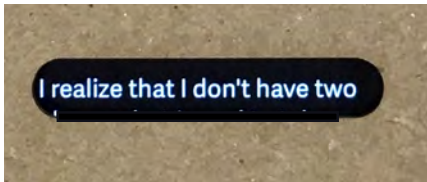


Wall, An Eccentric Design Probe: Exploring and Exposing the Sense-and-Extract Paradigm



The Last Word display showing a sentence with a presumed error.

James Pierce

University of Washington Seattle,
Washington, USA
jppierce@uw.edu

Gabrielle Queen

University of Washington Seattle,
Washington, USA
gabrique@uw.edu

Kristin Tapang

University of Washington Seattle,
Washington, USA
ktapang@uw.edu

This work is licensed under Creative Commons Attribution-NoDerivs International 4.0. DIS '25, July 5–9, 2025, Funchal, Portugal © 2025 Copyright is held by the owner/author(s). ACM ISBN 979-8-4007-1485-6/2025/07. <https://doi.org/10.1145/3715336.3735411>

Abstract

We describe our process of conceptualizing, designing, prototyping, and using Wall, an eccentric smart device intended to encourage reflection on the design qualities of surveillant sensing systems. As a result of our conceptual design process, we first summarize four design qualities that define a sense-and-extract interaction paradigm. We then document our process of developing Wall. Finally, we reflect on our experiences using and living with Wall, and the experiential and theoretical insights we gained through our self-use studies.

Author Keywords

Design, research through design, surveillance, smart devices, speculative design, probes

CSS Concepts

Human-Computer Interaction

Introduction

Framed with our broader research program that investigates the design qualities of spatial sensing systems, this Pictorial describes our research through design (RtD) process [9,26,29] of conceptualizing, designing, building, using, and reflecting upon a functional technology probe called Wall.

Wall is an unusual everyday smart device that invites reflection from users, designers, and researchers on the experiential, social, and ethical aspects of surveillant spatial sensing systems. Wall continu-

ously listens, learns, and displays information gleaned from auditory environmental data. While we have designed many variations of Wall, including advanced versions that use large language models (LLMs) to predict future conversations, here we focus on a simpler functional version of Wall that displays transcripts of nearby speech and auditory events (e.g., laughing, coughing).

We outline our RtD process, beginning with our foundational theoretical inquiry and concluding with grounded, empirical reflections on our use of Wall.

In conclusion, we discuss how our work illuminates a new interaction paradigm we call sense-and-extract. We reflect on the design qualities of Wall and other zone interfaces, and present empirically grounded findings, including themes of aesthetic surveillance and value of technical errors and imperfections for meaning-making and privacy.

Conceptual Backdrop

Wall is part of a larger set of design concepts we produced to explore the qualities of surveillant sensing systems (page 2). In future publications, we intend to detail our process of designing, reflecting, and articulating theoretical insights about new interaction paradigms. Here we summarize insights developed through our conceptual design process, and which we explore concretely using Wall:

1. Sense-and-Extract Interactions

Traditional user interfaces represent an *act-then-react* paradigm. Press a button and a screen updates, for example. We refer to these as *point interfaces*. New technologies with sensors, connectivity, and AI capabilities have propelled a new paradigm

of *sense-and-extract* interfaces. For example, a camera and microphone sense environmental data, the device runs object recognition software, and an alert is sent to a user elsewhere notifying them that a specific person's face was detected. Sense-and-extract system employ what we call a *zone interface*, which involves spatial sensor fields.

2. Design qualities of zone interfaces

We have identified four qualities that characterize zone interfaces and the sense-and-extract paradigm, and which help us articulate its design opportunities, limitations, and concerns:

→ **Sensitization:** They transform physical spaces into zonal input fields. *Space becomes an interface.*

→ **Extraction:** They capture data and infer information from the environment and its actors, objects, and events. *Space becomes exposed to data-capture.*

→ **Imposition:** They project extractive capabilities onto physical spaces (e.g., with cameras or mics), or inject them into physical objects (e.g., with GPS). As tangible buttons are replaced with invisible sensor fields, *inherent points of interactional consent, control, and awareness disintegrate.*

→ **Disjunction:** They can lead to outputs that lack a clear and reliable cause-effect relationship. *Input/output relations become looser, less correlated, and harder to predict.*

We return to these concepts in the conclusion, and refer to them throughout. Next we describe the eccentric family of devices we designed, which led us to these four qualities of a sense-and-extract interaction paradigm.

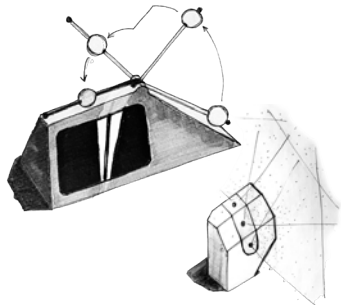
A FAMILY OF ECCENTRIC PRODUCTS

We conceived of Wall as part of a family of devices for exploring properties and potentials of smart sensing products, such as smart speakers with voice assistants, GPS trackers for locating objects and people, and security cameras. We refer to these as spatial sensing systems. Through our conceptual exploration we iteratively identified experiential themes and theoretical concepts for understanding the opportunities, limitations, and concerns of these powerful yet controversial technologies.

Design qualities. Building on prior research [1,15,18,17 23], here we define “design qualities” with a three-part framework: *[1] Opportunities* - what are the positive affordances, uses, and values that spatial sensing systems support?, *[2] Limitations* - what are the practical, technical, and experiential limitations of spatial sensing systems?, and *[3] Concerns* - what issues should we consider with regard to values such as consent, control, privacy, autonomy, and transparency? This framework was in part inspired by foundational design curriculum, including Industrial Design handbooks that summarize “design opportunities” and “technical limitations” of materials and processes such as injection molding, slip casting, and wood joinery [27]. Our framework adds a social, experiential, and ethical dimension to an understanding of design qualities.

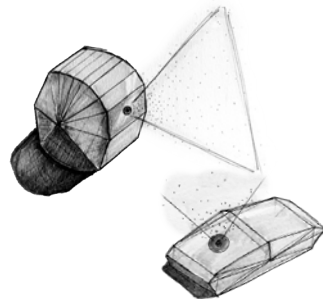
Eccentric design approach. We adopt an eccentric approach similar to that outlined by Pierce [22] in which the researcher develops thematic systems of unusual devices, along with variations of each design, in order to investigate a design research space. A key aspect of this process is to allow aesthetic and affective potentials to drive design forward, rather than letting rational, functional norms overly constrain the exploration. Our approach is further inspired by various experimental and speculative approaches to RtD including ludic design [10], counterfunctional design [24], slow design [20], reflective design [25], phatic design [3], speculative enactments [6], and material speculation [28].

Articulating theoretical concepts through conceptual designs. We emergently searched for concepts to help us articulate qualities of these strange devices, their “personalities,” and why we found them intriguing. This dialogical process led us to the concept of a sense-and-extract paradigm and the qualities of sensitization, extraction, imposition, and disjunction.



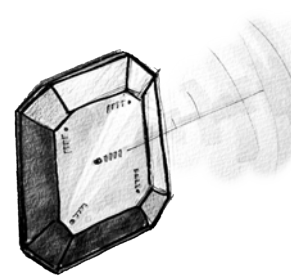
01 - FRAMES

Frames is an eccentric surveillance camera device that televisually displays a spectrum of extraction states ranging from basic video streaming to advanced object detection and synthetic media generation. Frame has two units: Frame 1 (From) and Frame 2 (To), which function as sensor and display units respectively.



