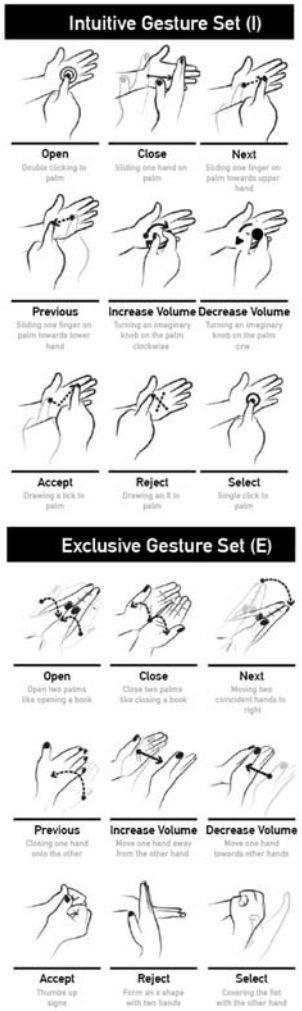

Figure 1: Intuitive (I) and Exclusive (E) gesture sets.

Introduction

Gesture elicitation studies are done by extracting designs from the users. These studies emerged because it was stated that system designers mostly bore technical concerns in their designs, more than meeting the users' needs [8]. Therefore, as the users designed, their elicited gestures were found easier to perform and more comprehensible [5]. Yet, these 'new' designs were influenced by the gestures used with the current technology (mouse, touch surface). This concept is called *legacy bias* [5]. There are diverging views and speculative studies on the topic that demand inquiry. Also, there's a significant lack of exploration on the effects of legacy bias on end users' preferences.

The conflicting views are as follows. First, it was claimed that legacy bias hinders the exploration of the novelty of innovative systems [4]. The other view argued that legacy systems are easier to adapt to, with less load on the mind and the body [3]. Yet, our previous findings [1] challenged that claim: the elicited non-legacy gestures were less tiresome.

Despite these different views, there weren't any studies, that we know of, which compared user elicited gestures from the end users' perspective. If the proposed gestures are to be used by the general population, evaluating with participants, other than those who designed the gestures, is necessary. We made 36 participants evaluate the pre-designed legacy and non-legacy gesture sets (Figure 1) from a previous study of ours [1]. We based our gesture selections to our previous work, since it was the only study that made participants create both legacy and non-legacy gesture sets. Also, our contradictory findings in favor of non-legacy gestures called for further exploration.

In the context of legacy bias, as users favor their familiar ways, other features of gestures are still uncovered. Therefore, we formulated our study from commands we define as 'gesture memory' and 'no-gesture memory' ones. To explain, *gesture memory* commands are *gesticulated* with touch screen technologies such as swiping or tapping (e.g. next, open); and *no gesture memory* commands are controlled via their allocated *signs* such as tick or cross (e.g. accept, close). We speculated that with gesture memory commands, the legacy gestures would be favored mostly due to strong gesture familiarity. Whereas with no gesture memory commands, both gesture alternatives would be presented in a more objective light and their features would be discovered.

Overall, we explored the end users' preferences and the reasons behind gesture selections by asking the questions: (1) "How would participants that did not partake in the design process evaluate the gestures?" (2) "Which features of gestures are valued when gesture memory is not prominent?". To the best of our knowledge, this study is the first to evaluate different user-elicited gesture sets from end user's perspective.

Method

36 undergraduate students of varying majors (22 female, 14 male) and ages (M=20.29, SD=1.46) voluntarily participated in our study. Our previous study [1] is the only study made participants create two gesture sets and classify them as legacy (Intuitive – I) and non-legacy (Exclusive – E). In this study we made a comparison between these two gesture sets; 9 commands were chosen, and 18 gestures were tested. Our command selection included commands that have *gesture memory* and *no-gesture memory.* Two-step test provided both quantitative and qualitative data to

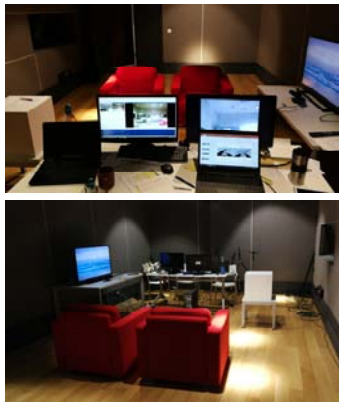

Figure 2: Setting of the user test.

