Open Sesame: Re-envisioning the Design of a Gesture-Based Access Control System

Michael Karlesky

Polytechnic Institute of New York
University (NYU•Poly)
5 MetroTech Center
Brooklyn, NY 11201
mkarle01@students.poly.edu

Edward Melcer

Polytechnic Institute of New York
University (NYU•Poly)
5 MetroTech Center
Brooklyn, NY 11201
emelce01@students.poly.edu

Katherine Isbister

Polytechnic Institute of New York
University (NYU•Poly)
5 MetroTech Center
Brooklyn, NY 11201
isbister@poly.edu

Abstract

We present results of an exploratory study employing a Wizard of Oz mockup of a new gesture-based access control system we are constructing for our lab's entryway. Among user interactions witnessed, we have identified several behaviors of interest to security researchers and HCI researchers alike. We discuss our security system design approach as an extrapolation of two identified trends, demonstrating the potential for the felt experience of pleasurable and playful systems to help solve difficult interaction problems. We also show the great value of prototyping a mockup to reveal designers' assumptions about human interactions with new technology use cases.

Author Keywords

Kinect; movement-based interaction; gesture-based access control system; behavioral biometrics; wizard-of-oz mockup; pleasure; Third-wave HCI.

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms

Human Factors; Design; Security.

Introduction

We participate in an interdisciplinary security research group. Our contribution is *Open Sesame[1]*, a gesture-based access control system for our lab that replaces existing electronic key cards with behavioral biometrics.

Trendspotting: New Approaches for Security Design
Designing for secure systems is terrifically challenging. Security policies and practices are rarely popular with administrators or users leading to breaches. Open Sesame is the embodiment of two trends that could eventually fundamentally alter secure system design.

In the so-called Third Wave of HCI [3], users' emotional state and measures other than productivity are preeminent. If we marry this trend to security design, then security begins with usability, catering to users' desired actions and real behavior in favor of restricting users to a security technology's apparent limits. Thus, we conceived Open Sesame to create momentary, pleasurable, felt experiences with the hope of finding happy users who also adopt secure behaviors.

Depth sensing, gesture recognition, and biometric technologies are going mainstream with wide adoption by consumers (primarily in gaming). We offer that it may soon be more economical in physical security to choose two installations of such technologies (one for training and one for sensing) than to produce and manage a security token for each user. Open Sesame is an exploration of this scenario and its tradeoffs.

Figure 1: An early, staged photo reveals many of our initial and ultimately incorrect interaction design assumptions: users would stop to gesture; users would always walk in the hallway approach; gestures would be solely upper body motions; users would gesture at a distance from the door; and users would be aware of the invisible sensing envelope.

---

[1] "Open Sesame" is the secret phrase uttered to unseal a magic cave in *Ali Baba and the Forty Thieves* of medieval Arabic literature.

Initial (Incorrect) Assumptions for Interaction Design
Open Sesame is a work in progress. To our knowledge, no system in use or in the literature utilizes gesture to unlock doorways. Consequently, we have no prior work to guide us. We will detail our flawed assumptions (see Figure 1) as revealed by an exploratory design study with a Wizard of Oz mockup and draw conclusions to guide design for momentary gesture spaces.

Background and Motivation

Usability and Security
The growth of literature on the topic of usability in security indicates it is becoming a significant academic concern [6]. A broad review of publications reveals that researchers typically view usability and security to be in tension (e.g. [2]). That is, they hold a pessimistic view assuming more secure systems are less usable while more usable systems are less secure. Our aim is to create a working counterexample to this viewpoint.

Wellbeing and Pleasurable / Playful Interfaces
In recent decades positive psychology has emerged with a focus on maximizing human experiences of contentment, happiness, satisfaction, etc. Work to integrate these ideas and HCI is a new direction [9]. Noting the engagement and fun present in game interactions, HCI researchers have sought to introduce lessons learned in the gaming world into non-game systems [11]. With Open Sesame, we are repurposing gaming technology (i.e. Microsoft Kinect) with the intent to create a pleasurable, momentary interaction that may also be fun [5]. Framing the issue in Maslow's hierarchy of needs, we are working to address usability in a security system by leapfrogging from functionality over usability to pleasurability [8].