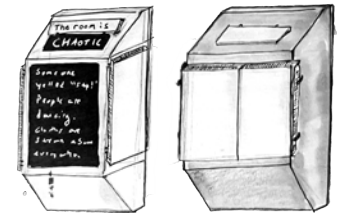
02 - SHELL

Shell is an eccentric camera, mic, and GPS device that cannot be powered off. It continuously extracts data from its surroundings. Users can physically move Shell to alter its sensing capabilities. Shell is positionally unstable and tends to float around. It uses loud noises and bright lights to compel people to recharge it when its power is low.



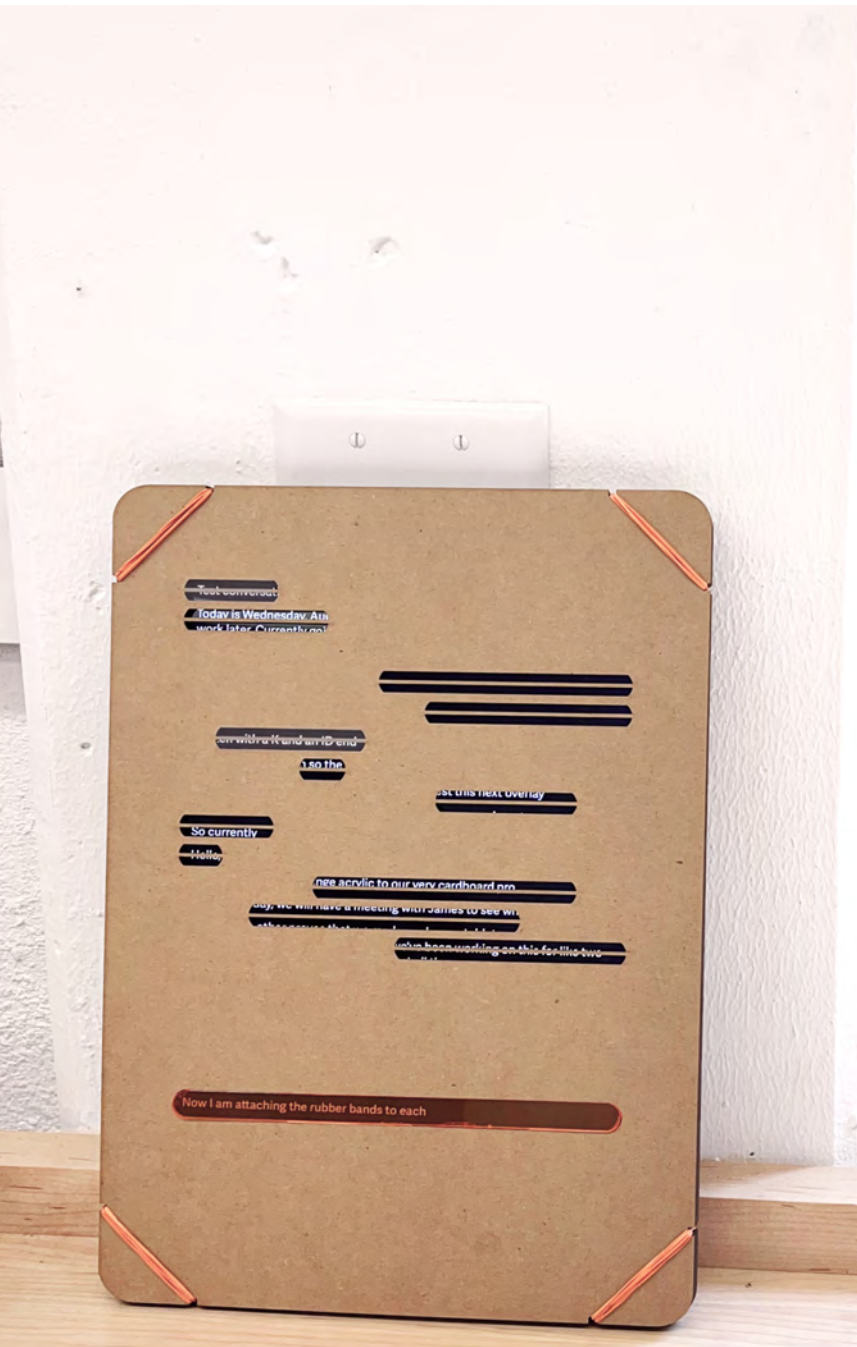
03 - WALL

Wall is an eccentric smart speaker system that continuously listens, learns, and textually displays its interpretations and speculative extrapolations to nearby users. Depending on the selected mode, Wall may or may not let those nearby know it is listening.



04 - ROOM

Room is an eccentric environmental monitoring system that interprets the social, emotional, and behavioral micro-climate of an architectural space. Room is constantly attempting to “read the room.”



PROBING FURTHER...

Among our family of eccentric device concepts, we decided to bring one them to life through functional prototyping. Rather than adopt a high-fidelity prototyping or product development approach, we decided to follow an approach inspired by technology probes [12], cultural probes [8], and other prototyping strategies that cleverly combine simple materials and off-the-shelf components. We were particularly inspired by Chatting et al's innovative uses of mobile phones in constructing Ritual Machines [3] and Paulo Cardini's Monotask phone covers [7].

We chose to develop lightweight probes rather than technically complex prototypes for two reasons. First, we wanted to move quickly and test our conceptual design in the wild for extended periods of time. While custom hardware and software would allow us to test more advanced features, it would also greatly extend our timeline.

Second, we saw an opportunity to develop a form language that supported our goals of reflection and provocation by partially exposing the underlying off-the-shelf hardware and software systems we employed, including iPhones, iPads, Zoom, and an accessibility app called Ava that transcribes conversations and identifies auditory events, such as "water running" and "coughing."

← A version of Wall that uses an iPad running a text-to-speech app called Ava. Our probes utilize simple and inexpensive materials, such as cardboard and rubber band, and they "expose the seams" of their construction to encourage reflection on the current technologies that literally lie beneath the surface.

Among our eccentric product family we determined that Wall was the ideal concept to pursue for the following reasons.

Engaging - We anticipated that Wall would be entertaining and thought-provoking to engage with across many different user demographics and contexts, including home, work, and public.

Simple - Compared to other concepts, the backend and frontend computing components can be easily implemented with off-the-shelf products (e.g., Zoom, smart phones, external USB microphones).

