Description of design dimensions and evaluation for Ambient Displays

Morgan Ames, Anind Dey
UC Berkeley Computer Science

Abstract

We first present a set of potential design dimensions for ambient displays, which are ubiquitous computing devices which monitor and display information in a peripheral, non-obtrusive way, and are meant to reduce demand on one's memory and overloaded senses (such as vision). We developed these design dimensions based on our experience in building them and expectations for their future use. Second, we apply the design dimensions to a set of ambient displays that have been built by other groups. Last, we survey existing evaluation techniques in light of our design dimensions, evaluating their efficacy for ambient displays and making suggestions for modifications.

Introduction

Computers are now embedded in our dishwashers, telephones, lamps, and home entertainment systems. They drive our cars, protect our homes, water our lawns, and keep track of our purchases. As processors appear in more and more devices, the paradigms for computer interface design will have to shift to account for the burden of being surrounded by so many computing devices.

Current desktop interfaces assume that they can demand a user's attention at will. They make little distinction between notifications that are important and those that can be put off until later. They often do not show a history of commands or changes, leaving it to the user to remember the state of the system. If a user is interested in knowing when a data source reaches a certain state - a stock breaks 100, or cars on the freeway are traveling over a certain speed - the user must poll that information source; there is no easy way to tell a desktop computer, "notify me when this information reaches this state." Moreover, most computer interfaces are heavily visual, overloading this sense while leaving the rest virtually unused. These systems don't support multi-tasking well; they assume that they are a user's only task.

Imagine these design standards extended to ubiquitous computers. Every appliance, every gadget could demand our attention as if it were the only one doing so. The result would be chaotic and stressful - one can already get a flavor for what this would be like with flashing PDAs, ringing cellular phones, and chiming wristwatches all vying for attention.

Imagine instead that these many devices work together to enable, rather than hinder, multi-tasking. They display information with obtrusiveness relative to their importance: your boss' e-mail will get through immediately, while the final score of that soccer match can wait until you have a free moment. Some devices may display information (time, traffic, weather) constantly, but alert you with a more salient cue when a certain state is reached (like an alarm clock) or when information changes (like traffic noise or ambient light). Events are shown at an appropriate time scale and with an appropriate refresh rate: a few seconds for a phone call, minutes or hours for e-mail, all day for a birthday reminder. If it is important to compare the present with the past (*e.g.*, stock market trends, word processor document changes) or with the future (weather forecasts), the device reduces the need to remember these states by displaying them as well. Similarly, a device can provide an overview of an information source in the periphery, which resolves to an appropriate level of detail when given more than a glance (for example, the ambient light you see through a window gives you an idea of the time of day and the weather, while people's clothing and the movement of trees give you more detailed information). Finally, these devices could take advantage of a sense that is not overloaded with other tasks (*e.g.*, ICQ's sounds indicating the actions of other members, or a car's engine noises).

The field of ambient displays is one of the research fields studying alternatives to this current paradigm of task-oriented design. Ambient displays are ubiquitous computing devices that provide a continuous stream of information in a peripheral, non-obtrusive way. Ambient displays are particularly good at monitoring and displaying in a simple manner the status of a complex system, but can provide us with any information about the world that we do not need or want to directly attend to. Though they constantly display information, ambient displays are meant to be pleasing: though a window tells us time, season, weather, and other information, window offices are coveted. Ambient displays reduce a user’s cognitive load by alerting the user to an “interesting” development, rather than requiring the user to occasionally check the status of an information source. [1][2]

Design Dimensions

The dimensions our group finds important to ambient displays are:

- **Intrusiveness:** displays do not demand attention, but provides information with a level of intrusiveness appropriate to the information's importance
- **Notification:** devices display information constantly, but alert with a more salient cue when a certain state is reached or when information changes - they do not demand the same amount of attention all the time
- **Persistence:** displays show information at an appropriate time scale and an appropriate refresh rate
- **Temporal context:** if comparison with past or prediction of future is important, displays show it, reducing cognitive demands on user by not requiring them to remember other states
- **Overview to detail:** displays show the right amount of detail: get an overview at a glance, and more detail if one pays attention
- **Modality:** displays show information in a mode (that is, using a sense) that is not already overloaded
- **Level of abstraction:** displays do not show information directly, but rather in an abstract or indirect manner. The method of displaying information should be clearly linked to the nature of the information
- **Interactivity:** displays are appropriately interactive (or not), without demanding too much from the user
- **Location:** displays reflect sensitivity to location and their surroundings in general, such as a quiet room vs. a noisy public plaza
- **Content:** displays show information that the user cares about, or are flexible in content
- **Aesthetics:** apart from being useful or valuable as information sources, the displays are also pleasing

The overarching goal of ambient displays is to reduce a user's cognitive load by externalizing memory. They aid in multi-tasking by externalizing the demands many tasks put on one's memory. An ambient display does not provide an answer to the task-oriented query "I want to find out about this" as much as it supports the request of "I want to be notified when this happens". The success of an ambient display is difficult to pin down: it may reduce cognitive load by automatically providing the same information that one had to look up before, or they may change one's awareness of some information that can be measured in a behavior change. These and other factors make them particularly difficult to evaluate with traditional task-oriented evaluation techniques.

