# Big Gestures?: Factors that Influence Gesture Visibility

Adrian Reetz
University of Saskatchewan
110 Science Place
Saskatoon, Canada, S7N 5C9
adrian.reetz@usask.ca

Carl Gutwin
University of Saskatchewan
110 Science Place
Saskatoon, Canada, S7N 5C9
gutwin@cs.usask.ca

## ABSTRACT
In many scenarios that involve digital systems, it is beneficial to maintain awareness of other people's actions. During face-to-face communication with fellow humans, this is accomplished by gestures. The recent interest in gestural interfaces offers the possibility to transfer this paradigm to human-computer interaction. Previously, researchers exclusively discussed gesture size as a contributing factor for awareness maintenance through gestures. However, this might be only one piece of the picture, as other factors might be equally important. We studied small (tablet-sized), medium (monitor-sized), and large (full-arm) gestures. Our study showed that, although size does have significant effects, there are other factors that influence awareness maintenance. Our results provide empirical guidance about the ways that gesture size affects awareness, show that other factors, such as gesture morphology, influence awareness, and suggest that gestural interaction has potential for improving group awareness in co-located environments.

## Keywords
Gestures; group awareness

## 1. INTRODUCTION
In 1993, Don Norman describes the usefulness of "big controls and big actions" for shared work:

> The critical thing about doing shared tasks is to keep everyone informed about the complete state of things […] Each pilot […] must be fully aware of the situation, of what has happened, what is planned. And here is where those big controls come in handy. When the captain reaches across the cockpit over to the first officer's side and lowers the landing-gear lever, the motion is obvious: the first officer can see it even without paying conscious attention. The motion not only controls the landing gear, but just as important, it acts as a natural communication between the two pilots, letting both know that the action has been done. […] Automatically, naturally, without any need for talking. [14, p. 142]

What Norman means is that large gestures are useful for system interaction in a group setting because they enable coworkers to maintain awareness of their colleagues' actions through observation. When arguing for the advantage of large gestures, researchers mostly referred to serious and failure-critical work areas, for example, airplane cockpits [14], power-plant operations [6], or train control rooms [10]. Group awareness, however, is useful in other scenarios as well. When playing video games, it is advantageous to know about the actions of teammates or opponents; when using large public displays or art installations, it is useful to coordinate ones actions with other bystanders; even when sitting in a living room and watching movies, it can be useful to see when a friend wants to pause the movie or change the volume. Unfortunately until recently, none of the interactive systems in these scenarios supported gestural input. These days, however, an increasing number of gesture- and motion-capture systems are available to and affordable for a general audience. Supported gesture types range from finger- (e.g., on touch screens) over arm- (e.g., Wii Remote) to full-body-gestures (e.g., Kinect). Now we are able to employ gestural input in a wide range of group scenarios and use large gestures for creating group awareness. With this awareness, we would be able to better coordinate action between co-located people and maintain a better understanding of the state of our digital systems.



What we do not know, however, is how to design for gestures that foster the type of awareness Norman was talking about. He and his colleagues used observational studies to identify *gesture size* as the primary factor gesture visibility. Although this makes sense intuitively, there might be other factors that also have profound influence on gesture visibility, which is important for creating awareness [15]. Ignoring these factors can lead to gesture designs that do not allow people to maintain high levels of awareness and thus strip gestural interfaces from this useful property.

For this research note, we investigated the role of gesture size on gesture visibility. More precisely, we were interested in how reliably people can observe other's gestures (*observation rates*) and how reliably people can identify the performed gesture (*identification rates*). We conducted a user study in which we tested observation and identification rates of differently sized gestures (tablet-, monitor-, and full-arm-sized) from different locations in an environment. At the same time, we sought for other factors that influenced observation and identification rates.

Our findings suggest that gesture size is indeed an important contributing factor for gesture observability and identification. However, we discovered other factors, which, if not considered, can drastically decrease gesture observation and identification rates and, therefore, reduce the amount of awareness that co-located people can create from other's gestural commands. With this note, we contribute to the understanding of awareness creation through gestures. We show that gesture size is an important, but not the only factor for maintaining awareness of other's actions.