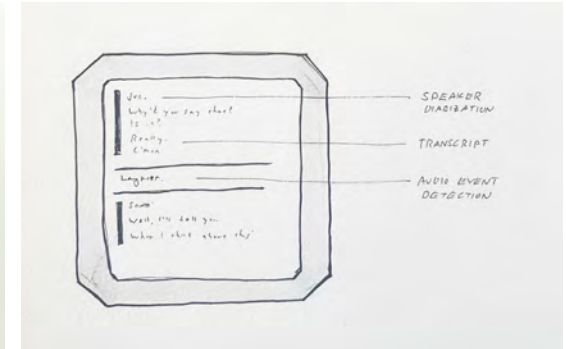
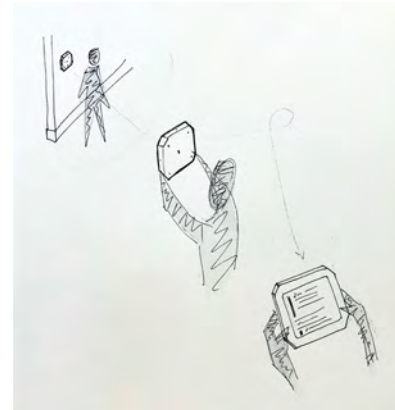
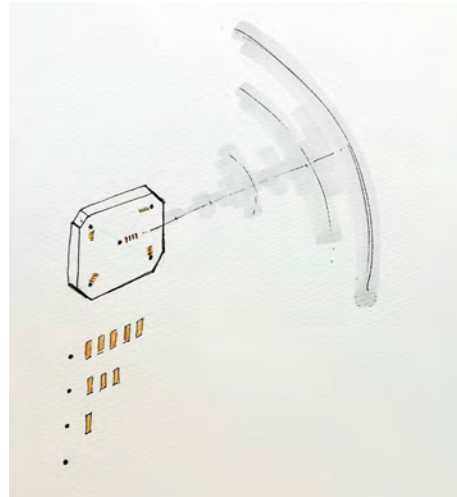
Social - We liked that Wall was potentially a more social than strictly personal device, allowing it to be deployed across many different situations, such as cafes, households, and our work studios.

Aural - We also liked that Wall utilized microphone and text, but not image and video. In a world saturated with video and imagery, we were keen to explore less visually dominant designs.

Original Concept Vision

The original concept for Wall envisioned a device that hung on the wall like a picture frame. It could be configured in several display and sensing modes, ranging from highly transparent and visible to completely concealed via disguise. In the most visible mode, anyone nearby could read the live captions and other information displayed. In the concealed modes, only users who knew about the object's secret capabilities would periodically remove Wall to review its surreptitiously extracted information.

The original vision supported a range of extraction modes, each with increasingly more advanced powers. These ranged from a basic transcription mode to a predictive mode that uses generative AI to speculatively display future conversations.



Early Prototype Explorations

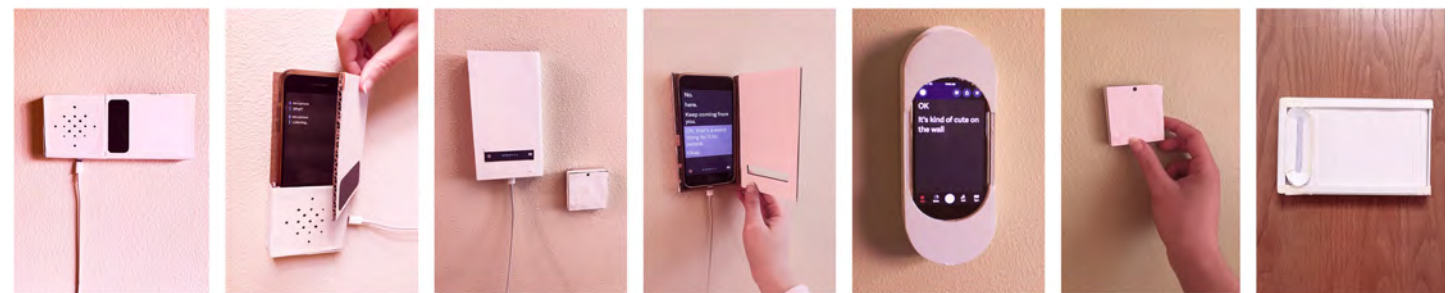
As we pivoted to our probe approach, we narrowed our concept to a simple transcript version of Wall that would utilize commercially available smartphones and software applications.

Using iPhones, iPads, and various apps such as Zoom, iOS Live Captions, and Ava, we explored many different form factors using rapid prototyping.

These early explorations considered different patterns for how users would interact with Wall: Frequently or infrequently? Quickly or for extended periods of time? Through reading the display or more engaged forms of mode-switching? High or low visibility?

Our early prototypes used hinge and slider mechanisms to facilitate adjusting the visibility of the display. However, we eventually

converged toward a “set it and forget it” interaction model that encouraged users to keep Wall in a single display and extraction mode, but did allow them to alter these with some effort by removing the phone or tablet from the enclosure and adjusting the app running in the background.



Refining the Form & Form Language

Once we determined the “set it and forget” cadence of user interaction, we simultaneously considered Wall’s form language (the semiotics of how its materials and shape communicate) and the practical aspects of its form factor—including viewability, positioning, routing of power cords, and so on.

We determined that Wall would be the first of a family of eccentric probes, and so the form language should be applicable to other types of hardware platforms and form factors such as Shell, Room, and Frames (see Page 2). We also decided that we wanted to differentiate our eccentric probes from mainstream consumer smart products. Instead, we sought a form language that featured economical small-scale production, and that communicated our design philosophy of eccentric sensing and lightweight exploratory prototyping.

We took inspiration from a wide variety of designers who thoughtfully use low-cost materials found at hardware and art supply stores, including the Dutch conceptual design studio Droog, Google’s Cardboard AR viewer, and aforementioned works by Chatting et al and Paulo Cardini.

We leaned into the use of cardboard, paper, magnets, rubber bands, tape, and other low-cost everyday materials. We converged upon a form factor that centered on a cardboard enclosure supplemented with functional and aesthetic details in acrylic, rubber bands, and rice paper.

Functional Features

In parallel with our form language development, we explored various features and requirements. The main feature set involved a set of overlay screens cut to expose elements of the underlying iPhone/iPad speech-to-text app. These cutouts ranged

from exposing the entire screen to fully concealing it.

We maintained the initial vision of enabling a spectrum of display modes, from fully exposed to concealed. Conceptually, this was intended to allow users to explore varying levels of imposition on subjects within the environment, and corresponding levels of consent and awareness.

We also tried to support as much of the original vision as possible with regard to extraction modes, but there were limited options using low-cost consumer applications. Users could switch between a Zoom version with simple transcription or other apps that also include event captioning. With more effort, users could use Otter AI to produce speaker diarization (differentiating between speakers).

Our final version struck a balance between enabled users to alter the display while encouraging them to let Wall listen, record, and reveal without user interruption.



WALL - THE FINAL PACKAGE

We produced six usable sets of Wall probes. Each set contained five overlay screens for two different platforms, an iPhone X and an iPad Pro. Each probe includes a main sleeve that holds the iPad or iPhone in place, with a specialized opening for insertion and removal. The five overlays for each sleeve and are both attachable and stowable with their respective methods (rubber bands for iPad and a tabs for iPhone).

Each of the three authors adopted one of each set to use and test within their homes and our studio workplace.



rubber bands

overlay

iPad



more overlays

tabs

overlay

iPhone



more overlays

OVERLAY MODE SWITCHING

In creating the overlays for each probe, we decided to focus on five different points of interaction that provided varying levels of the transcript's visibility and interactive elements, like start/stop/pause recording.

Since there are two distinct devices, we used this as an opportunity to explore different variations of each overlay concept (i.e. redacted, transcript, etc.).

Due to its size, the iPad overlays are secured with four rubber bands. The iPhone overlays use four tabs, which work better at this scale.