Existing Ambient Displays

Several groups have been building ambient displays at various levels of abstraction. Below is a sampling of these ambient displays, described in terms of the dimensions above with specific details provided about any evaluation performed, to give a flavor of the work that is being done in the field.

Ambient Displays from MIT Media Lab

An Interactive Poetic Garden
http://acg.media.mit.edu/projects/stream/index.html
http://acg.media.mit.edu/projects/stream/InteractivePoeticGarden.pdf
Tom White and David Small, Aesthetics and Computation group

The "Stream of Consciousness" Interactive Poetic Garden projects words into a six-foot square pool with a waterfall, allowing words to flow and be stirred up with the water, creating spontaneous and transient poetry. Hands or other objects held above the water repel words and create a blue aura. If a word is directly pressed, it glows, swells, and spawns related words one at a time. Letters in a word behave as if attached by springs - for example, letters undulate as words circulate through the system, and words swirled vigorously "break" and form into new words. Old words are removed.

Design dimensions

- Intrusiveness: somewhat intrusive (peripheral animation with swirling words may be distracting, water gurgling and splashing)
- Notification: none
- Persistence: real-time response; old words are removed eventually (time scale not specified)
- Temporal context: good sense of past interactions, based on words currently present in the pool (words are removed over time, and words are fairly stable when not directly manipulated); no sense of future interactions
- Overview to detail: sound of water to appearance of words in the pool
- Modality: vision (words moving), sound (water splashing)
- Level of abstraction: clear link between the motions of water and words; interactivity is less clear
- Interactivity: very responsive
- Location: appropriate for a semi-public or public display
- Content: the words displayed are flexible; the content is less relevant than relaxing/amusing; no specific information source
- Aesthetics: very pleasing

Evaluation: testimonials and casual observation

Lumitouch
http://tangible.media.mit.edu/projects/LumiTouch/lumitouch.htm

http://tangible.media.mit.edu/papers/LumiTouch_CHI01/LumiTouch_CHI01.pdf
Angela Chang, Brad Koerner, Benjamin Resner, XingChen Wang, and Professor Hiroshi Ishii

A Lumitouch picture frame glows when its counterpart frame's touch sensors are activated. Different colors are displayed depending on length, location, and intensity of the touch. Motion is also monitored with an infrared sensor and displayed as a glow at the bottom of the remote frame. When a frame is sending information, it glows along the top.

Design dimensions

- Intrusiveness: minimal with just glow; possible peripheral animation
- Notification: glow or sequence of colors
- Persistence: granularity of seconds; glows for as long as the touch lasts or as long as motion is detected
- Temporal context: none
- Overview to detail: overview is a glow, detail could be the picture in the frame or the particular sequence of colors
- Modality: visual (tactile for transmission)
- Level of abstraction: fairly abstract; user needs to know what glowing means for the display to have meaning (some users develop various codes for different color combinations)
- Interactivity: very interactive - remotely responds to touch and motion
- Location: appropriate as a personal display, as it is meant to foster person-to-person contact; less appropriate as a semi-public or public display
- Content: fixed content (glow responds to touch only); content is presumably important to user; content focuses on connecting people (rather than reducing cognitive load)
- Aesthetics: somewhat pleasing

Evaluation: preliminary testing with random pairs of users

Pinwheels
http://tangible.media.mit.edu/projects/pinwheels/pinwheels.htm
Sandia Ren, Phil Frei, Seye Ojumu, Rujira Hongladaromp and Professor Hiroshi Ishii

One or more pinwheels spin in response to an information source, such as the stock market, web site hits, wind movement, server packet types/destinations, movement of people, et cetera. Different configurations of pinwheels can be used for different effects.

Design dimensions

- Intrusiveness: somewhat intrusive (peripheral animation, breeze of cold or dry air)
- Notification: changing rate of spinning; may be difficult to discern
- Persistence: granularity of seconds or less; spin continuously, and at different speeds depending on state of information source (so current state of information is always accessible)
- Temporal context: very little (only the air movement in the room)
- Overview to detail: minimal, unless the configuration of pinwheels communicates some information
- Modality: vision, touch (feeling of wind and coolness)
- Level of abstraction: abstract, unless the configuration of pinwheels communicates some information
- Interactivity: not interactive, unless the information source involves the user (for example, the pinwheels spin more in response to movement in the room)

- Location: appropriate for private, or semi-public displays, not as good for public displays (abstract)
- Content: flexible content; can be people-focused or information-focused
- Aesthetics: pleasing enough (motion or airflow could be irritating)

Evaluation: personal testimonials from museumgoers in Tokyo, some observation

AmbientROOM/Ambient Fixtures
http://tangible.media.mit.edu/projects/ambientROOM/ambientROOM.htm
http://tangible.media.mit.edu/projects/Ambient_Fixtures/Ambient_Fixtures.htm
http://tangible.media.mit.edu/papers/Tangible_Bits_CHI97/Tangible_Bits_CHI97.pdf
http://tangible.media.mit.edu/papers/Ambient_Disp_CoBuild98/Ambient_Disp_CoBuild98.pdf
Professor Hiroshi Ishii, Craig Wisneski, Matt Gorbet, Scott Brave, Andrew Dahley, Brygg Ullmer, and Paul Yarin

A variety of ambient information sources, described below, are brought together into one cubicle-like room with a workstation. Small bottles "contain" other information to be displayed, and a large clock will make the room display events of the past or future when its hands are rotated.