## 2. RELATED WORK
Groups are sets of two to five people who carry out tasks in medium-sized workspaces [9]. Whenever people engage in collaborative activities, they have to split their attention between their working task and awareness maintenance. Since these two tasks compete for people's attention, groupware designers try to minimize the cognitive load from awareness maintenance and find a balance between these two tasks [8].

Awareness is the perception and comprehension of the state of the environment [6,9]. Group awareness, that is an understanding of the activities of others, provides context for people's activities and is critical to successful collaboration [5]. Two factors determine the level of group awareness: the actor's nimbus (the space in which actors make their activity available to others), and the observer's focus (the space which is covered by the observer's attention) [2]. When nimbus and focus overlap, observers go through a three-phase process to gain group awareness: perception of an action, comprehension of the situation, and projection of the future status [7].

In HCI, the term "gesture" is used for a wide variety of concepts (overview in [33]). A classification that fits best within the scope of this work might be a combination of 2-D plane-based finger movements [1] and 3-D mid-air arm movements [4]. Executing a gesture involves three phases: a ballistic motion, a corrective phase, and a final acquisition phase where a hand or finger is moved within the bounds of a target or region [12]. Very few papers have implicitly investigated gestures as a method for creating group awareness in co-located environments (e.g., [13]).

# 3. STUDY OF GESTURE OBSERVABILITY

To determine how factors such as gesture size and location affect people's ability to see and interpret these gestures, we carried out a controlled experiment. We picked these factors because they are of high importance for the observability of gestures [15]. We chose a dual-task study setup, i.e. participants had to perform a primary working task at the same time as observing the secondary observation task. We did so because it embodies the trade-off between primary working task and awareness-maintenance task. In our simulated collaborative scenario, the study participants (the *observer*) carried out an attention-demanding choice reaction task (word selection) as their primary task in a room where another person (the *actor*, a confederate) executed various gestures. The observer's job was to maintain awareness of the actor, but without reducing their performance on their primary task.

## 3.1 Participants and Apparatus

We recruited 18 participants from a local university, ages 19 – 45 ($\bar{x} = 29$), 9 female, 9 male. These participants were all experienced with traditional computer systems ($\bar{x} = 35\,hrs/wk$), and were all familiar with gestures on touch-based devices such as mobile phones and tablets.

The study was carried out in a large laboratory (approximately $10 \times 10\ m^2$). The actor's station held a 22" monitor and remained stationary during the study. The observer's cart held a 7" MiMo touch screen, on which the primary task was displayed, and on which the observer indicated their observations and identifications of the actor's gestures.

## 3.2 Conditions: Gesture Size and Location

We defined three gesture sizes: small touch-gestures performed on a 7" hand-held tablet; medium hover-gestures performed approximately 1 $cm$ above a 22" horizontal screen; and large full-arm pointing gestures.

For each of the gesture sizes, we created 6 different gestures (see Figure 1 and Table 1). For small and medium gestures, we chose two gesture types that can be found on most touch screens (tap and swipe) and one geometric gesture (circle). We based our large gestures on systems such as Virtual Shelves [11] or Air Pointing [4], which partition the space around the user into zones. For our system, we used six zones that were arranged in front of the actor ($-90°$ to $+90°$ horizontally, and $-45°$ to $+45°$ vertically).