Each format allows users to change overlay modes, but this is subtly discouraged due to the moderate effort involved. This "set it and forget it" interaction cadence was a deliberate choice. We wanted to encourage users to engage with a mode for an extended period of time, even if it initially appeared boring or uneventful. Silent mode, for example, demands patience and later engagement with transcripts. Open and Last Word/Line Modes allow for immediate meaningful feedback.



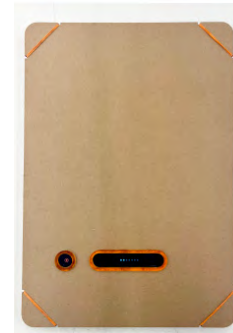
01 - OPEN

The entire screen is available for users to view and interact with, including live transcripts, start/stop/pause buttons, and review of archived transcripts. All nearby actors are invited to engage with Wall in Open mode.



04 - REDACTED

Lines of live transcription text can be seen, but they are obfuscated by overlays that render them difficult or impossible to read.



02 - MUTED

Users can only see that the probe is active through our exposing of the "listening" icons. Users can also start/stop/pause the extraction process. Unfamiliar users may notice that a digital device is active, but have no other clues as to Wall's extractive activities.



05 - SILENT

Users can see a faint glimmer of activity, but nothing else. Unfamiliar users may assume Wall is a decorative object, and not a surveillance device.

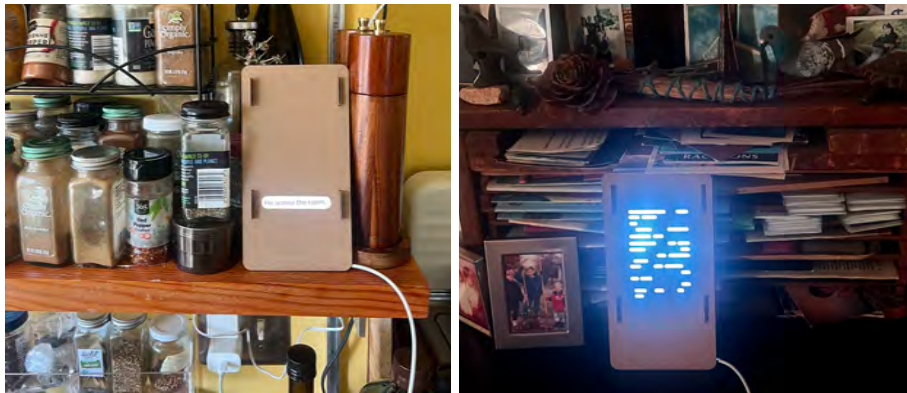


03 - LAST WORD/LINE

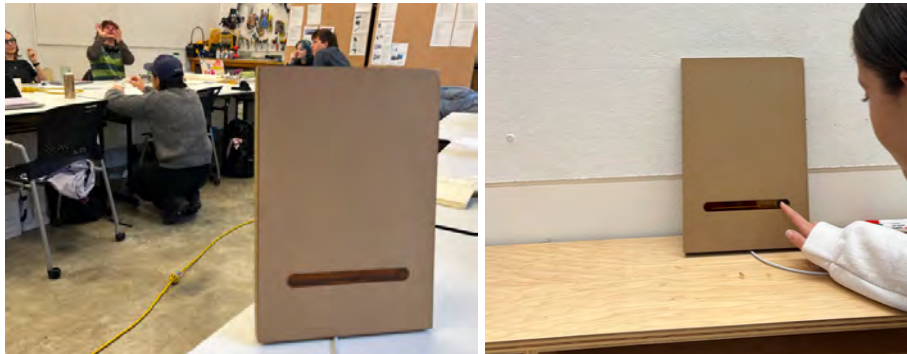
Only the most recent line of transcribed text is viewable. This display field is transient because as soon as new words are captured, the Last Word/Line will be replaced. All nearby actors can view clear and immediate feedback regarding Wall's imposed, extractive functionality.

DISGUISED (NOT INCLUDED)

We decided to omit the fully concealed or disguised version of Wall included in the original vision, both for ethical concerns and because Silent mode sufficiently enabled a highly imposed mode of low visibility and awareness extraction.

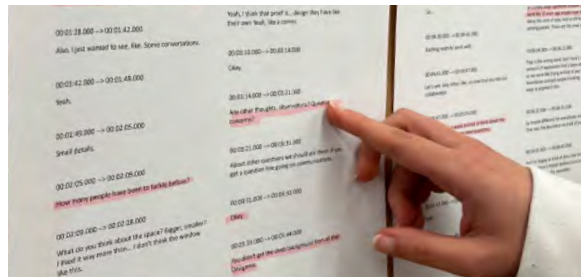


Wall positioned in Author 1's two favorite spots: Kitchen area near the dining table (left) and living room near the sofa (right).



Authors 2 and 3 experiment with Wall in the Industrial Design Studio at [institution withheld].

Authors 2 and 3 experiment printing transcripts and posting them on our design studio at [institution withheld].



SELF-USE STUDIES

We decide to adopt a first-person [5], or autobiographical [19], approach to assessing firsthand use of our Wall probes. We took this self-use approach for several reasons.

First, self-use offers particular advantages including those that are especially relevant to the sensitive, intimate nature of our project. Self-use allows the research team to deeply reflect on personal experiences in ways that are impossible using traditional qualitative methods. In our case, we were interested in the content of intimate conversations. We sought to probe issues of consent and privacy by positioning Wall within sensitive areas of the home.

Second, our inquiry combines theoretical, empirical, and what we think of as material inquiry. Material inquiry is practice-based research that involves designing and making artifacts. The empirical strand of inquiry is but one of three core threads. We felt that a lab or field study with external participants was not needed for this particular study and might, in fact, overload our reporting in this Pictorial with too much empirical data. Instead, we view our self-use deployments as a complement to our theoretical and material inquiry.

Third, self-use offers a flexible method that allowed us to pause and resume our studies at our own pace. This enabled us to switch between periods of operational use, dedicated reflection, theorization, and stepping away from the project entirely to focus on other activities.

The major limitations of self-use include the biases of the researchers (including a confirmation and selection bias towards an artifact we created). The data input from self-use studies is also limited to the number and perspectives represented by the

research team. Diverse and outside perspectives are not readily included.

Authors 1, 2, and 3 engaged in multiple forms of self-use. We deliberately did not develop a coherent plan or roadmap. Instead, we iteratively experimented with different forms and durations of self-use in manner very similar to a conventional iterative form exploration and prototyping within a design process. That is, we chose to be highly experimental in our self-use.

Our self-use studies included the following formats.

Short Experiments. Our initial round of self-use studies were similar to rapid user testing sessions. For example, we conducted a series of 1-2 hour experiments where we setup Wall in different situations (e.g., at a coffee shop, when friends were over for dinner). We then wrote 1-2 page documentations of our experiments that included summaries of "What was interesting" and "Areas for improvement."

Social Deployments. On several occasions, we setup Wall within different areas of our studio spaces at the University of Washington, including our research studio and one of our Industrial Design studios. We also experimented with ways of engaging with transcripts, including printing out booklets and posting them to walls of the studio. These deployments were of limited duration, and we explained the operation and purpose of the devices to our colleagues who entered into these environments.

Extended Home Use. Our deepest insights came from extended forms of home use, which we describe next.

6 Months Living On and Off with Wall

In this section we focus on the experiences of Author 1, and occasionally refer to similar experiences shared by Author 2 and 3. This allows us to go deeper into Author 1's experience of using Wall for 1-4 week durations from August 2024 - January 2025. Participation was limited to close members of our households. The functionality of Wall was fully explained, consent was given, and household members had full control over Wall including the ability to pause or stop its recording.