Energy

4.70
4.25
3.80

Happiness

5.10
4.50
3.90

Powerful

4.70
4.35
4.00

"In Charge"

5.00
4.50
4.00

Anxiety

2.80
2.60
2.40

Comfortable

5.30
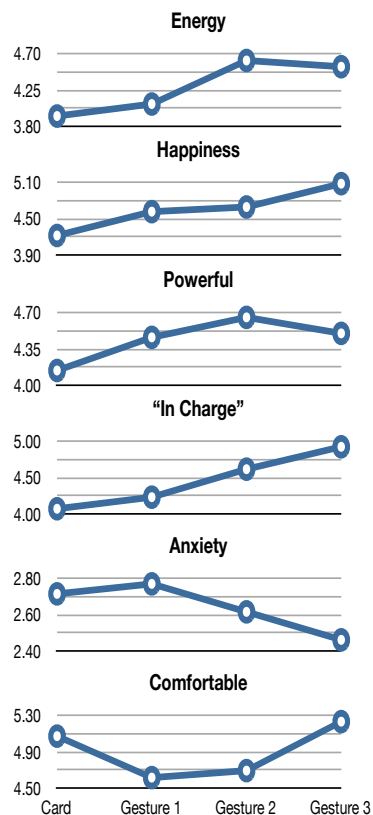4.90
4.50

Card   Gesture 1   Gesture 2   Gesture 3

Figure 2: Mean values for all participants' self reports measures for Likert-style questions across four interactions (use of existing card reader plus three self-selected gestures). Note: not a statistically significant sample size.

Related Work

Novel Access Control Systems

Mobile phone-based research projects are numerous in the access control literature and not only implement wireless unlocking but also schemes to confer authorization to users lacking it. An example of such a system is *Grey*; Grey is also notable for its in-depth analysis with respect to usability in security [1].

Accelerometers in mobile phones, game controllers, and specialized hardware are at the heart of a number of research projects exploring behavioral biometrics for in-air gesture use. The *uWave* project is an example [10]. Others have used the multitouch surfaces of smartphones and tablets to capture the behavioral biometrics of touch-based gestures; of note, work of ours serves as an example in this category [15].

Motion and Emotion

Research clearly links bodily movement / posture to human emotional state. Evidence of this has been demonstrated, for example, in self report [16] and in measurable biochemical changes [4].

Exploratory Design Study Method

We performed a non-statistically significant study with fourteen participants to discover faulty assumptions, elicit user reaction, and gauge operational bounds.

Wizard of Oz Setup

The study setup was situated at the entryway to the NYU•Poly Game Innovation Lab — home of the eventual working system. Gesture recognition was performed by the study administrator (an author) with the door unlock action triggered by remote control.

Instruments

We used three paper instruments. Two of the instruments, the Affect Grid [14] and Self-Assessment Manikin (SAM) [12], are quick assessments of arousal and valence. The SAM instrument also attempts to measure "dominance" (i.e. how "big" or "small" a participant feels). A third instrument includes a series of Likert-style survey questions as well as a small number of questions asking for free-form responses.

Basic Structure of Study

The study was comprised of seven doorway interactions and a final semi-structured interview. These interactions included using the existing card system, play-acting predefined emotional states, and self-selecting moods for use with the entryway. Participants were left to individually construct a concept of the system and interact with purely imagined sensors.

Results and Observations

To be clear, our study results are but a design exploration to guide initial development and a more rigorous study thereafter — to see what we could see.

Summary of Quantitative Results

On self-reports of energy, happiness, power, and "in charge", the mean for all users was greater for self-selected gestures than for card reader interactions (see Figure 2). We find this consistent with existing work connecting emotion and bodily motion: movements larger than a card swipe increase arousal and valence.

On self-reports of anxiety and comfort, we see a complementary pattern. Anxiety initially peaks with the first self-selected gesture interaction then declines below that of the card reader exercise. Similarly, self

reported comfort dips after the card reader interaction and then climbs above it (see Figure 2). We surmise the anxiety and discomfort associated with the novelty and performative elements of self-selected gestures abated with repeated acts, ultimately at levels lower than the familiar card swipe action. We interpret these results to suggest promise in presenting a security interface users find actively pleasurable. Of course, issues of novelty, accuracy, and efficacy can only be rigorously addressed with the final, working system.

Study Participant Gestures
INTERACTION FOCUSED ON DOOR
All fourteen participants focused their attention and interactions on the door. If a gestural interaction was a sentence with the participant as its noun and the gesture as a verb, the direct object receiving the gesture's action was the door itself. Gaze, body language, and post study comments make this plainly clear — the participants interacted with the door itself.

