Making Big Gestures: Effects of Gesture Size on Observability and Identification for Co-Located Group Awareness

Adrian Reetz and Carl Gutwin
Department of Computer Science, University of Saskatchewan
110 Science Place, Saskatoon, Canada, S7N 5C9
adrian.reetz@usask.ca, gutwin@cs.usask.ca

ABSTRACT
Co-located work environments allow people to maintain awareness by observing others’ actions (called consequential communication), but the computerization of many tasks has dramatically reduced the observability of work actions. The recent interest in gestural interaction techniques offers the possibility of recreating some of the noticeability of previous work actions, but little is known about the observability and identifiability of command gestures. To investigate these basic issues, we carried out a study that asked people to observe and identify different sizes and morphologies of gestures from different locations, while carrying out an attention-demanding primary task. We studied small (tablet sized), medium (monitor-sized), and large (full-arm) gestures. Our study showed that although size did have significant effects, as expected, even small gestures were highly noticeable (rates above 75%) and identifiable (rates above 69%). Our results provide empirical guidance about the ways that gesture size, morphology, and location affect observation, and show that gestural interaction has potential for improving group awareness in co-located environments.

Author Keywords
Gestures; consequential communication; group awareness

ACM Classification Keywords
H.5.2. Information interfaces and presentation (e.g., HCI): Interaction styles

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 or member of the control team 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. [25, p. 142]

This kind of information flow, also called consequential communication, has been shown by several researchers to be an important part of the natural way in which people maintain awareness in a group [12, 30, 33]. However, consequential communication depends on large easily-observable actions and controls, which are no longer common in most workplaces. Instead, most tasks are now carried out on general-purpose computers with standard graphical user interfaces. On these computers, activities that once had characteristic actions and artifacts (e.g., getting a file from a cabinet, using a Rolodex to find a telephone number, drawing a diagram, or entering numbers in a ledger) now all look very similar to an observer—that is, they all look like a person sitting at a computer monitor and moving a mouse.

Researchers in distributed groupware have looked at the problem of reduced observability (since people’s bodies are not visible in a distributed setting), and have proposed visualization techniques to make others’ actions in a shared workspace more obvious [13]. However, these enhancements often work only when people are observing the same part of the shared workspace, and the techniques do not provide a solution in situations where people are carrying out loosely coupled work in a co-located setting.

A recent development, and one that could potentially improve observability, is the rise of gestural interaction techniques. Gestures are now common on touch-screen devices of all sizes, and larger gestures that involve full-body interaction have also been extensively studied (e.g., Kinect-based interaction, Virtual Shelves [19], Air Pointing [6], and Skinput [16]). Gestures and full-body interactions bring large easily-observable actions to general-purpose computers, and could thus be a solution to the problem of observability for collocated environments—they could be one way that designers help people maintain group awareness.

There is little information available, however, about whether gestural commands are in fact observable and interpretable, and what size of gesture is needed for an observer to notice the gesture while carrying out other tasks. That is, how should gestures be designed to make possible the kind of group awareness that Norman described?
We carried out an experiment to answer this question. We measured how well participants were able to observe and identify gestures of various sizes and morphologies from different locations in a co-located environment, while carrying out an attention-demanding primary task. We tested three gesture sizes: small gestures that took place on a tablet-sized device; medium gestures that occurred on a monitor-sized display; and large gestures that involved full-arm pointing to different locations in the room.

The study provided three main results:

- Gesture size did significantly affect both observation and identification. However, small gestures had surprisingly high rates of both observation (74%) and identification (69%). Medium and large gestures showed no differences for observation (82% vs. 83%), but large gestures were more accurately identified (92% for large vs. 83% for medium);
- Observation and identification were strongly affected by both location (both rates ranged from 65% to 91% by location) and gesture morphology (observation ranged from 59% to 99%, and surprisingly, the least-observed gesture was one of the large full-arm pointing gestures);
- Participants subjectively rated larger gestures as requiring significantly less effort than smaller gestures, and as being more observable and identifiable; people also strongly preferred the large gestures for the task of maintaining group awareness.