What follows is not a formal analysis of participant data, but rather a personal reflection, organized into themes. First-person narrative is used throughout, and intimate details are intentionally included to allow the reader to understand the authors' firsthand experiences of Wall.

Dynamics: Shifting Access, Consent, Control

I (Author 1) and my spouse (pseudonym: Anne) lived with Wall for interspersed durations of several weeks over the course of about 6 months. The lack of consistent use was related to several factors, including travel, busyness, and personal issues described further below.

I first setup Wall at my home in August 2024. I provided Anne with a basic overview of Wall and explained that it was always or often recording. Anne was accustomed to occasionally living with my unusual research prototypes and experiments, and had no concerns about privacy. We found Last Word mode to be the most enjoyable for daily use, as it struck a balance between inhibiting us from over-reading what was said but also provided a level of visibility that created intrigue and delight.

A dynamic was soon established where I was the de facto primary user who decided when Wall was active, where it was positioned, and which Mode it was in. Anne's engagement was typically limited to two

forms: [1] viewing and discussing the Last Word displayed, [2] passively interacting via natural speech activities. We both usually spoke freely around Wall with limited conscious awareness of its presence. However, we also limited Wall's placement to more public areas of our home, and often deactivated Wall for lengthy periods of time.

To date, only I have read all of the recorded transcripts in detail. This raises questions (although not concerns, per se) about household consent and transparency. As is often the case with household smart technologies [e.g., 14], one person often takes the lead while other household members lack access or interest in obtaining it.

Display: Magical Moments

The Last Word display led to many humorous and meaningful moments, some of which we shared together and some of which we experienced independently. One of the first meaningful shared moments occurred soon after we adopted Wall. We were having a conversation Sunday morning over breakfast, which is often a time for casually reflecting on our lives and making short and long-term plans. We had been discussing the importance of friendship, establishing roots in an area, and our recent loss of friends and neighbors who had moved away. As we were doing dishes, the word "Community" appeared in the Last Word display.

With an astonished look, Anne pointed to the word silently. (The Last Word display is transient; once a new word is spoken, the Last Word is replaced with the new last word spoken.) We later discussed its significance as a "sign"—a marker of meaning, or a call to action. Here, the surveillant autonomy of Wall was felt to play a role as an external force that we imbued with mild prophetic meaning.

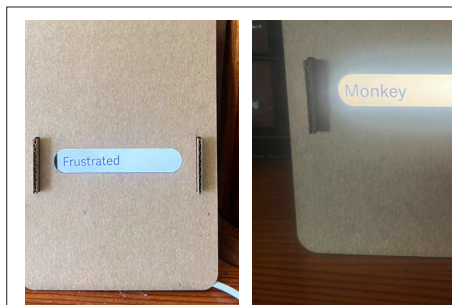
In other cases, the Last Word was less memorably charged with meaning but nonetheless prompted moments of joy, reflection, and meaning-making. I developed a habit of snapping a photo of the display when a particularly meaningful Last Word was displayed—which variously included words produced by me, Anne, the television, and unidentifiable sources.

Breakdowns: Learning to Love Errors

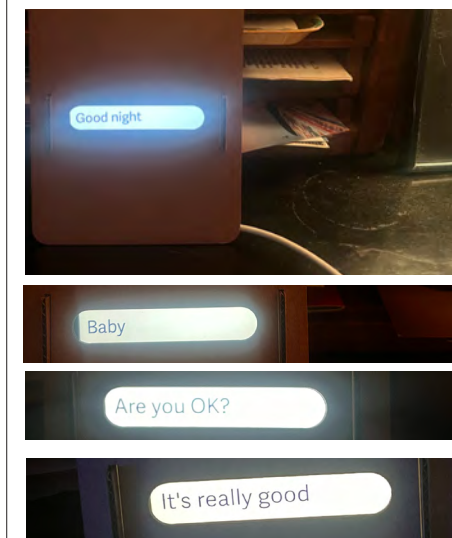
Anyone who reads Wall's display will quickly notice the glaring number of transcription errors, which we estimate can be as high as around 50%. The error rates of current low-cost text-to-speech algorithms are compounded by Wall's lack of advanced omni-directional microphone systems employed in products like Amazon Echo. However, these errors provided an unexpected source of pleasure, meaning, and privacy safeguards that increased our trust and comfort using Wall within the sensitive spaces of our homes. Usually Wall's errors were quite funny. We often felt they enhanced our use of Wall by creating an *impressionistic rather than literal record* of what was said.

When reviewing transcripts, Authors 1, 2, and 3 similarly described an impressionistic experience of recognizing parts of the conversation but not others. We described reading transcripts as similar to a "dream," a "fuzzy memory," an "alien interpretation," or "psychedelic experience." In some cases, it took several minutes to recall the conversation. In other cases, the recall was more immediate. Some conversations could not be recalled at all.

Often the transcripts served as potent memory aids allowing us to remember past events, but also left gaps and uncertainties in our recollection. Some phrases



Meaningful moments captured with Wall. (Right) The word "frustrated" appears in the Last Word display, providing a moment of levity during a tense moment between Author 1 and their spouse. (Left) The nickname for Author 1's cat, "Monkey," was an amusingly frequent occupier of the Last Word display.



Before-bed Last Word captures. When placed in the living room, Wall often captured a Last Word between Author 1 and their spouse, or their cat. [Accuracy varies wildly.]

would appear accurate, others were clearly incorrect but easy to decipher, and the rest were unintelligible but nonetheless provided a sense of conversational duration and intensity.

Importantly, these high error rates served an emergent protective, trust-building function. Our initial concerns about revealing sensitive information within our households or to third-party service-providers quickly dissipated when we realized that we ourselves struggled to decipher many conversations. We found that everything captured by Wall was open to contention and plausible deniability. In effect, Wall created a welcome balance between accurate and revealing extraction, and a privacy-preserving disjunction between our actual speech inputs and Wall's text outputs.

Review: Complicated Reminiscence

Shifting to a different temporal scale, we turn to a story of perusing the transcripts archival. First, it is important to note that due to a variety of intended and unintended factors, only about 20 hours of transcription was available at the time of this event. While this was initially disappointing to me (Author 1), I soon found a new appreciation for how this simplified my focus and enhanced the meaning of the transcripts that were available. One transcript in particular was source of deep significance as it revolved around a very difficult and emotionally painful time period when an immediate family member was undergoing medical treatment, and another was hospitalized with a devastating injury. In fact, I discontinued use of Wall for several months in response to these challenges.

When I reviewed some of the transcripts months later, the family member with a life-threatening illness had been cured, although the other was still struggling and required intensive caregiving. I was struck

by both how much had changed and how much had remained the same, or gotten worse. The transcripts ambiguously conflated conversations with family, me, and Anne. The combined records read like a fever dream that vacillated between vividly accurate dialogue and hazy, inaccurate interpretations. The banality of the conversations was striking, as live conversations are rarely recorded in daily life.