THREE SPACES: PERSONAL, THEATER, AND EFFICIENCY
We identified three styles of relating to the entryway space: "personal space", "theater space", and "efficiency space" (see Figure 3).

Eleven of our study participants approached closely to the door itself before gesturing to unlock it. When asked, consistent with results seen by Reeves and Nass [13], these participants perceived a "personal space" of the door to be engaged at a socially appropriate distance. Four of our participants viewed the space in front of the door as a performative theater space, employing large, sweeping actions including twirls and spins. A single test subject performed an "efficient" gesture at a distance, thus announcing their intent to unlock the door. After making this gestural announcement, this user then expected their approach to trigger an unlock event once the (now pre-authorized) user was within an arm's reach of the door.
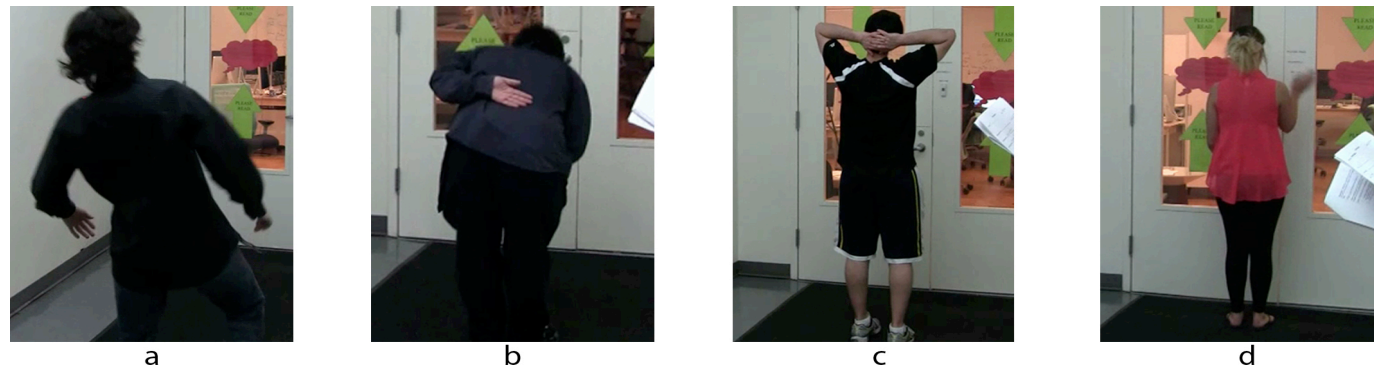

a   b   c   d

Figure 3: (a) Theater Space — mid spin. (b) Theater Space — finishing with a bow. (c) Personal Space — stretching / yawning. (d) Personal Space — one of many witnessed variations on waving "hello."

Violation of Expectations

Before encountering our potential users' wide variety of interaction styles, we, the designers, envisioned a simple interaction mechanic and a technical architecture amenable to it. We had assumed users would walk into sensor range and immediately gesture using only their torso (see Figure 1). Consequently, we originally saw interactions as the sum of a bisection of each user — lower body for gait analysis and upper body for gesture analysis. We came to these assumptions by way of our knowledge of the Kinect's range (approx. 1.2–3.5 m) and the challenges of writing complex software. We expected separate walking actions and upper body gestures because these are easier to process.

Discussion, Design Changes, Implications

Design Changes

To address the incongruity between our initial assumptions and real user behavior, we must supplement our use of Kinect cameras with another depth-sensing technology able to operate in a shallow range within inches of the door itself. Secondly, as users have enacted gesture variety far exceeding our expectations, we have abandoned gait recognition in favor of analyzing whole body motion.

User Nouns, Gesture Verbs, and (Direct) Objects

Third wave HCI has arrived [3]. Gesture-based interfaces are on the rise. From our study we see that in momentary interactions with in-air gesture-based systems our participants naturally and consistently looked to identify a receiver of their gesture actions. Consequently, we see cause for designers to create in-air gestures systems with this principle in mind. In fact, referencing our earlier observation, we find that a simple sentence structure is an effective construct for such systems: users are nouns, gestures are verbs, and artifacts in interaction are the direct objects receiving a user's gesture actions. Framing gestures with this construct is the (literal) embodiment of the language/action perspective of interaction [18]. Effective gesture designs must consider the noun and direct object — not only verbs.