Our results suggest that gestural interaction techniques can provide the foundation for the consequential communication that underlies Norman’s idea of ‘big actions’—and that although gesture size does matter, even small gestures on small devices are surprisingly visible. We considered how smaller gestures could be so well observed; our analysis showed that even when command gestures are small, the preparatory and staging actions that precede them are often much more observable—for example, moving one’s hand up to a mobile device is noticeable action that draws attention to the upcoming command gesture.

Our study provides empirical evidence about the benefits of gestural interfaces for collaborative activity, and is the first to analyze the specific effects of gesture size, morphology, and location on observability and identification. Our work indicates that designers could use gestural interaction techniques as a way to improve the natural communication and awareness that occur in co-located work environments.

RELATED WORK

Groups and Mixed-focus Collaboration

Groups are sets of two to five people who carry out tasks in medium-sized workspaces [14]. 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 (see next section for examples) and find a balance between these two tasks [13]. When people work in a group, they often engage in mixed-focus collaboration, i.e. people shift frequently between loosely and tightly coupled activities during a work session [7]. Coupling refers to the degree with which people have to interact to progress with their work [29]. When people are loosely coupled, they have to interact less with each other to complete their task as when they are tightly coupled [29]. However, even during loosely coupled work, people still need to be aware of others’ activities [26].

Group Awareness and Consequential Communication

Awareness is the perception and comprehension of the state of the environment [8, 14]. Group awareness, that is an understanding of the activities of others, provides context for people’s activities and is critical to successful collaboration [7]. 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) [3]. 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 [9]. There are several methods for creating group awareness in colocated environments [12]: direct communication, indirect productions, consequential communication, feedthrough, and environmental feedback.

Consequential communication occurs through visible or audible signs of interaction with a workspace [30]. The size of the actions necessary to operate controls makes actions public and creates situation awareness, which is important in many collaborative real-world tasks [30]. In HCI research, consequential communication is frequently mentioned as an awareness mechanism, and observational studies show that it is frequently used in real-world situations [18, 30]. However, it is rarely explored in controlled studies and occasionally considered to be of little importance [31]. This is in contrast to other fields, which showed that consequential communication plays a crucial role throughout life, for example, as facilitator for learning through observation and imitation [15, 28].

Gestures

In HCI, the term gesture is used for a wide variety of concepts (see [5] for an overview). 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 movements [6].

Initially, swipe gestures were used to move elements around a touch screen [20]. Researchers extended these simple gestures to include geometric forms, script, and multi-touch input [10, 17]. Next, researchers created technologies that enables the use of gestures on different surfaces, such as the human arm [16] and virtual space in front of the user [11, 21]. Two-D plane-based gestures have been investigated in great detail. Researchers have evaluated, for example, learnability [23], ergonomics [24], and social as-
pects [27] of gestures (see [35] for a recent overview). Very few papers, however, have implicitly investigated gestures as a method for creating group awareness in co-located environments. Morris et al., for example, used collaborative gestures to create awareness in a co-located work environment [22].

Full 3-D input using arm gestures was first mentioned by Bolt [4] and explored by Baudel and Beaudouin-Lafon [2]. Unfortunately, most papers in this area focus on technical aspects, and only two articles explore human factors of full-arm gestures in HCI: Virtual Shelves [19] and Air Pointing [6]. Neither of these works have investigated the consequential communication aspect of full-arm gestures.

**A STUDY OF GESTURE OBSERVABILITY**

To determine how factors such as gesture size, morphology, 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 [32]. We chose a dual-task study setup 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 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. With our setup, the primary task required the majority of participants’ attention. The task required foveal vision focus on the touch screen, thus forcing participants to temporarily dedicate their full attention. Furthermore, participants completed the primary task 2.5x as often as the actor performed gestures.

**Study Methods**

**Participants and Apparatus**

We recruited 18 participants from a local university, ages 19 – 45 (̅x = 29), 9 female, 9 male. These participants were all experienced with traditional computer systems (mean 35 hours / week), 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 m x 10 m), in which we placed two moveable carts holding the study computers. The actor’s cart held a 22” monitor and remained stationary during the study. The observer’s cart was moved to several different locations during the session (see Figure 2). It 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.