A few meaningful yet partly inaccurate (and thus privacy-preserving) transcriptions are include below:

→ "Processing his emotions in real time.. He doesn't know things like secretary"

→ "It's like a meal like the perfect thing I miss so much emotionally you know"

→ "Reality is actually not reality anymore it's like"

→ "I don't think it's gonna be like suddenly she's like on dust that's not works unless"

On the one hand, I experienced shock and disbelief with how Wall enabled and perhaps compelled me to relive details of a very painful period. On the other hand, the initial discomfort of reliving these events quickly transitioned into an appreciation for the opportunity to do so. I felt that Wall had inadvertently selected an appropriate balance of detail and omission to provide me with an impressionistic mode of reminiscence, and without overburdening me with every literal detail.

EXPERIENTIAL INSIGHTS

Aesthetic Surveillance

We (Authors 1, 2, and 3) found that Wall's

continuous extractive and imposed recording of environmental speech and conversations was both fun and dangerous—and that the danger can be part of the fun.

In general, we identify a need to move beyond positive/negative binaries when considering user experiences of surveillant sense-and-extract systems. Our firsthand experiences of Wall are not easily slotted into the common either-or distinctions used to normatively assess digital surveillance, privacy, and data ethics. Instead, our experiences combined or elided distinctions between good/bad, benefit/harm, consensual/non-consensual, transparent/opaque, and safe/vulnerable.

Upon deeper reflection, we conclude that our positive and negative affective experiences of Wall were not isolated incidents but rather imbricated, reciprocally influencing forces. We refer to this as **aesthetic surveillance**—an affective experience with surveillant technologies that blends or oscillates between positive and negative valences.

Within an aesthetic surveillant experience, a *lack* of consent, transparency, and safety may counterintuitively form a source of positive emotions such as thrill, serendipity, or surprise. Critical media and surveillance scholars have alluded to this dynamic [e.g., 4,11].

Given the polarizing cultural debates over digital privacy and surveillance, the concept of aesthetic surveillance may seem a contradiction in terms. But consider again Author 1's difficult but ultimately rewarding experience of reviewing transcripts collected while dealing with family health issues. A more mundane example is any of the numerous instances where Wall captured a humorous but mildly embarrassing Last Word without first asking for renewed consent.

Yesterday • 03:02PM • 3:39m 30s

Tomorrow, yeah September 4 Wednesday or noon

Oh yeah September 4th Wednesday

44

Or

Can you change it?

She doesn't wanna go

Always

I'm not sensitive

Definitely

Been telling the story when you were a toddler

About to get a shot

Crying

So fat milk, milk

Stay here

I

It's OK

OK, I'm gonna

Party, September

You had to be there: A typical transcript. This conversation between Author 2 and her family represents a typical Wall transcript when running the speech-to-text app Ava.

This text provided a source of meaningful reflection for Author 2, but is unlikely to resonate with a general audience.

The concept of aesthetic surveillance has two uses within HCI and design. First, the concept can be used as an experiential state that researchers may identify within studies of current technologies, or design for via future technologies. Second, we hope the term can help us rethink binaries that have emerged around safety/vulnerability, consent/non-consent, and benefit/harm. There are indeed deeply troubling realities and trends involving digital surveillance [e.g., 11,21,30], and we in no way intend to minimize these harms. At the same time, we find that lived experiences of act-and-extract systems are rarely black and white. The HCI community has designed and documented many cases of surveillance with varying shades of ambivalence by users [e.g., 2,13,14,16].

Aesthetic surveillance crystallizes a more general idea that surveillance is often both good *and* bad. The word surveillance holds negative connotations in most arenas. Within everyday contexts, surveillance refers to close, systematic, and sometimes automated observation or monitoring of people, places, and activities typically for the purposing of managing or influencing them. However, surveillance can also serve protective functions, such as disease surveillance in medicine or crop surveillance in agriculture. The concept of aesthetic surveillance allows us to maintain attention on the negative connotations of surveillance while also acknowledging its many positive and ambivalent uses, and their complex interplay.

Much like the word surveillance, the qualities we identify of sensitization, extraction, imposition, and disjunction are not inherently negative, yet they do always hold the potential for serious harm. These qualities of zone interfaces, such as Wall's, may trigger negative connotations. Yet their actual impacts upon the lived experiences of actual people really does depend on the details of context, and design.

Details Make or Break Acceptability

Wall illustrates how specific user interface details significantly affect user experiences of trust, control, benefits, harms, and more. We find that these interface details, rather than a system's generic surveillance capabilities (e.g., "always-recording"), are the difference between an act-and-extract system that is perceived as dangerous or exploitative, and one that's embraced as useful and desirable.

A key insight from our design and self-use process is that a highly surveillant system can be made acceptable and even desirable with the right user interface details. This insight suggests both opportunities as well as concerns, for it highlights how surveillant companies and products can manipulate users by foregrounding benefits while minimizing exposure to harmful or risky backend processes [e.g.16].

Here we reflect on techniques we employed to design an acceptable and even enjoyable everyday technology whose *raison d'être* is to surveillantly record conversations.

Covering up surveillance. Most of the time, Wall is designed to literally cover up the third-party software systems that power its operation (Zoom, Ava, iOS, and iCloud). Primary users who engage in swapping covers or reviewing transcripts will periodically encounter these backend systems. However, the situation is very different for adjacent users whose engagement is limited to viewing the display screen or may even be unaware of Wall's activity. Such adjacent users are inhibited from accessing or knowing about recorded transcripts or risks associated with Wall's cloud-based third-party apps.

Disarming with charm. The form language of Wall employs inexpensive everyday materials to create a humble and intimate do-it-yourself aesthetic. This

form language was designed to communicate that Wall was not a corporate or conventional R&D product but rather an exploratory, artistically-inflected project. Feedback from colleagues and acquaintances confirmed that Wall's physical and interactive forms disarmed users' suspicions or objections through its charming, whimsical, intriguing, and bespoke personality.

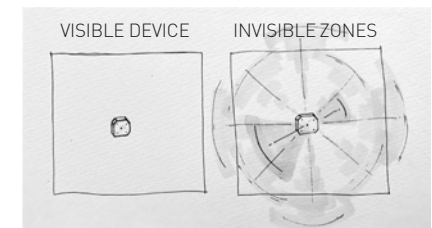
Enlisting a trusted champion user. Dove-tailing with the techniques above, our process of deploying Wall has involved identifying an interested user—namely ourselves or a close friend or family member. In practice, we found that the friends, family, and colleagues we exposed to Wall's surveillant sense-and-extract capabilities rarely objected or expressed concern. In large part, this was because they trusted us. We came to see ourselves as "trusted champion users" who experienced few barriers to configuring a highly surveillant device within sensitive contexts such as our homes and workplaces.

Upshots of Errors & Imperfections

Wall's extractive and imposing capabilities are far less omnipotent than they first appear. In reality, the system is riddled with errors, inconsistencies, and practical limitations that resulted in us being unable to achieve the high levels of surveillance we at times desired, such as accurate, long-duration transcripts.

Sometimes Wall is very accurate, but often it is comically off target. Initially, we felt Wall's error rate was disappointingly high, and that it represented a mismatch with our original vision. However, we soon identified numerous upshots of these imperfections:

→ *Increased privacy/security:* Transcriptions are easily contested and plausibly



The Vision: A device that continuously imposes speech extraction all around.