Security, Usability, Pleasure, and Artifacts

We interpret our findings to show promise in designing for pleasure (even playfulness) in security to yield good usability as a desired side effect. At this early stage we offer Open Sesame as a possible example of the radical HCI intervention Dourish, et al. advocate for security use in ubiquitous computing specifically and security interfaces of all sorts generally [7]. In fact, the socialness and expressivity witnessed in our user study go beyond security considerations and handily reveal the performative and intertwined nature of our relationship with our artifacts [17].

Conclusion

We have presented a Wizard of Oz design study of our in-air, gesture-based access control system Open Sesame currently in development. We detailed our goals for the system and demonstrated the immense value of an exploratory study to uncover assumptions regarding user behavior and preference and gauge operational bounds. Looking ahead to new systems with momentary gesture interactions, we offered suggested design principles for such systems. Finally, in addressing usability in security systems, we made a preliminary case for the value of designing for pleasure so as to concurrently maximize usability.

Acknowledgements

References

[1]   Bauer, L., Cranor, L.F., Reiter, M.K., and Vaniea, K. Lessons learned from the deployment of a smartphone-based access-control system. ACM (2007).

[2]   Braz, C., Seffah, A., and M'Raihi, D. Designing a trade-off between usability and security: A metrics based-model. *Human-Computer Interaction–INTERACT 2007*, (2007), 114–126.

[3]   Bødker, S. When Second Wave HCI Meets Third Wave Challenges. ACM Request Permissions (2006).

[4]   Carney, D.R., Cuddy, A.J.C., and Yap, A.J. Power Posing: Brief Nonverbal Displays Affect Neuroendocrine Levels and Risk Tolerance. *Psychological Science 21*, 10 (2010), 1363–1368.

[5]   Carroll, J.M. Beyond fun. *interactions 11*, 5 (2004), 38–40.

[6]   Cranor, L.F. and Garfinkel, S. *Security and Usability: Designing Secure Systems That People Can Use*. O'Reilly Media, Cambridge, 2005.

[7]   Dourish, P., Grinter, R.E., Delgado de la Flor, J., and Joseph, M. Security in the wild: user strategies for managing security as an everyday, practical problem. *Personal and Ubiquitous Computing 8*, 6 (2004), 391–401.

[8]   Jordan, P.W. Human factors for pleasure in product use. *Applied Ergonomics 29*, 1 (1998), 25–33.

[9]   Kanis, M. and Brinkman, W.P. HCI for Positive Change. *the 13th Computer-Human Interaction Netherlands Conference*, (2009), 127–130.

[10] Liu, J., Zhong, L., Wickramasuriya, J., and Vasudevan, V. uWave: Accelerometer-based personalized gesture recognition and its applications. *Pervasive and Mobile Computing 5*, 6 (2009), 657–675.

[11] Malone, T.W. Heuristics for designing enjoyable user interfaces: Lessons from computer games. *Proceedings of the 1982 conference on Human factors in computing systems*, (1982), 63–68.

[12] Morris, J. "Observations: SAM: The Self-Assessment Manikin-An Efficient Cross-cultural Measurement of Emotional Response." *Journal of Advertising Research 35*, 6 (1995), 63–68.

[13] Reeves, B. and Nass, C. The Media Equation. In *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*. Center for the Study of Language and Information, 2003.

[14] Russell, J.A., Weiss, A., and Mendelsohn, G.A. Affect Grid: A single-item scale of pleasure and arousal. *Journal of personality and social psychology 57*, 3 (1989), 493.

[15] Sae-Bae, N., Ahmed, K., Isbister, K., and Memon, N. Biometric-rich gestures: a novel approach to authentication on multi-touch devices. ACM Request Permissions (2012).

[16] Strack, F., Martin, L.L., and Stepper, S. Inhibiting and Facilitating Conditions of the Human Smile: A Nonobtrusive Test of the Facial Feedback Hypothesis. *Journal of personality and social psychology 54*, 5 (1988), 768.

[17] Suchman, L. *Human-Machine Reconfigurations: Plans and Situated Actions*. Cambridge University Press, New York, 2006.

[18] Winograd, T. A language/action perspective on the design of cooperative work. *Human-Computer Interaction 3*, 1 (1987), 3–30.