**Study Conditions: Gesture Size, Morphology, 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.

![Figure 1: Overview of small gestures (top), medium gestures (center), and large gestures (bottom)](image)
Naturally, the index finger travelled the longest distance: \( \bar{x} = 0.46 \, \text{m} \) (small gestures), \( \bar{x} = 0.94 \, \text{m} \) (medium gestures), and \( \bar{x} = 1.65 \, \text{m} \) (large gestures). Small gestures were performed in \( \bar{x} = 1.9 \, \text{s} \), medium gestures in \( \bar{x} = 2.3 \, \text{s} \), and large gestures in \( \bar{x} = 1.7 \, \text{s} \).

Table 1: Gestures with mean magnitude and execution time

<table>
<thead>
<tr>
<th>Small</th>
<th>Medium</th>
<th>Large</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tap: top left corner (0.40 m, 1.8 s)</td>
<td>Tap: top right corner (0.77 m, 1.7 s)</td>
<td>Point: left, high (2.09 m, 1.7 s)</td>
</tr>
<tr>
<td>Tap: top right corner (0.30 m, 1.6 s)</td>
<td>Tap: bottom right corner (0.60 m, 1.9 s)</td>
<td>Point: front, high (1.55 m, 1.7 s)</td>
</tr>
<tr>
<td>Circle: top half (0.58 m, 2.1 s)</td>
<td>Circle: left half (1.10 m, 2.6 s)</td>
<td>Point: right, high (2.05 m, 1.7 s)</td>
</tr>
<tr>
<td>Circle: bottom half (0.50 m, 2.0 s)</td>
<td>Circle: right half (1.06 m, 2.5 s)</td>
<td>Point: left, low (1.16 m, 1.7 s)</td>
</tr>
<tr>
<td>Swipe: left edge (0.55 m, 1.9 s)</td>
<td>Swipe: top edge (1.03 m, 2.3 s)</td>
<td>Point: front, low (1.16 m, 1.7 s)</td>
</tr>
<tr>
<td>Swipe: right edge (0.44 m, 1.9 s)</td>
<td>Swipe: left edge (1.05 m, 2.4 s)</td>
<td>Point: right, low (1.88 m, 1.6 s)</td>
</tr>
</tbody>
</table>

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. Six locations formed a symmetry: L1–L7, L2–L6, and L3–L5. The first two pairs, however, differ in a way that in L1 and L2 participants are behind the actor; in L6 and L7 they are in front.

Observer’s Primary Task

In order to simulate a realistic work environment, we created an attention-demanding primary task for the observer to perform during the experiment. 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 another message.

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 handheld 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 to 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.

Study Design and Hypotheses

The study used a 3 x 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 \((12 + 12 \times 7) \times 3 = 288\) gestures per participant. The observer’s system recorded all gesture observations and identifications, and tracked the participant’s performance on the primary task. After the experiment, participants filled out a basic demographic questionnaire, one NASA TLX form per gesture size, and one ranking questionnaire.

We formulated four hypotheses for our study:

H1. Larger gestures will be observed at a higher rate than smaller gestures.

H2. Larger gestures will be identified more accurately than smaller gestures.

H3. Larger gesture size will reduce the negative effects of occlusion.

H4. Facing the actor will have higher observation and identification rates.
Data Analysis
We performed a univariate ANOVA to investigate the effect of Gesture Size and Location on primary task performance measured as reaction rate. To determine the effect of factors Gesture size and Location on observation and identification rates, we analyzed the trials in a 3 × 7 repeated-measures ANOVA. We carried out separate analyses of our dependent measures by gesture morphology (since morphologies were not the same across sizes) with a 3 × 6 RM-ANOVA. We evaluated the TLX data using a repeated-measures ANOVA, and we analyzed the rank data using a Friedman test for k related samples. All post-hoc tests used Bonferroni corrections.