10:53:53 Cause like, if I'm talking over her
10:53:49 Definitely showing up on my over he
10:53:56 Is that showing up.
10:53:57 That showing up.
10:53:58 It is.
10:53:58 It is fun!
10:54:01 But I can't put it to the.
10:54:03 Yeah, so it's still accurate.
10:54:01 But I can't wait till like who said
10:54:06 Yes! Oh.
10:54:05 Yeah, it's still accurate. I just
10:54:08 But it's

The Reality: An interactive device with disjointed input-output relationships, and fluctuating accuracy and responsiveness.

denied, leading to increased trust and comfort with Wall in sensitive spaces.

→ *Distancing and defamiliarization*: Imperfect transcriptions prompted reflection and deeper engagement.

→ *Dreamy, surreal, and humorous experiences*: Wall's specific, and largely accidental, fluctuations of extractive accuracy led to a distinctive experience that provided levity, dream-like memories, and impressionistic engagements.

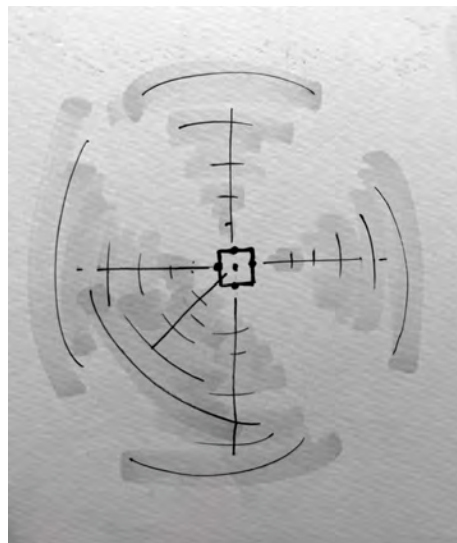
→ *Agentive associations*: At times, we felt as though Wall was inadvertently using judgment and mystical force to decide the appropriate level of accuracy.

We also note that other practical constraints—such as the tedious process for downloading transcripts and the challenges of routing power cables—also created unexpected opportunities for increased privacy and slower, deliberate forms of reflection

THEORETICAL INSIGHTS

We designed, developed, and used Wall to materially and empirically explore concepts we articulated through our conceptual design process (pages 1-2). Our case study of Wall provides an illustration of these concepts, while also allowing us to develop these concepts deeper and at times challenge them. For example, our original vision of Wall presumed it was capable of highly accurate and continuous extraction. In practice, multiple technical and practical limitations resulted in extraction that was quite imperfect.

In conclusion, we return to our core concepts for understanding the design qualities



A formative conceptual diagram of Wall's sensitizing fields, which transform nearby space into an extractive interface.

of surveillant spatial sensing systems, and place these into conversation with Wall.

Human-computer interfaces have expanded from an act-then-react interaction paradigm to a sense-and-extract paradigm. Traditional digital user interfaces follow an *act-then-react* model wherein a user knowingly acts upon a perceptually well-defined interface by touching a recognizable interactive affordance such as button or dial. The computer system then reacts with perceptible output, such as changing pixels on a screen, modulating the speed of a servo motor, or activating sound waves from a speaker. We refer to these as *point interfaces* because user interaction is constrained to clearly defined physical areas, such as buttons on a touch screen, keyboard, or mouse.

Newer technologies equipped with sensors, networked connectivity, and advanced data-

processing capabilities such as AI have paved the way for a new interaction paradigm we refer to as *sense-and-extract*. For example, a smart camera system may detect a person's face, identify their name, and remotely display this information in nearly "real-time" on a screen to a user far away—entirely unbeknownst to the unaware and non-consenting subject. We refer to systems that enact a sense-and-extract paradigm as *zone interfaces*, because they create a fuzzy, spatial zone of interactivity that sharply contrasts with traditional point interfaces.

Through our eccentric design approach, we identified qualities that characterize the sense-and-extract paradigm and help us articulate its opportunities, limitations, and concerns. Our prototype and self-use deployments of Wall enabled us to dig deeper into these four qualities grounded in lived experiences and concrete illustrations.

Sensitization: *Physical spaces are transformed into digital interfaces.* Wall uses a smart microphone system to transform physical space into a sensing interface. Wall creates an auditory spatial sensing zone that functions as an *input field* for interactive computing.

Extraction: *Environments and those within them become exposed to automated digital data-capture and information gathering.* Wall extracts data and information from the environment, primarily from human speech. Wall can continuously extract audio data, filter speech, and translate this data into text captions that appear on its display or are stored locally or remotely "in the cloud."

Imposition: *Extraction is thrust upon the environment and actors within it. Inherent points of consent, control, and awareness disintegrate as tangible buttons are replaced with invisible sensor fields.* Wall can continuously impose its extractive potentials onto its nearby environment. As we experienced, nearby users are likely to overlook or forget Wall and its ongoing imposition of speech extraction. However, the degree of imposition that Wall

exerts upon subjects and object depends heavily on details of its configuration. For example, Open and Last Word display modes offer moderate forms of consent and awareness to nearby users, whereas the Silent mode gives little clues as to Wall's imposed, extractive powers.

Disjunction - *Input/output relations become looser, more uncorrelated, and harder to predict.* Wall is a very loose input/output device. It is not as predictable as a conventional act-then-react interface. Some speech is successfully extracted, while other conversations are missed. Among the speech that is captured, errors are very common. Furthermore, nearby users cannot perceptually ascertain the precise boundaries of the zonal field of sensing and extraction. It forms a fuzzy gradient rather than a clearly delineated button, screen, or keyboard.

The sense-and-extract paradigm creates many new user experience design opportunities, as well as practical limitations and ethical concerns. Articulating differences between conventional point interfaces (act-then-react) and zone interfaces (sense-and-extract) is a key to understanding complex HCI issues with power and agency—including concerns over user consent, control, awareness, trust, benefits, and prevention of harms.

For example, consider recent controversies around facial recognition technologies. These debates can be seen to revolve around the qualities of extraction (facial recognition gathers sensitive personal information), imposition (it is thrust upon users in ways that are often impossible to see), and disjunction (both accurate and inaccurate identifications can have harmful consequences).

We hope that designers and researchers can use and extend this vocabulary across the areas of critical, empirical, and pragmatic design-oriented activities.