**Table 1: Gestures with mean magnitude and execution time**

| Small | Medium | Large |
|---|---|---|
| Tap: top left corner (0.40 $m$, 1.8 $s$) | Tap: top right corner (0.77 $m$, 1.7 $s$) | Point: left, high (2.09 $m$, 1.7 $s$ ) |
| Tap: top right corner (0.30 $m$, 1.6 $s$) | Tap: bottom right corner (0.60 $m$, 1.9 $s$) | Point: front, high (1.55 $m$, 1.7 $s$) |
| Circle: top half (0.58 $m$, 2.1 $s$) | Circle: left half (1.10$m$ , 2.6 $s$) | Point: right, high (2.05 $m$, 1.7$s$ ) |
| Circle: bottom half (0.50 $m$, 2.0 $s$) | Circle: right half (1.06 $m$, 2.5 $s$) | Point: left, low (1.16 $m$, 1.7 $s$) |
| Swipe: left edge (0.55 $m$, 1.9 $s$) | Swipe: top edge (1.03 $m$, 2.3 $s$) | Point: front, low (1.16 $m$, 1.7 $s$) |
| Swipe: right edge (0.44 $m$, 1.9 $s$) | Swipe: left edge (1.05 $m$, 2.4 $s$) | Point: right, low (1.88 $m$, 1.6 $s$) |
| Average for small: (0.46 $m$, 1.9 $s$) | Average for medium: 0.94 $m$, 2.3 $s$ | Average for large: 1.65 $m$, 1.7 $s$ |

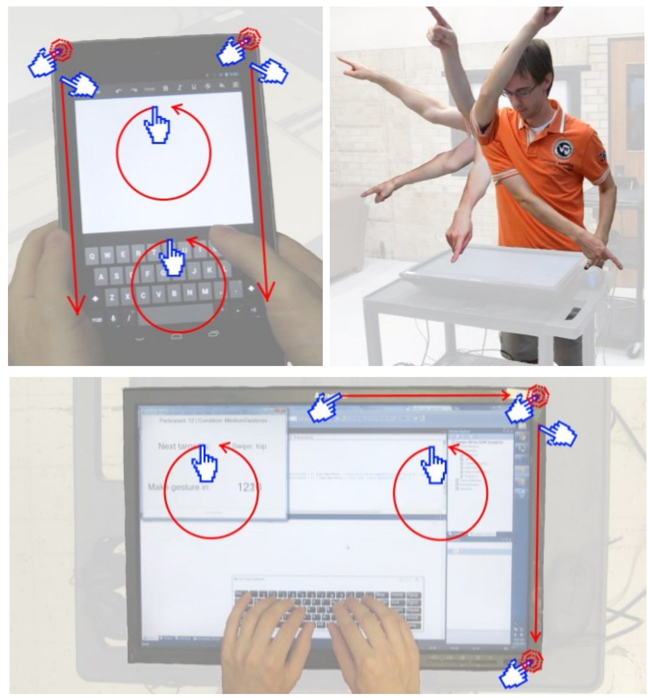

**Figure 1: Overview of small (top-left), large (top-right), and medium gestures (bottom)**

To gain quantifiable values for each gesture size, we measured magnitude (travel distance of index finger) and execution time of the actor's arm movement with an IR-based motion-tracking system (see Table 1 for overview and per-size averages).

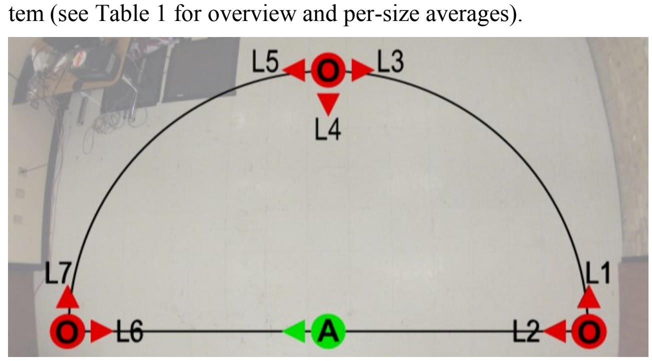


**Figure 2: Observer locations (O) and actor location (A)**

Participants observed the actor from seven different locations (L1–L7), comprising three positions arranged in a semicircle around the actor, and either two or three orientations at each position (facing the actor, or facing perpendicularly away). Figure 2 shows these locations.

### 3.3 Observer's Primary Task

We created an attention-demanding primary task for the observer to perform during the experiment. The rationale behind this is that in a realistic scenario, people are normally involved in a primary task. The task involved repeatedly selecting one of four possible buttons indicated by a written message displayed on the observer's display (displayed on a 7" MiMo touch screen, see Figure 3). Participants were given a short period to complete the selection ($1.0\ s - 2.0\ s$, randomly chosen); if they did not finish their selection in time or made a wrong selection, the system would play a warning sound. After each correct selection, the system would wait $1.0\ s$ and then display the next word-selection task.