RESULTS
Primary Task Performance
We did not find effects of Gesture size ($F(2,216) = 1.1, p > .1$) and Location ($F(6,216) = 1.8, p > .1$) on primary task performance. We found, however, a significant effect of Participant x Gesture size for 8 participants ($F(34,216) = 4.3, p < .01$). When looking more closely at this finding, we saw that all affected participants performed significantly worse with the first gesture size they saw during the experiment. We concluded that the training phase was too short for them to achieve their highest level of proficiency. Since we counter-balanced the order of gesture sizes between participants and therefore controlled for this factor, we felt confident that primary task performance was independent from Gesture size and Location. As a result, we omitted it from all further analyses.

Observation Rate and Identification Rate
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. Sphericity was violated for observation rate by both Gesture size and Location (Mauchly’s test: $p = .00$), and for identification rate by Location (Mauchly’s test: $p < .01$). For these analyses, we use Greenhouse-Geisser corrections.

Effects of Gesture Size
On average, participants showed the highest observation (see Table 2 and Figure 4) and identification rates (see Table 3 and Figure 5) with large gestures, followed by medium and small gestures.

Observation Rate
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 ($p < .05$).

Identification Rate
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 ($p < .01$).

Table 2: Observation and identification rates [%]

<table>
<thead>
<tr>
<th>Gesture size</th>
<th>Small</th>
<th>Medium</th>
<th>Large</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observation rate: Mean ± Std. err.</td>
<td>74 ± 5.0</td>
<td>82 ± 4.3</td>
<td>83 ± 3.0</td>
</tr>
<tr>
<td>Identification rate: Mean ± Std. err.</td>
<td>69 ± 3.8</td>
<td>82 ± 3.5</td>
<td>92 ± 2.0</td>
</tr>
</tbody>
</table>

Table 3: Observation and identification rates [%]

<table>
<thead>
<tr>
<th>Location</th>
<th>L1</th>
<th>L2</th>
<th>L3</th>
<th>L4</th>
<th>L5</th>
<th>L6</th>
<th>L7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observation rate: Mean ± Std. err.</td>
<td>65 ± 4.5</td>
<td>76 ± 4.2</td>
<td>80 ± 4.5</td>
<td>89 ± 2.1</td>
<td>78 ± 5.7</td>
<td>91 ± 2.5</td>
<td>79 ± 5.5</td>
</tr>
<tr>
<td>Identification rate: Mean ± Std. err.</td>
<td>68 ± 4.5</td>
<td>77 ± 3.4</td>
<td>91 ± 5.0</td>
<td>91 ± 1.8</td>
<td>85 ± 3.7</td>
<td>89 ± 2.5</td>
<td>76 ± 3.3</td>
</tr>
</tbody>
</table>

Effects of Location
Observation Rate
ANOVA showed a significant effect of Location on Observation rate ($F(2.4, 40.0) = 15.8, p = .00$). As shown in Table 1, the different locations were associated with a wide
variety of observation rates: the highest at L6 and L4, and the lowest at L1 (see Figure 6 for a map of observation rates by location). Follow-up analyses showed that the observation rate at L1 was significantly worse than from all other locations ($p < .05$), and L6 had a higher observation rate than its symmetric counterpart L2 ($p < .01$).

**Identification Rate**

ANOVA also showed a significant effect of Location on Identification rate ($F(3.2, 53.7) = 13.4, p = .00$). As shown in Table 3 and Figure 7, participants had the highest identification rate from L4, followed by L6, and the worst observation rate from L1. The identification rate from L1 was significantly worse than from L3 through L6 (all $p < .05$), and L4 and L6 had significantly higher identification rates than L1, L2, and L7 (all $p < .01$).

**Gesture Size x Location Interaction**

**Observation Rate**

ANOVA showed a significant interaction between Gesture size and Location ($F(5.7, 97.4) = 4.1, p < .01$) for Observation rate. As shown in Figure 4, small gestures were significantly better observed from L4 and L6 (both $\bar{x} = 0.90$) than from L1 ($\bar{x} = 0.52$) and L2 ($\bar{x} = 0.62$) (all $p < .01$). Observation rate from L1 was significantly worse than from all other locations except L2 (all $p < .05$). As expected, mean differences were high between symmetric locations L1–L7 (.22) and L2–L6 (.28) and low between L3–L5 (.06) and L4–L6 (.00).