1) The water lamp projects light through a pan of water onto the ceiling. The water is agitated in response to information (such as the activity of the lab's hamster), throwing ripple shadows on the ceiling.

Design dimensions
- Intrusiveness: somewhat intrusive (peripheral animation)
- Notification: ripple shadows
- Persistence: best for refresh rates of minutes to hours (shorter timescales would result in overlapping ripples); information fades quickly
- Temporal context: sense of immediate past as ripples die down, no sense of future; further past or future can be accessed through the manipulable clock described above
- Overview to detail: possible, if information was encoded in the type or intensity of the ripples
- Modality: visual; possibly aural (splashing water)
- Level of abstraction: very abstract; user would have to know information source
- Interactivity: not interactive
- Location: appropriate for private or semi-public displays; difficult for public displays (abstract)
- Content: flexible content
- Aesthetics: somewhat pleasing

2) Patches of light projected on the wall of the room move when people outside of the room are moving.

Design dimensions
- Intrusiveness: somewhat intrusive (peripheral animation)
- Notification: movement of patches
- Persistence: real-time; patches last for as long as a person is moving outside the room
- Temporal context: none by itself; past or future can be accessed through the manipulable clock
- Overview to detail: peripheral animation to location of spot
- Modality: visual
- Level of abstraction: somewhat abstract; can be learned through observation

- Interactivity: not interactive
- Location: good for a private, semi-private, or public display
- Content: fixed content
- Aesthetics: somewhat pleasing

3) The scratching of pens on the group whiteboard is transmitted into the room.

Design dimensions
- Intrusiveness: not very intrusive (could be, if one is sensitive to sound)
- Notification: scratching sound
- Persistence: real-time sounds
- Temporal context: none by itself; past or future can be accessed through the manipulable clock
- Overview to detail: none
- Modality: aural
- Level of abstraction: fairly clear link between scratching/squeaking sounds and whiteboard pens; less clear link to particular group whiteboard
- Interactivity: not interactive
- Location: good for low-noise private or public settings
- Content: fixed
- Aesthetics: not applicable

4) The room's light and a soundtrack of birds and rainfall changes with "approximate" information such as unread e-mail messages.

Design dimensions
- Intrusiveness: not very intrusive (subtle lighting changes, sound changes)
- Notification: change in light levels or volume of sound
- Persistence: best for gradual refreshing; information is continuously displayed
- Temporal context: none by itself; past or future can be accessed through the manipulable clock
- Overview to detail: none
- Modality: visual, aural
- Level of abstraction: very abstract; user would need to know information source
- Interactivity: somewhat interactive, depending on information (e.g. if the system is monitoring number of unread e-mails, reading e-mails will make the system respond)
- Location: good for a quiet private or semi-public display, not good for a public display (abstract)
- Content: flexible content
- Aesthetics: somewhat aesthetic

Evaluation: observation of people in ambient room, experimentation for the background-to-foreground threshold for individual displays (no details provided)

Personal ambient displays
http://tangible.media.mit.edu/projects/Personal_Ambient_Disp/Personal_Ambient_Disp.htm
[no paper available]
Craig Wisneski, Professor Hiroshi Ishii, Will Logan, Jeff Steinheider, Blair Dunn

A small device such as an accessory (jewelry, watch) or pocket keychain/trinket conveys information through heat, movement/vibration, or other tactile modes.

Design dimensions
- Intrusiveness: somewhat intrusive, depending on tactile mode
- Notification: yes (any response from the object is a notification)
- Persistence: lingering heat if information is displayed that way; otherwise, information is displayed for the length of the data input only
- Temporal context: only lingering heat, if any
- Overview to detail: none
- Modality: touch
- Level of abstraction: fairly abstract - user must know what information the device conveys for it to be meaningful
- Interactivity: not interactive, unless the information source involves the user
- Location: private, portable display
- Content: flexible content; more specific information was not specified
- Aesthetics: variable; jewelry and trinkets can be quite aesthetic

Evaluation: none known

Musicbox
http://tangible.media.mit.edu/projects/musicbox/musicbox.htm
[no paper available]
Brygg Ullmer and Professor Hiroshi Ishii

This device is a wooden box that glows and plays music in response to light, movement, and live music around a remote piano. One can also browse through past songs played by the Musicbox.

Design dimensions
- Intrusiveness: somewhat intrusive
- Notification: constantly running; plays music when piano starts
- Persistence: real-time
- Temporal context: songs are saved; the rest (movement, light) is discarded
- Overview to detail: not known
- Modality: light, sound
- Level of abstraction: songs and light levels directly map to song and light levels at remote location; however, it's not clear what drives the system
- Interactivity: not interactive
- Location: appropriate for private, semi-public, or public settings
- Content: content is fixed, and is meant to give a sense of a remote place
- Aesthetics: somewhat aesthetic

Evaluation: none known

Ambient Displays from Georgia Institute of Technology

Digital Family Portrait
http://www.cc.gatech.edu/fce/ecl/projects/dfp/index.html
http://www.cc.gatech.edu/fce/ecl/projects/dfp/pubs/dfp-chi2001.pdf
James Rowan, Beth Mynatt

A picture frame displays the state of a remote loved one over time through the position and size of icons around the frame. The icons can be four sizes and can change their distance to the (digital) picture in the frame. Each icon represents a day, with the current day white. Touching an icon brings up more information about that day in the place of the picture. Trends are shown by dots around the outside of the frame, and "alarms" are shown by the position of icons. The data measured, which may be split between multiple frames with different icons, is meant to be what a person living nearby would naturally notice: health (sleep, diet, activity level), environment (weather, repair of house), and social interactions (in person, by phone, in letters), and would ideally be gathered by infrared, sonar, tactile, and radio sensors; cameras; microphones; utilities metering; and appliance monitoring.