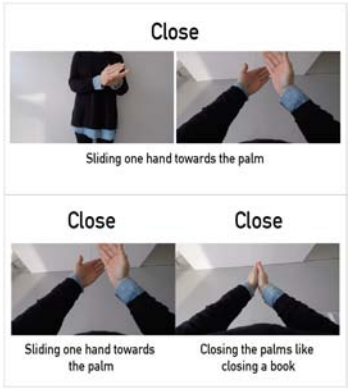

Figure 3: Videos shown during the first (a) and the second (b) steps of the test.

understand the user's mental models. The first step started with the participants standing in front of a TV, with the Wizard of Oz (WoZ) out of their field of vision [2] (Figure 2). A pre-recorded tutorial of a randomly selected gesture was shown three times, which lasted about 1 minute (Figure 3-a). After watching the tutorial, participants performed the gesture and the experimenter initiated the command. For each gesture, a Likert Scale of four criteria (memorability, appropriateness to command, social environment consideration and physical ease) rated from 1 to 7, was given. Participants believed that they controlled a TV, a computer and a sound system, based on the command. In the second part of the test, while the participants were seated, the two gesture options were shown next to each other (Figure 3-b). We conducted a semi-structured interview where participants stated their choices between the gestures (or remain indecisive) and walked us through their selection criterion.

Results & Themes

Comparison of Gesture Sets

In the paired sample T-test the Intuitive gesture set was significantly higher in *memorability* ((I): M=6.30, SD=0.66, (E): M=5.90, SD=0.85, t(35)=3.78, p=.001), *appropriateness to the command* ((I): M=6.21, SD=0.66, (E): M=5.66, SD=0.8, t(35)=3.74 p=0.001), *social environment considerations* ((I): M=6.47, SD=0.77, (E): M=6.02, SD=0.91, t(35)=4.68, p<0.001), *physical ease* ((I): M=6.16, SD=0.92, (E): M=5.79, SD=1.32, t(35)=2.36, p= 0.024.

Gesture by Gesture Analysis

The one by one, paired sample t-test analysis of gestures showed that gestures with 'no memory' (close, volume commands, reject) contradicted with general findings. *Close (E)* (M=6.47, SD=0.91) was significantly more memorable than *Close (I)* (M=5.56, SD=1.23); t (35) =-4.48, p<0.001. Also, participants found *Close (E)* (M=6, SD=0.91) as significantly more appropriate to the command than *Close (I)* (M=6.47, SD=0.91); t (35) =-5.76, p<0.001. (Table 1, green highlight). For the other commands with no gesture memory (*volume commands, reject*), the Intuitive gestures were not significantly different or better than the Exclusive in all four components; p>0.05. (Table 1, orange highlight)

Themes & Discussion

We transcribed and coded the semi-structured interviews for our analysis. This yielded the two themes of this section, in relation with the previous section.

Familiarity over Usability (QA1)

As speculated, among 36, 28 participants chose legacy gestures with 'gesture memory' mostly because of familiarity. On the other hand, 9 out of this 28 people did elaborate on other features. These participants stated that some gestures from the non-legacy set were physically, aesthetically, semantically superior. Much to our surprise, they still selected legacy gestures purely due to familiarity. A participant commented; *"This gesture (Volume (E)) is physically easier and also logical… On the other hand, the knob gesture is something that is familiar. I don't know. I think I will choose the knob (Volume (I))"*.

We had speculated that with no gesture memory commands, the participants would elaborate on both alternatives longer. *Close (E)* was found more memorable and appropriate to the command than *Close (I)*. There were not significant differences between gesture commands for *Volume up/down* and *Reject,* participants found them equally favorable. Two Participants indicated that *Reject (E) w*as more

	I	E	Indecisive
Open	20	12	4
Close	6	27	3
Next	30	4	2
Previous	33	0	3
Increase Volume	15	19	2
Decrease Volume	15	19	2
Select	29	5	5
Accept	22	9	2
Reject	20	15	1

Table 1. Participants selections and gesture by gesture analysis. The highlights show significance between the gestures. (Orange: no difference between the legacy and non legacy gesture, Green: exclusive gesture significantly better)

practical. Therefore, as we suspected, commands with no gesture memory, made participants more objective while evaluating the other aspects.

While presenting their reasons for selecting gestures, 10 out of 36 participants changed their choices to non-legacy and stated that non-legacy evoked more application domains. This is related to the 'affordance' defined as a spectrum of activities made possible by the design [7]. We believe that when participants think of other domains (i.e. controlling curtains with gestures), they lose their familiarity and find non-legacy gestures more appropriate for these domains.