Medium gestures were best observed from L6 ($\bar{x} = 0.93$) and L4 ($\bar{x} = 0.91$) and worst observed from L1 ($\bar{x} = 0.69$). Observation rates from L1 were significantly worse than from L3 through L6 (all $p < .05$) and from L2 worse than from L6 and L4 (both $p < .05$). Compared to small gestures, observation rate in L2 improved close to average ($\bar{x} = 0.79$). As expected, mean differences became lower between symmetric locations L1–L7 (.12) and L2–L6 (.14) and stayed low between L3–L5 (.03) and L4–L6 (.01).

**Effects of Gesture Morphology**

We analyzed Gesture morphology separately within each gesture size (since they were different across sizes). Since
sphericity was violated for all measures (Mauchly’s test: all $p < .05$), we use Greenhouse-Geisser corrections.

For small gestures, ANOVA showed a significant effect of Gesture morphology on Observation rate ($F(3,6,61.1) = 5.7, p < .01$) and on Identification rate ($F(3,0,51.4) = 6.1, p < .01$); for medium gestures, ANOVA showed a significant effect of Gesture morphology on Observation rate ($F(3.5,60.0) = 5.7, p < .01$) and on Identification rate ($F(2.7,46.5) = 5.3, p < .01$); for large gestures, ANOVA showed a significant effect of Gesture morphology on Observation rate ($F(2.4,61.0) = 17.1, p = .00$), but not Identification rate.

**Observation Rate**

![Image of Observation Rate graph](image)

**Figure 8: Observation rates per gestures (small: blue / left; medium: green / center; large: red / right)**

**Identification Rate**

![Image of Identification Rate graph](image)

**Figure 9: Identification rates per gestures (small: blue / left; medium: green / center; large: red / right)**

We found that participants observed the small gesture circle: top significantly more often than the gestures tap: top left, tap: top right, and circle: top (all $p < .05$). For medium gestures, participants observed tap: bottom right significantly less often than the gestures circle: left, circle: right, tap: top right, and swipe: top (all $p < .05$). Among large gestures, point: left, low had a significantly lower observation rate than any other large gestures (all $p < .05$), and was the least-observed gesture at any size.

Participants showed a significantly higher identification rate for the small gesture swipe: right than for the gestures swipe: left and tap: top right (all $p < .05$). For medium gestures, participants identified swipe: top significantly less often than all gestures except tap: bottom right ($p < .05$). For large gestures, there were no significant differences.

**Subjective Measures**

Participants rated their experience using the NASA TLX questionnaire. Overall, participants felt that larger gestures were less effortful and less frustrating than smaller gestures.

**Figure 10: Participant preference rating**

We found a significant difference in mental demand between all three gesture sizes (all $p < .01$). For physical demand and frustration, there were significant differences between large gestures and medium and small gestures (all $p < .05$). Finally, participants rated small gestures as more effortful than medium and large ones (both $p < .01$).

**Figure 11: NASA TLX results**

We also asked participants to rank the different gesture sizes in terms of perceived visibility, recognition accuracy, and their preference to work with (Figure 11). (We discarded the data from one participant because the questionnaire was not filled out correctly.) A significant majority of participants ranked large gestures most visible ($\chi^2(2,17) = 21.5, p = .00$) and most recognizable ($\chi^2(2,17) = 19.9, p = .00$). Overall, 14 of 17 participants preferred to work with large gestures over small and medium ones ($\chi^2(2,17) = 17.3, p = .00$).
DISCUSSION

In this discussion, we first explain how our results confirm our hypotheses and discuss some additional insights we gained from analyzing our results. Then, we come back to our premise and lay out how our findings support Norman’s idea of “big controls and big actions”. We describe some use cases, mention potential directions for future work, and address issues that come with the use of big gestures. Finally, we list the limitations of our work.