Design dimensions
- Intrusiveness: non-intrusive
- Notification: none
- Persistence: updated once a day
- Temporal context: good sense of past: icons for the past 4 weeks are displayed, one week per side of the frame; trends can give a sense of possible future
- Overview to detail: touching an icon brings up more detailed information about what was sensed
- Modality: vision
- Level of abstraction: somewhat abstract; the link between the icons and the person in the picture can be inferred, but the user needs to know what the icons represent
- Interactivity: touch icon for details
- Location: appropriate for a personal or semi-public display; targets a home setting
- Content: health, environment, or social interactions can be displayed; icons are chosen
- Aesthetics: fairly pleasing

Evaluation: A design inquiry began with interviews of family counselors at assisted living facilities to address the issues elderly people and their families face, and interviews with families to address sensing, privacy, and handling aging in general. The group iterated on their design while continuing interviews, conducted field trials were conducted on the first prototype, and distributed surveys on appropriate icons and icon size discrimination.

Office of the Future peripheral displays
http://www.cc.gatech.edu/fce/ecl/projects/kimura/index.html
http://www.cc.gatech.edu/fce/ecl/publications/kimura-uist2001.pdf
Dr. Elizabeth Mynatt, Dr. Blair MacIntyre, Dr. Gregory Corso, Stephen Voida, Michael Terry; Research assistants: Ron Barbas, Amanda Lyons; Past Researchers: Klaus Marius Hansen, Joe Tullio

"Kimura" projects different "working contexts" (a set documents and communications for one task, such as a project or conference) on a whiteboard in an office as images from activity logs, to give the user a sense of past work. It keeps track of outstanding e-mail messages or print jobs relating to each working context, as well as the relative importance of different documents, and allows users to rearrange and write notes to different contexts. If a user selects one of these working contexts from the board, their computer screen displays the last state of that context, and the projection displays some of the history of the user's actions, again gleaned from activity logs, close to the screen.

The group is also experimenting with how humans can detect changes peripherally.

Design dimensions

- Intrusiveness: non-intrusive
- Notification: printer/e-mail icons pop up with pending documents or unread messages
- Persistence: Past actions are perusable, and sometimes displayed as one of the images in the working context
- Temporal context: good sense of past, through screenshots of all documents in a working group
- Overview to detail: abrupt change from a peripheral working context to desktop
- Modality: visual
- Level of abstraction: mapping between collection of pictures and task is clear; possible interactions must be learned
- Interactivity: working contexts can be selected, annotated, or moved; it's not clear whether their collections of pictures can be modified
- Location: appropriate for a private office display
- Content: fixed content
- Aesthetics: not very aesthetic; not addressed

Evaluation: none known; design was inspired by a number of other more-evaluated systems

Semipublic Displays
http://www.cc.gatech.edu/fce/ecl/projects/semipublic/index.html
http://spam.cc.gatech.edu:8080/ecl/uploads/86/Ubicomp2002Final.pdf
Elaine M. Huang, Elizabeth Mynatt

Semipublic displays make visible the activities, outstanding requests, and collective interests of an already-close-knit group of people. Peripheral displays are especially relevant to a close-knit group because they share context. The SMART Board displays four quadrants of information: requests and notes, collaborative space, a group portrait whose people fade out when they don't spend time in the lab, and flowers whose petals display planned attendance of a seminar or event.

Design dimensions

- Intrusiveness: nonintrusive; little peripheral animation (scrolling)
- Notification: requests and collaborations are scrolled through
- Persistence: refreshed with weekly status reports; requests and collaborations are cycled through, each shown for a few minutes; portraits fade to white over a few days when a member isn't in the lab
- Temporal context: past week is displayed by requests, collaborations, and group portrait; future events are displayed by attendance flowers
- Overview to detail: little overview to detail - group portrait and flowers are meant to give an overview with no detail available, while requests and collaborations are detailed without an overview
- Modality: visual
- Level of abstraction: clear link for requests and collaborations, minimal learning needed group portrait and attendance flowers
- Interactivity: collaborative space and attendance flowers are directly interactive; others are indirectly interactive, through status reports and presence
- Location: designed for a close-knit group
- Content: fixed content; content is relevant to a closely-collaborating work group

• Aesthetics: somewhat pleasing display

Evaluation: observation of group discussion generated by reminders; planned surveys, interviews, and logs of e-mail status reports and system usage

Design and Evaluation Techniques

While interface design and evaluation techniques are well-developed and backed by years of research for task-oriented, visual interfaces, there has been little work done on evaluating ubiquitous computing devices and even less on ambient displays, which are, by design, not task-oriented. Some have debated the efficacy of these traditional techniques for non-traditional computing devices, or have modified these techniques to address distributed computing, non-visual interfaces, or non-task-oriented devices [12]. We consider an evaluation technique to be effective if it addresses the design guidelines described above, though we expect that both the evaluation techniques and the guidelines will evolve over time. We will summarize these major design inquiry and evaluation techniques, noting how suited they are to ambient displays: interviews and surveys, contextual inquiry, ethnography, heuristic evaluation, cognitive walkthrough, GOMS, various kinds of user studies, and Scandinavian Design.