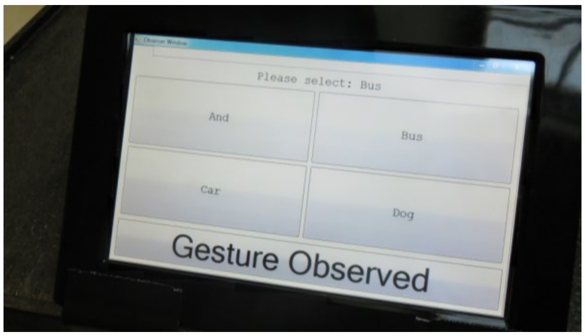


**Figure 3: User interface for the primary working task**

### 3.4 Procedure

After completing a demographics survey and being introduced to the system, participants completed 12 training trials. Participants were then moved to the starting location and asked to start the primary task; they were instructed to maintain awareness of the actor's activities, and report any gestures they observed using their interface (see Figure 3).

The actor then started performing typical tasks at his station, which acted as distractor tasks in between gestures that the observer had to report. The actor texted on the hand-held tablet (small gestures); he typed using the on-screen keyboard on the horizontal screen (medium gestures); and he fidgeted and moved objects around at the cart (large gestures). Within these typical activities, the actor performed a total of 12 gestures (each gesture twice, randomized order) per location. The actor's UI indicated when to perform the next gestures; the interval was randomly chosen (from $1.5\ s - 6.0\ s$). When participants noticed a gesture, they could pause the primary working task, and specify the gesture they just observed from the UI.

### 3.5 Study Design and Hypotheses

The study used a $3 \times 7$ within-participants factorial design with factors Gesture size (small, medium, large) and Location (L1–L7, as shown in Figure 2). Order of gesture size was balanced using Latin square between participants; order of locations was randomized within gesture size.

The actor performed a total of 288 gestures per participant. The system recorded all gesture observations and identifications, and tracked the participant's performance on the primary task.

Our main hypothesis is that size has a great influence on gesture observation and identification rates. However, we strongly assume that other, previously ignored or underestimated factors, play an important role and can ultimately render gesture size as factor.

### 3.6 Data Analysis

To determine the effect of *Gesture size* on observation and identification rates, we analyzed the trials in a $3 \times 6$ repeated-measures ANOVA with Bonferroni-corrections for all post-hoc tests.

## 4. RESULTS

Observation rate is the number of gesture observations made by a participant divided by the number of gestures performed by the actor. Identification rate is the number of gestures correctly identified by a participant divided by the number of observations.

**Table 2: Observation and identification rates [%]**

| **Gesture size** | Small | Medium | Large |
|---|---|---|---|
| **Observation rate: Mean ± Std. err.** | 74 ± 1.9 | 82 ± 1.7 | 83 ± 4.4 |
| **Identification rate: Mean ± Std. err.** | 74 ± 3.1 | 83 ± 3.4 | 92 ± 1.4 |

ANOVA showed a significant effect of *Gesture size* on *Observation rate* ($F(1.1, 18.9) = 9.7,\ p < .01$). Follow-up analyses showed that medium and large gestures had a higher observation rate than small gestures ($p < .05$). ANOVA also showed a significant effect of *Gesture size* on *Identification rate* ($F(2, 34) = 51.2,\ p = .00$). Follow-up analyses showed that all three gesture sizes were significantly different ($p < .01$).