Hypotheses

We confirmed all of our hypotheses.

H1: Larger gestures lead to higher observation rates

As predicted, participants showed significantly higher observation rates with large and medium gestures than with small gestures, and higher observation rates with large gestures than with medium gestures. While this result is true on the (categorical) gesture-size scale (small—medium—large), we also found a similar pattern when looking at the (continuous) gesture magnitude. Figure 12 illustrates the logarithmic relationship between gesture magnitude and observation rate ($F(1,15) = 119.0, p = .00, R^2 = .89$). However, our regression analysis revealed that one large gesture (“point: left, low”) was a residual outlier. For the curve fit, we removed this outlier (case-wise analysis with 3σ cutoff); we talk about this case later in the discussion.

We want to emphasize that the logarithmic relationship continues across different gesture sizes and morphologies (2D touch and hover gestures as well as 3D pointing gestures). This means we can generalize our findings to most gestures that fit our definition in the related work section.

We found a logarithmic relationship between magnitude and identification rate, similar to the one between magnitude and observation rate ($F(1,15) = 19.3, p < .01, R^2 = .56$). Not surprisingly, the effect is smaller because there are other factors that affect identification rate. Again, our regression analysis revealed, that one gesture (medium size, “swipe top”) was a residual outlier. For the curve fit, we removed this outlier (case-wise analysis with 2σ cutoff); we will come back to this particular case later in the discussion.

Figure 12: Observation rate per gesture magnitude

H2: Larger gestures lead to higher identification rates

Participants showed significantly better performance with large gestures than with medium and significantly better performance with medium than with small gestures. The overall identification rate of larger gestures is better than that of smaller gestures; even when observed, larger gestures are easier to identify than smaller gestures.

Identification rates of all gestures were affected in similar ways by occlusion than observation rate. For medium gestures, differences in identification rates between L1 and L7 and between L2 and L6 decreased with increasing gesture size. This implies that small gestures suffer significantly from occlusion and that this effect diminishes with increased gesture size. With an unobstructed view to the actor, gesture size does not affect performance. However, in multi-display environments where people move around freely, it is likely that occlusion will occur; in this case, larger gestures can enable higher group awareness.

H3: Larger gestures are less affected by occlusion

In locations L1 and L2, gestures were occluded by the actor’s body. A comparison of symmetrical pairs L1–L7 and L2–L6 therefore shows how much occlusion affected participants’ observation rate. Our results showed that the mean differences in observation rate between L1 and L7 and between L2 and L6 decreased with increasing gesture size. This implies that small gestures suffer significantly from occlusion and that this effect diminishes with increased gesture size. With an unobstructed view to the actor, gesture size does not affect performance. However, in multi-display environments where people move around freely, it is likely that occlusion will occur; in this case, larger gestures can enable higher group awareness.

H4: facing the actor leads to higher performance

In locations L2, L4, and L6, participants were facing the actor, in locations L1, L3, L5, and L7, they were perpendicularly seated to the actor. When pairwise comparing L1–L2, L3–L4, L5–L4, and L6–L7, we found that participants performed on average better when facing the actor. However, most of these comparisons showed no significant differ-

Figure 13: Identification rate per gesture magnitude
ence. While these results might sound surprising, they are in accordance with theory. Vision research has shown that human response to rapidly moving targets is almost invariant with its location in the field of vision [34].

Additional Findings and Research Questions

Are all gestures of one size equally easy to observe?

For small gestures, we found that “circle: top” was the easiest gesture to observe, significantly easier than both “taps” and “circle bottom” (all \( p < .05 \)). This was most likely because it had the longest execution time (2.1 s) and largest magnitude (0.58 m) among all small gestures.

For medium gestures, we found that “tap: bottom right” was significantly harder to observe than any other gesture except “swipe: right” (all \( p < .05 \)). Contributing factors were its low execution time (second lowest in its category: 1.9 s) and its small magnitude (smallest in its category: 0.60 m).