Design Inquiry

Interviews and Surveys

The best-known and most-used design inquiry techniques are in-person interviews and surveys. A study participant is given ten to ninety minutes of questions, and then usually asked for comments, either by a questioner or on a form.

Interviews are relatively easy to perform, but they have a few drawbacks that are especially crucial to ambient displays. Much depends on the quality of the questions: the interviewer must carefully plan questions that stay on-topic but do not slant the participant's answers. Participants feel like they are at the mercy of the interviewer: they usually offer only enough information to answer the question, and no more, so if questions are not well-formed, the interviewer will be disappointed after hours of work [13]. Moreover, interviews are usually conducted in a place removed from the participant's usual settings and tasks, so if a designer plans to design a display for one of these settings, they will miss important details of the participant's work habits that are important for any design inquiry and especially important for ambient display inquiry. If asked, for example, what time she takes lunch, a participant may answer, "usually 12:30," without specifying that she takes lunch at 12:30 Monday through Wednesday, meets a friend at 12 on Thursday - unless the friend calls and reschedules, which happens half the time - and goes to a company lunch meeting at 11:30 on Fridays.

One place interviews may be more appropriate is to gauge reactions to ambient display ideas the designers already have. Interviewers can get quick feedback on a range of ideas, and use other techniques to refine the list further or evaluate the displays of this culled list as they are developed. There is a danger of participants only having a faint conception of how they would actually interact with the display - there is a good chance of false positives and missed problems.

While some ambient displays are designed for a specific, closed audience, others are public displays meant for anyone who happens by. Of all the design inquiry techniques, interviews and surveys are most appropriate for dealing with the latter situation: they can reach a large audience, perhaps even *in* the target setting; they call for casual interaction, rather than the more intense "master-apprentice" interaction of contextual inquiry or ethnography; and they can be combined with observational data to produce a better picture of the participant's "environment".

Interviews and surveys can be fast and easy, but they may miss important details of participants' tasks. Many of their limitations for ambient displays are limitations for any interface. Interviews and surveys could be used for ambient displays and are more apt for public displays.

Contextual Inquiry

Contextual inquiry begins with an interview and an overview of techniques, and then goes into three hours or more of shadowing, where the designer is a nosy "apprentice" and a participant is the "master." Designers do not necessarily begin the inquiry with preformed design ideas - in fact, the fewer preconceptions the designer has about the participant's work environment, the better - but they mention design ideas as they come up, gauging the participant's reaction and especially watching for non-committal responses ("Umm, maybe", "yes, I suppose", etc.) or anything but immediate enthusiasm. Designers must carefully watch the participant, and write nothing off to chance or idiosyncrasy. At the end of the shadowing period, the designer conducts a closing interview, summarizing findings and design ideas with the participant and discussing future development. [14]

Contextual inquiry, with its focus on the details of a participant's task and short time-span, is well suited to most "office displays", "personal displays", and other ambient devices for a closed set of people. The designer gets a fairly thorough understanding of the user's work environment in terms of common tasks and superficial organization in usually a day. Participants, put in a position of authority and working in their normal environment, are more inclined to give detailed descriptions of their tasks.

However, a contextual inquiry should be conducted by a designer who is experienced in both inquiry techniques and ambient display design, and appropriate users must be found and convinced to participate in the contextual inquiry. Participants must be convinced to "act naturally" while being scrutinized, which is difficult, but especially important to ambient displays. Finally, a contextual inquiry would not be suited for a display that relies on, or possibly modifies, a group's social strata - it is difficult to gather enough data on a group's social and organizational issues in one day of observation.

Contextual inquiry is a good choice for ambient displays meant for a closed set of people, and is relatively fast and easy to carry out by an experienced designer.

Ethnography

This in-depth study is conducted like a weeks-long contextual inquiry by ethnography specialists, often psychologists or sociologists. They begin with no preconceptions of what will be built and shadow a person or group incessantly, writing, taping, and video recording everything from meetings to lunch break. For example, Victoria Bellotti from PARC will go through people's trash to find out more about a work environment. Eventually the ethnographers fade into the background, and participants do not have to be coached to act naturally. At the end of the study, the ethnographers present their findings to the rest of the design team. [15]

Ethnography provides a much richer source of information than a contextual inquiry, especially about a group's social and organizational structure, which would be excellent for group ambient displays. It also provides more details about a participant's everyday actions, especially as the participants become accustomed to having the ethnographers watching. As in a contextual inquiry, the ethnography is on the participant's terms and in the participant's work environment, which inclines a participant to give more details about tasks.

These studies, however, require a huge time commitment, and good ethnographers charge a lot for their services. Ethnographies are usually conducted by large companies wanting to install or replace a major software product, and it may be difficult to find people who agree to participate in such a study for an ambient display. Moreover, Scandinavian Design techniques show us that it's beneficial to have designers directly involved in the evaluation process, rather than getting information from a third party - they gain a better understanding of a participant's environment, and have a better idea of what interfaces would be successful.

Ethnographies - a rich data source, but expensive and time-consuming - seem beyond the scope of most ambient displays except the ones that need detailed information on a group's social structure.