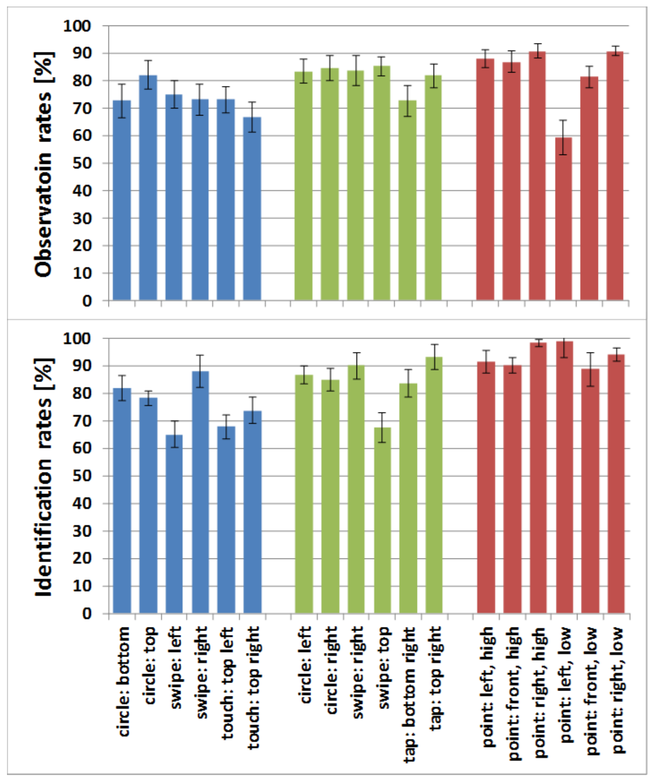


**Figure 4: Observation and identification rates per gestures (small: blue / left; medium: green / center; large: red / right)**

## 5. DISCUSSION

The data supports our hypothesis that size is a major factor for predicting user's gesture observation and identification rates. The plot of each gesture's magnitude (travel distance of index finger) shows a logarithmic relationship between magnitude and observation and identification rates (both $R^2 = .49$) (Figure 5). However, there are some outlying values that cannot be explained by magnitude.

## 5.1 Observation Rates

The large gesture "point: right, low", for example, shows, despite its magnitude, the lowest overall observation rate (60%), and lies $> 3\sigma$ below -the expected predicted value. We speculate that this gesture was performed too closely to the actor's body to be distinguished from the background fidgeting and movements. All large gestures that aimed "high", in contrast, resulted in very good observation rates ($> 87\%$). We conclude that not only the magnitude, but also the direction of a gesture plays an important role in gesture observability: gestures that lead away from the body seem tomight be more observable.

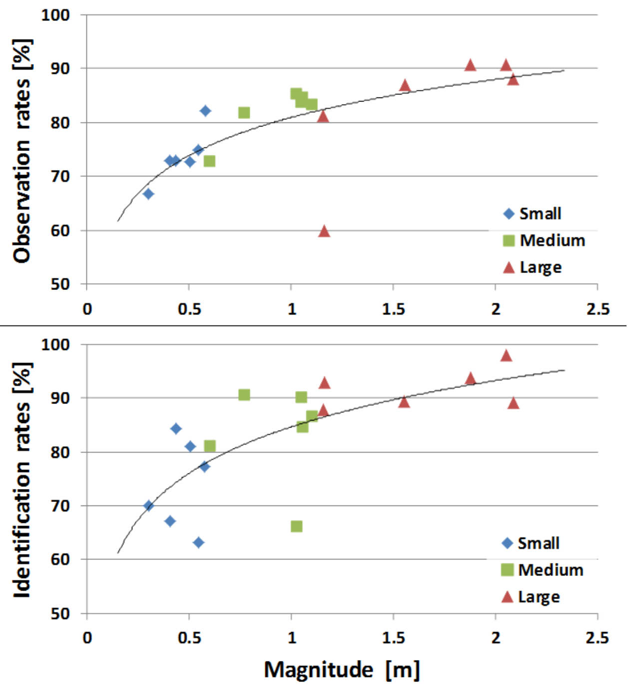


**Figure 5: Observation and identification rates per gesture magnitude**

Our data revealed a cluster of four medium-sized gestures of similar magnitude and execution time: the two "circle" and the two "swipe" gestures. Nonetheless, the shorter and faster "tap: top right"-gesture showed a similar good performance, whereas "tap: bottom right" was less well observed. Again this hints toward our assumption that gestures close to the body are less observable than gestures away from the body. Neither morphology nor magnitude of the gesture seems to play a decisive role.