For large gestures, we found that participants showed a significantly lower observation rate with “point: left, low” than with any other gesture (all \( p < .05 \)). As before, we assume that mostly execution time (1.75 s) and lack of magnitude (1.16 m) are responsible for this effect. In addition, the gesture was performed very close to the body, which made it more difficult to spot than other large gestures, which were all performed away from the actor’s body.

Are all gestures of one size equally easy to identify?

For small and large gestures, we found no gesture that was consistently better or worse than the other ones.

For medium gestures, however, we found that participants performed significantly worse with “swipe: top” than with any other gesture except “tap: bottom right” (all \( p < .05 \)). A detailed analysis showed that we can attribute more than half of the errors to confusing this gesture with the gesture “tap: top right”. These two seemingly different gestures share a similar post-stroke hold and retraction phase. Apparently, participants oftentimes required the preparation and stroke phase of the actor’s gesture to shift their attention from their primary working task to the perception phase of consequential communication. To make gestures more distinguishable, we therefore recommend avoiding gestures that end with similar strokes and the same post-stroke hold and retraction. For example, the small-gesture swipes were rarely confused with the small-gesture taps.

Are large gestures generally easy to observe and identify?

Ironically, the least likely observed gesture in our study was a large one. A good strategy to make large gestures visible is to make them lead away from the actor’s body.

Did the labels “left” and “right” confuse participants?

All directions in gesture descriptions were meant to be relative to the actor. As a result, there was a danger that participants confused left and right and top and bottom when they were in front of the actor (e.g., his “left” became their “right”). We analyzed all errors in conditions L6 and L7; no participant systematically confused any of these labels.

“Big Controls and Big Actions”

Norman’s original idea was that big controls and big actions create awareness. Our results showed that gestures, independently from their size, are indeed observable and can therefore improve group awareness: people know that something has happened. When looking at identification rate, we can also give an initial estimation for the next step toward group awareness, knowing what exactly has happened. Our results indicate that people can distinguish between at least six gestures. We also showed that identification rate depends on more factors than observation rate. A more thorough investigation of these factors could give us more insight about potential limitations, such as upper limits of an alphabet of discernible gesture, as well as guidelines for designing distinguishable gestures. Another important issue is finding gesture sets with different levels of observability, so that interaction designers can select a gesture that matches an action’s desired publicity.

There are many cases in which people would want to make their actions public. Public gestures can be part of, for example, co-located multiplayer games where the group should be aware of certain actions. Another example are Scrum-teams where the team should know about the completion of a single Scrum-task. Likewise, there are many cases in which people want to keep actions private or do not want to distract others. As said before, our findings show that people can control the publicity or privacy of their actions through gesture size.

There are, however, some disadvantages to large gestures. For example, they require more physical effort, and there are some socio-cultural restrictions to the use of big gestures. Again, we assume that large gestures will mostly be used in group environments, where each member accepts and understand large gestures in the context of their work.

Limitations

There are a couple of limitations to our study. While we selected our gesture sizes to reflect a broad variety of gestural interfaces, we only used a typical set of gestures within each size and not a broad variety of all possible gestures. This allows us to only give an initial assessment and lower boundary about identification rates, leaving a more systematic approach to future work. Common contextual and semantic knowledge, for example, can increase identification rates. In addition, our study took place in a controlled laboratory environment. We are confident but cannot guarantee how much our findings apply to a real-world scenario.

Conclusion

In this paper, we demonstrated that gestural interaction techniques can be used for creating visible device interaction, thus laying the groundwork for providing consequential communication to co-located collaborators. We measured observation and identification rates of different gestures and showed that even small gestures are visible and could create consequential communication. However, larger gestures are more easily observable, mainly due to a re-
duced effect from occlusion. In addition, increasing size makes gestures more easily identifiable. Considering our findings, we encourage interaction designer to include gestural interfaces in their groupware applications, so that users can benefit from having consequential communication as an implicit method for gaining group awareness.

REFERENCES
10. Goldberg, D. Richardson, C. Touch-typing with a stylus. Proc. of the CHI ’93, 80–87
17. Hauptmann, A. Speech and gestures for graphic image manipulation. Proc. of CHI ’89, 241–245