Evaluation techniques

Heuristic Evaluation

Heuristic evaluation, developed by Jakob Nielsen, is perhaps the easiest evaluation technique reviewed here. A number of designers familiar with heuristic evaluation (Nielsen recommends three to five experts or ten to fifteen novices [15) step through an interface in pre-defined scenarios, applying Nielsen's 10 "usability heuristics" to the interface with the target user in mind. [17] These ten heuristics, recently updated, are:

- Visibility of system status
- Match between system and real world
- User control and freedom
- Consistency and standards
- Error prevention
- Recognition rather than recall
- Flexibility and efficiency of use
- Aesthetic and minimalist design
- Help users recognize, diagnose, and recover from errors
- Help and documentation

The nature of these heuristics changes between desktop applications, web applications, and embedded applications, and some are more applicable than others in a specific situation. The ones most important to ambient displays are:

• Visibility of system status
The purpose of an ambient display is to make the state of any system more visible and accessible. The status-event model that ambient displays exploit to reduce cognitive load focuses on making any system change visible in a time-frame appropriate to the time-frame of the event.

• Recognition rather than recall
Ambient displays are meant to reduce cognitive load, which is impaired when users must remember what states changes in the display mean.

• Match between system and real world
While some displays are quite abstract, they are designed to be easily learnable, which can be facilitated by a clear, understandable mapping between the display and the world.

• Error prevention
The errors that an ambient display should prevent are errors in interpretation.

• Aesthetic and minimalist design
As both a piece of art and a source of information, ambient displays should certainly be aesthetic. They need to be easily learnable, which a minimalist design facilitates.

The following heuristics aren't as applicable to ambient displays:

- Help and documentation
- User control and freedom (not important for fixed-information displays)
- Consistency and standards - lower cognitive load
- Help users recognize, diagnose, and recover from errors
- Flexibility and efficiency of use

Possible additions include:
- Is the display better than what it replaces?
- Is the information displayed relevant to its users?
- How intrusive is the display?
- Does the display's event time scale match the information?
- How learnable is the display?

Heuristic evaluation, though easy to perform and often useful for traditional interfaces, may not be suited to ambient displays. It would be difficult to define "tasks" for a participant, as there is no direct interaction with the display. Also, if evaluators are much different from the target audience of a display, they will report "problems" that wouldn't be issues with the target audience ("false positives") and miss other problems entirely.

<u>Cognitive Walkthrough</u>

In a cognitive walkthrough, as in a code walkthrough, a designer gives a peer evaluator familiar with interface design and ambient displays a detailed walkthrough of the planned interface, given in writing or drawings. The evaluator logs all the places that the interface would be especially good or bad, based on a set of heuristic-like facts about interface use. [18]

There is no need to find target users in a cognitive walkthrough. However, if the evaluators and users are too different, the results of the walkthrough, like the results of a heuristic evaluation, may be full of false positives and missed problems. Moreover, ambient displays are not meant to be the focus of one's attention, and evaluators may have difficulty assessing the obtrusiveness of the display. Moreover, this technique would be especially bad for public or shared ambient displays, as it does not support the evaluation of group effects.

<u>GOMS</u>

GOMS summarizes its components: Goals, Operators, Methods, and Selection rules. A user's high-level goals are often called "tasks" in other evaluation techniques, and can be discovered through a contextual inquiry, interview, or survey followed by a task analysis. These goals are broken into subgoals, which in turn are broken down until every subgoal is an "atomic" action (like a keystroke or mouse click) or operator. GOMS handbooks provide times for each of these operators. Methods are sufficiently habitual, automatic groups of atomic actions that speed up use, and when there is more than one method for a particular goal or when it's not clear if there is a suitable method, one must use selection rules to decide which method to use. The purpose of GOMS is to specify these four components to provide a way to analyze a system: it shows that an interface provides a method for every goal, that the operators are ordered to support goals, and that the time it takes to complete a goal is "reasonable."

Tasks that can be best analyzed with GOMS have these characteristics: directed goals or subgoals (that is, concrete tasks that a user focuses on), routine and well-learned components, and sequential actions (except for CPM-GOMS, summarized below). GOMS is good for either passive (user-controlled) or active systems, though there may be fewer goals (and thus fewer operators, methods, and selection rules) in active systems. Since GOMS predicts user response time, it provides no way to evaluate learnability, account for environmental limitations and prior knowledge, or measure errors. However, even casual-use systems need efficient methods that make them fast and easy to use, and a common goal in systems is "recover-from-error," which does come under the scope of GOMS. [19]

The four primary types of GOMS are:

- CMN-GOMS, which focuses on creating a hierarchy of goals and specifying methods, operators, and selection rules from that.
- KLM-GOMS (Keystroke-Level Model), which is a simplified version of CMN which uses only "keystroke-level" operators and no goals, methods, or selection rules.
- N-GOMS-L, which gives a more specific procedure for identifying GOMS components, expressed like a programming language. It includes advice for the number of steps in a method, how goals are set and terminated, and how to estimate cognitive load.
- CPM-GOMS is good for analyzing parallel activities, using Cognitive, Perceptual, and Motor operators in a "Critical-Path Method" schedule chart (often known as a PERT chart or timeline).