## 5.2 Identification Rates

The medium-sized gesture "swipe: top" showed a surprisingly low identification rate (66%). An analysis revealed that participants confused this gesture in 18% of all observed cases with "tap: top right". Looking deeper into the interaction between actor and the observer can give us an explanation. During the first part of the ballistic pointing phase, the observer recognizes through peripheral vision that a gesture is about to occur. Then the observer has to switch her foveal vision and cognitive focus toward the gesture. Depending on the duration of the gesture, the actor might have already completed the ballistic phase and entered the corrective or even final acquisition phase. Subsequently, the observer does not perceive the entire gesture but only its final phase. When having two or more gestures with similar final phases, observers will have difficulties telling them apart. This was the case with "swipe: top" and "tap: top right". However, it is noticeable that this confusion was not symmetrical, i.e., very few participants mistook "tap: top right" for "swipe: top".

For small gestures, there appear to be two distinct groups, whose difference cannot be explained by gesture magnitude. A promising approach here might be looking at gesture morphologies, but future research is necessary to gain a better understanding.

# 6. CONCLUSION

In this note we showed that gesture size is an important factor for gesture observability and identification. However, there are other factors involved that can have severe influence, e.g., directionality and similarity of gestures, timing during gesture phases, and discriminability between purposefully executed gestures and unintentional movements (similar to *Midas touch*).

With this note, we wanted to bring back Norman's idea of "big gestures" into the design of gestural interaction, but also show that gesture size alone is not enough to guarantee awareness of other's actions. We hope this work will spark further research into the area of gesture observability and identification since group awareness is a useful tool for collaboration and gestural interfaces will become even more ubiquitous in the future.

# 7. REFERENCES


1. Baecker, R., Grudin, J., Buxton, W., and Greenberg, S. Touch, gesture, and marking. In *HCI: Toward the Year 2000*. Morgan Kaufmann Publishers, 1995.
2. Benford, S. and Fahlén, L. A spatial model of interaction in large virtual environments. *Proc. of ECSCW '93*, 109–124.
3. Cadoz, C. and Wanderley, M. Gesture: music. In *Trends in gestural control of music*. IRCAM, 2000.
4. Cockburn, A., Quinn, P., Gutwin, C., Ramos, G., and Looser, J. Air pointing: Design and evaluation of spatial target acquisition with and without visual feedback. *IJHCS 69*, 6, 401–414.
5. Dourish, P. and Bellotti, V. Awareness and coordination in shared workspaces. *Proc. of CSCW '92*, 107–114.
6. Endsley, M.R. Design and Evaluation for Situation Awareness Enhancement. *Annual Meeting of the Human Factors and Ergonomics Society 32*, 2 (1988), 97–101.
7. Endsley, M. Toward a theory of situation awareness in dynamic systems. *JHFES* 37, 1 (1995), 32–64.
8. Gutwin, C. and Greenberg, S. Design for individuals, design for groups: tradeoffs between power and workspace awareness. *Proc. of CSCW '98*, 207–216.
9. Gutwin, C. and Greenberg, S. A descriptive framework of workspace awareness for real-time groupware. *Computer Supported Cooperative Work 11*, 3–4 (2002), 411–446.
10. Heath, C. and Luff, P. Collaborative activity and technological design: Task coordination in London Underground control rooms. *Proc. of ECSCW '91*, 65–80.
11. Li, F., Dearman, D., and Truong, K.. Virtual shelves: Interactions with orientation aware devices. *Proc. of UIST '09*, 125–128.
12. Meyer, D., Abrams, R., Kornblum, S., Wright, C., and Smith, K.. Optimality in human motor performance. *Psychological Review 95*, 3, 340–370.
13. Morris, M., Huang, A., Paepcke, A., Winograd, T. Cooperative gestures: multi-user gestural interactions for co-located groupware. *Proc. of CHI '06*, 1201–1210
14. Norman, D. *Things that make us smart: defending human attributes in the age of the machine*. Basic Books, 1993.