Technological Skepticism (QA2)

In our study participants were not aware of the WoZ, and there were no hiccups during the executions to prime their opinions. However, 16 participants thought that technology might not be able to catch some of the gestures, even though we stated the opposite. We believe that this is an interesting finding, since system designers were the ones criticized for designing gestures with technical restrictions in mind. We see that even participants can be affected by their technological knowledge (we did not find relation between this skepticism and the participants' background) while evaluating novel systems.

Conclusion & Further Work

Our study aimed to answer (1) "How would participants that did not partake in the design process evaluate the gestures?" and (2) "Which features of a command are valued when gesture memory is not prominent?".

For the first question, our study showed that legacy gestures made users relate with the new technology, endorsing previous studies [3, 5]. However, the singular analysis of the gestures revealed that non-designer participants favored 3 gestures of the non-legacy set, coinciding with our previous study [1]. Yet, we also think that the 'design' of the gestures influenced the decisions as much as bias. To illustrate, certain gestures were hard to perform (i.e. Next (E)), regardless of being logical towards functionality. Since we wanted to test user-elicited gestures, we used them without making any changes. Yet, had benchmark analysis as in Nielsen et al. [6] or alternative gesture generation methods been applied, our results could have been different.

Our second question was aimed to discover gestures features and was done by picking *gesture memory* and *no gesture memory* commands. With no memory commands we discovered that some non-legacy gestures featured practicality and more 'affordances'. However, when the gesture memory existed, participants tended to choose legacy gestures, even if they thought the non-legacy version was easier to perform (QA1). This full-on biased finding raises the question; is familiarity alone enough for usability? We plan to discover this and further look at gestures' affordances in a scenario-based experiment. Further work should seek the reasons behind a participant's selections and not limit themselves with the users perceived technological restraints (QA2).

We believe that different means to extract gestures and analysis of the results should be sought out, which we plan to do so. We plan to extend this work with in-depth studies with different demographics to understand how we can enhance the elicitation methods by controlling the effects of legacy bias on the perception of participants. Our aim is to contribute to HCI field by putting forth elicitation methods that can yield to more appropriate outcomes for gesture control.

References

1. Idil Bostan, Oğuz Turan Buruk, Mert Canat, Mustafa Ozan Tezcan, Celalettin Yurdakul, Tilbe Göksun, and Oğuzhan Özcan. 2017. Hands as a Controller: User Preferences for Hand-Specific On-Skin Gestures. *DIS'17*: 1123–1134. Retrieved from http://dl.acm.org/citation.cfm?doid=3064663.3064766
2. N. Dahlbäck, A. Jönsson, and L. Ahrenberg. 1993. Wizard of Oz studies - why and how. *Knowledge-Based Systems* 6, 4: 258–266.
3. Anne. Köpsel and Nikola Bubalo. 2015. Benefiting from legacy bias. *Interactions* 22, 5: 44–47. Retrieved from http://dl.acm.org/ft_gateway.cfm?id=2803169&type=html
4. Meredith Ringel Morris, Andreea Danielescu, Steven Drucker, Danyel Fisher, Bongshin Lee, C. Schraefel, and Jacob O Wobbrock. 2014. Reducing Legacy Bias in Gesture Elicitation Studies. *Interactions* 21, 3: 40–45. Retrieved from http://dl.acm.org/citation.cfm?doid=2608008.2591689
5. Meredith Ringel Morris, Jacob O. Wobbrock, and Andrew D. Wilson. 2010. Understanding users' preferences for surface gestures. *Proc. of GI'10*: 261–268. Retrieved from http://dl.acm.org/citation.cfm?id=1839260
6. Michael Nielsen, Moritz Störring, Thomas B. Moeslund, and Erik Granum. 2004. A Procedure for Developing Intuitive and Ergonomic Gesture Interfaces for HCI. 409–420. https://doi.org/10.1007/978-3-540-24598-8_38
7. Donald A. Norman. 1999. Affordance, conventions, and design. *Interactions* 6, 3: 38–43. Retrieved from http://portal.acm.org/citation.cfm?doid=301153.301168
8. Jacob O. Wobbrock, Meredith Ringel Morris, and Andrew D. Wilson. 2009. User-defined gestures for surface computing. *Proc. CHI'09*: 1083. Retrieved from http://dl.acm.org/citation.cfm?doid=1518701.1518866