The primary goal of an ambient display is to keep a user informed about some information. In GOMS analysis, this goal can be broken into both general subgoals involving the user's cognitive processes, and specific subgoals relating to the mode of finding information. GOMS operators for ambient displays, the atomic actions of these goals, are mostly cognitive and can be described by the dimensions of ambient displays; for example, an appropriate level of detail, an appropriate refresh rate, clear link between display and information, and, of course, the correct content. The GOMS methods would be primarily cognitive or perceptual habits, and would thus be hard to describe, as would selection rules. Ambient design issues are generally high-level and displays often don't include interactive components; in fact, there may be no keystrokes to measure in an ambient display. Thus, the types of GOMS that provide cognitive and perceptual operators - namely, CPM-GOMS, and to a lesser extent N-GOMS-L - are suited to ambient displays, while KLM-GOMS and the "vanilla" CPM-GOMS would not be as well-suited.

<u>User Studies</u>

There are several types of user studies, a few of which are outlined below. In general, user studies are not considered a "discount" evaluation technique as heuristic evaluation, cognitive walkthroughs, and GOMS are, because evaluators must obtain permission for studies, find and work with users, and process large amounts of data. They are often appropriate for one or a few evaluations of many in an interface's design cycle. Questions that evaluators should strive to answer for all types of user studies of ambient displays are:

- Do people want/need to know the information the display provides? How often do they want to look? How important is it relative to other things they're doing?
- Is the display easier to monitor and understand than the information-gathering technique it's replacing? Does it reduce cognitive load, does it improve the "quality" of interaction, and is it easy to map observation to understanding?
- Does the display show information at the right level of detail - are people likely to want to know more detail if they notice a particular state and look it up elsewhere, or is the amount of detail provided overwhelming?
- How important are past points of comparison or future predictions - does the display show enough of these, or too much?
- Is the display too obtrusive? Is it too peripheral to get people's attention when needed?

<u>User Studies - direct observation</u>

Participants are invited in to "use" a display and are often given sample tasks to complete with it, and are observed by a number of evaluators and possibly audio- or videotaped. Depending on the focus of the user study (*i.e.*, whether the evaluators are getting feedback on a preliminary design or timing a design near completion), participants may be encouraged to "think out loud" for the benefit of the evaluators. This can be done with anything from a low-fi prototype (for a project in the design stage) to an almost-complete interface. Evaluators note participants' emotions, exclamations, facial expressions, and other "qualitative" data, and take note of quantitative data such as time to complete a task or number of errors.

For ambient displays, this technique would be most effective if we could simply install the display in its target environment and log all interactions with it. "Thinking out loud" isn't very appropriate for ambient displays, as it too easily upsets people's concentration and awareness of peripheral information. "Tasks" should be normal tasks that aren't necessarily focused on the display.

These studies are especially good for evaluating working prototypes near the end of the design phase, which *can* be simply installed and monitored. Observing is time-consuming and often can only be done sporadically, and getting permission to watch interactions with the display can be difficult. Also, multiple observers mean multiple interpretations of what an "interaction" is, and there is a burden on the observer to catch everything possible. However, off all the evaluation techniques surveyed, this is the only way to catch natural interactions with the display.

<u>User Studies - automated observation</u>

Automated observation eases the burden on observers, though it is often more time-consuming to complete - and harder to get approved - if it involves analysis of videotape or audiotape. However, other automated evaluation techniques, such as pop-up screens, screen shots, or time logging - can automate information collection and analysis. The main drawback of these is that they miss the details of interactions that are important to evaluating ambient displays.

<u>User Studies - surveys</u>

Surveys are a convenient way to get feedback about an ambient display: it is relatively easy to convince large numbers of people to fill them out, it is easy to collate data (especially multiple-choice questions), and they are good for indirectly testing the learning effects of a display. However, as with design inquiry surveys, people tend to generalize and forget details when filling out a survey, and the value of what is collected depends on the quality of the questions: results are easily biased by leading questions, and how much data is collected depends on how thorough the questions are.

Surveys, as above, are good for public displays that have a large audience.

<u>User studies - controlled experiment</u>

In a controlled experiment, the evaluator identifies independent variables, dependent variables, and possible confounding variables. The evaluator then tries to limit as many confounding variables as possible, and varies independent variables to test their effect on dependent variables. One example of a controlled experiment for ambient displays takes the general form of a beeper study, and is as follows:

> Study participants are invited into a room where they have to perform a mundane task such as data entry (for displays that will be in an office) or ferrying things around (perhaps for displays in doorways or other places of transition). One or more ambient displays are displaying the status of various systems. The participant's actions are recorded during the tasks (including eyes wandering to the display), and the participant is interrupted from their tasks occasionally and asked the status of the information that is being displayed by the ambient display.

This may be especially good for testing learning, especially if we want a display to be immediately learnable. It is also nice to be able to control conditions to better pin down what change results in what reaction. One concern is that an experiment like this distorts a display's effectiveness by putting the experiment participant in a strange setting where they may be more or less vigilant than usual. Moreover, it's tempting to distort the information being displayed by the ambient display, to force it to change more often, for example. This may make the display more noticeable in this setting, while in a normal setting it would be ignored.