References

1. Arvola, Mattias. "Interaction Design Qualities: Theory and Practice." In Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries, 595–98. Reykjavik Iceland: ACM, 2010. <https://doi.org/10.1145/1868914.1868982>.
2. Bernd, Julia, Ruba Abu-Salma, and Alisa Frik. "{Bystanders'} Privacy: The Perspectives of Nannies on Smart Home Surveillance." In 10th USENIX Workshop on Free and Open Communications on the Internet (FOCI 20), 2020. <https://www.usenix.org/conference/foci20/presentation/bernd>.
3. Chatting, David, David S. Kirk, Paulina Yurman, and Jo-Anne Bichard. "Designing for Family Phatic Communication: A Design Critique Approach." In Proceedings of the 2015 British HCI Conference, 175–83. Lincoln Lincolnshire United Kingdom: ACM, 2015. <https://doi.org/10.1145/2783446.2783566>.
4. Chun, Wendy Hui Kyong. Control and Freedom: Power and Paranoia in the Age of Fiber Optics. MIT Press, 2008.
5. Desjardins, Audrey, Oscar Tomico, Andrés Lucero, Marta E. Cecchinato, and Carman Neustaedter. "Introduction to the Special Issue on First-Person Methods in HCI." ACM Transactions on Computer-Human Interaction 28, no. 6 (December 31, 2021): 1–12. <https://doi.org/10.1145/3492342>.
6. Elsdén, Chris, David Chatting, Abigail C. Durrant, Andrew Garbett, Bettina Nissen, John Vines, and David S. Kirk. "On Speculative Enactments." In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, 5386–99. Denver Colorado USA: ACM, 2017. <https://doi.org/10.1145/3025453.3025503>.
7. "Forget Multitasking, Try Monotasking | TED Talk Embed." Accessed January 20, 2025. https://www.ted.com/embed/paolo_cardini_forget_multitasking_try_monotasking.
8. Gaver, Bill, Tony Dunne, and Elena Pacenti. "Design: Cultural Probes." Interactions 6, no. 1 (January 1999): 21–29. <https://doi.org/10.1145/291224.291235>.
9. Gaver, William. "What Should We Expect from Research through Design?" In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 937–46. Austin Texas USA: ACM, 2012. <https://doi.org/10.1145/2207676.2208538>.
10. Gaver, William W., John Bowers, Andrew Boucher, Hans Gellerson, Sarah Pennington, Albrecht Schmidt, Anthony Steed, Nicholas Villars, and Brendan Walker. "The Drift Table: Designing for Ludic Engagement." In CHI '04 Extended Abstracts on Human Factors in Computing Systems, 885–900. Vienna Austria: ACM, 2004. <https://doi.org/10.1145/985921.985947>.
11. Hu, Tung-Hui. A Prehistory of the Cloud. MIT press, 2015.
12. Hutchinson, Hilary, Wendy Mackay, Bo Westerlund, Benjamin B. Bederson, Allison Druin, Catherine Plaisant, Michel Beaudouin-Lafon, et al. "Technology Probes: Inspiring Design for and with Families." In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 17–24. Ft. Lauderdale Florida USA: ACM, 2003. <https://doi.org/10.1145/642611.642616>.
13. H. Tan, Neilly, Richmond Y. Wong, Audrey Desjardins, Sean A. Munson, and James Pierce. "Monitoring Pets, Deterring Intruders, and Casually Spying on Neighbors: Everyday Uses of Smart Home Cameras." In CHI Conference on Human Factors in Computing Systems, 1–25. New Orleans LA USA: ACM, 2022. <https://doi.org/10.1145/3491102.3517617>.
14. Koshy, Vinay, Joon Sung Sung Park, Ti-Chung Cheng, and Karrie Karahalios. "We Just Use What They Give Us": Understanding Passenger User Perspectives in Smart Homes." In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, 1–14. CHI '21. New York, NY, USA: Association for Computing Machinery, 2021. <https://doi.org/10.1145/3411764.3445598>.
15. Lim, Youn-kyung, Erik Stolterman, Heekyoung Jung, and Justin Donaldson. "Interaction Gestalt and the Design of Aesthetic Interactions." In Proceedings of the 2007 Conference on Designing Pleasurable Products and Interfaces, 239–54. Helsinki Finland: ACM, 2007. <https://doi.org/10.1145/1314161.1314183>.
16. Lomborg, Stine, and Kirsten Frandsen. "Self-Tracking as Communication." Information, Communication & Society 19, no. 7 (July 2, 2016): 1015–27. <https://doi.org/10.1080/1369118X.2015.1067710>.
17. Lim, Youn-kyung, Sang-Su Lee, and Kwang-young Lee. "Interactivity Attributes: A New Way of Thinking and Describing Interactivity." In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 105–8. Boston MA USA: ACM, 2009. <https://doi.org/10.1145/1518701.1518719>.
18. Lomborg, Stine, and Kirsten Frandsen. "Self-Tracking as Communication." Information, Communication & Society 19, no. 7 (July 2, 2016): 1015–27. <https://doi.org/10.1080/1369118X.2015.1067710>.
19. Neustaedter, Carman, and Phoebe Sengers. "Autobiographical Design in HCI Research: Designing and Learning through Use-It-Yourself." In Proceedings of the Designing Interactive Systems Conference, 514–23. Newcastle Upon Tyne United Kingdom: ACM, 2012. <https://doi.org/10.1145/2317956.2318034>.
20. Odom, William, Mark Selby, Abigail Sellen, David Kirk, Richard Banks, and Tim Regan. "Photobox: On the Design of a Slow Technology." In Proceedings of the Designing Interactive Systems Conference, 665–68. Newcastle Upon Tyne United Kingdom: ACM, 2012. <https://doi.org/10.1145/2317956.2318055>.

21. O'neil, Cathy. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown, 2017.
22. Pierce, James. "Eccentric Sensing Devices: Using Conceptual Design Notes to Understand Design Opportunities, Limitations, and Concerns Connected to Digital Sensing." In *Creativity and Cognition*, 1–14. Virtual Event Italy: ACM, 2021. <https://doi.org/10.1145/3450741.3466775>.
23. Pierce, James, Richmond Y. Wong, and Nick Merrill. "Sensor Illumination: Exploring Design Qualities and Ethical Implications of Smart Cameras and Image/Video Analytics." In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–19. Honolulu HI USA: ACM, 2020. <https://doi.org/10.1145/3313831.3376347>.
24. Pierce, James, and Eric Paulos. "Counterfunctional Things: Exploring Possibilities in Designing Digital Limitations." In *Proceedings of the 2014 Conference on Designing Interactive Systems*, 375–84. Vancouver BC Canada: ACM, 2014. <https://doi.org/10.1145/2598510.2598522>.
25. Sengers, Phoebe, Kirsten Boehner, Shay David, and Joseph "Jofish" Kaye. "Reflective Design." In *Proceedings of the 4th Decennial Conference on Critical Computing: Between Sense and Sensibility*, 49–58. Aarhus Denmark: ACM, 2005. <https://doi.org/10.1145/1094562.1094569>.
26. Stappers, Pieter Jan, and Elisa Giaccardi. "Research through Design." In *The Encyclopedia of Human-Computer Interaction*, 1–94. The Interaction Design Foundation, 2017. <https://research.tudelft.nl/en/publications/research-through-design-2>.
27. Thompson, Rob. *Manufacturing Processes for Design Professionals*. Thames & Hudson, 2007. <https://books.google.com/books?hl=en&lr=&id=WehEEAAQBAJ&oi=fnd&pg=PP1&dq=rob+thompson+manufacturing+processes&ots=u9Gsb93Kb7&sig=ZOyPRMGYC-jilewRulozWIDTK9I>.
28. Wakkary, R. L., W. T. Odom, Sabrina Hauser, Garnet Hertz, and Henry Lin. "Material Speculation: Actual Artifacts for Critical Inquiry." In *5th Decennial Aarhus Conference on Critical Alternatives August (AA 2015)*, 97–108. Aarhus University, 2016. <https://research.tue.nl/en/publications/material-speculation-actual-artifacts-for-critical-inquiry>.
29. Zimmerman, John, Jodi Forlizzi, and Shelley Evenson. "Research through Design as a Method for Interaction Design Research in HCI." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 493–502. San Jose California USA: ACM, 2007. <https://doi.org/10.1145/1240624.1240704>.
30. Zuboff, Shoshana, Norma Möllers, David Murakami Wood, and David Lyon. "Surveillance Capitalism: An Interview with Shoshana Zuboff." *Surveillance & Society* 17, no. 1/2 (2019): 257–66.