Design and Evaluation techniques

Scandinavian Design

Scandinavian Design or Cooperative Design gives a framework for involving users in an entire design process, through contextual inquiry in the design stage to intensive "cooperative design workshops" in the evaluation stage. The contextual inquiry starts with a general overview of the users' jobs and the designers' goals. Designers then "shadow" users like an apprentice, asking for demonstrations and clarifications about their jobs, and collecting as many artifacts - notes, drawings, forms - as they could to get an overview of the existing systems and the users' environment. During this time users also learn about the possibilities of technology through seminars and discussions. [20]

Designers create current scenarios and then develop possible future scenarios, creating product prototypes (low-fidelity in the first iterations, increasingly complex in subsequent iterations) that support these scenarios. All user involvement from this point is in cooperative design workshops. Kyng recommends workshops first have an introduction of the prototypes; then a series of sessions where users individually rotate through each of the prototypes, running through scenarios and discussing problems and possible improvements; then a wrap-up. After a workshop, the developers who attended explain the results (with the help of notes and video) to the rest of the design team. Kyng suggests that the developers write a report with comments labeled as 'critique,' 'positive,' and 'proposals,' and prioritize the list of changes. Designers should also write reports about the current prototype as compared to the proposals from the workshop throughout development.

Kyng stresses the importance of choosing "ordinary users" (the user group doesn't have to be perfect, but "adequate enough" to make meaningful contributions) to help in the design process and of involving the designers themselves in the inquiry and workshops, rather than a third party that simply presents findings to the designers. He also gives suggestions for keeping the users' interest (compensation, hands-on work with prototype) and the designers' commitment (knowledge of user-centered design, learning by apprenticeship).

These design techniques, though developed for more traditional goal-oriented and visual interfaces, may prove to be effective in evaluating ambient displays. They must be tested and possibly modified to prove their efficacy. The descriptions above, the ambient display design dimensions, and the examples of ambient displays described in terms of the design dimensions are meant to be a basis and point of departure for further inquiry and experimentation.

References

[1] Wisneski, C., H. Ishii, A. Dahley, M. Gorbet, S. Brave, B. Ullmer and P. Yarin. “Ambient Displays: Turning Architectural Space into an Interface between People and Digital Information.” In *Proceedings of the First International Workshop on Cooperative Buildings, 1998.*

[2] Dix, A., J. Finlay, G. Abowd, R. Beale. “Status/Event Analysis.” Chapter 9 of *Human-Computer*

Interaction, 2nd edition, pp. 365-373

[3] White, T., and D. Small. “An Interactive Poetic Garden.” In *Extended Abstract of CHI 1998*, ACM.

[4] Chang, A., B. Resner, B. Koerner, X. Wang, H. Ishii. “Lumitouch: an Emotional Communication Device.” In *Proceedings of CHI 2001,* ACM.

[5] Hiroshi Ishii, Sandia Ren, and Phil Frei. “Pinwheels: Visualizing Information Flow in an Architectural Space.” In *Extended Abstract of CHI 2001*, ACM.

[6] Ishii, H. and B. Ullmer. “Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms.” In *Proceedings of CHI 1997,* ACM.

[7] Wisneski, Craig, H. Ishii, W. Logan, J. Steinheider, B. Dunn. “Personal ambient displays.” http://tangible.media.mit.edu/projects/Personal_Ambient_Disp/Personal_Ambient_Disp.htm

[8] Brygg Ullmer and Hiroshi Ishii. “Musicbox.” http://tangible.media.mit.edu/projects/musicbox/musicbox.htm

[9] Mynatt, Elizabeth, Jim Rowan, Annie Jacobs and Sarah Craighill. “Digital Family Portraits: Supporting Peace of Mind for Extended Family Members.” In *Proceedings of CHI 2001,* ACM.

[10] Hansen, K.M., MacIntyre, B., Mynatt, E.D., Tullio, J., and Voida, S. "Hypermedia in the kimura system: Using spatial, temporal, and navigational relationships to support multitasking and background awareness." Student Poster Presentation at the ACM Conference on Hypertext and Hypermedia.

[11] Huang, E., and E. Mynatt. “Semi-Public Displays to Support Collaboration and Awareness in Small Group Environments.”

[12] Nicole Yankelovich, Gina-Anne Levow, Matt Marx. “Designing SpeechActs: Issues in Speech User Interfaces.” In *Proceedings of CHI 1995,* ACM.

[13] Dix, A., J. Finlay, G. Abowd, R. Beale. “Evaluation Techniques.” Chapter 11 of *Human-Computer Interaction, 2nd edition*, pp. 405-442.

[14] Beyer, Hugh, and Karen Holtzblatt. *Contextual Design: A Customer-Centered Approach to Systems Design,* ch. 3: Contextual Inquiry. Morgan Kaufmann Publishers; March 2002

[15] Fetterman, David. *Ethnography: Step by Step.* Sage Publications; February 1998.

[16] Nielsen, J. “Finding Usability Problems through Heuristic Evaluation.” In *Proceedings from CHI 1992,* ACM, pp. 373-380.

[17] Nielsen, J. “Ten Usability Heuristics.” http://www.useit.com/papers/heuristic/heuristic_list.html

[18] Polson, P., C. Lewis, J. Rieman, C. Wharton. “Cognitive Walkthroughs: a method for theory-based evaluation of user interfaces.” In *Proceedings of Man-Machine Studies 1992,* Academic Press, pp. 741-773.

[19] John, B. and D. Kieras. ”Using GOMS for User Interface Design and Evaluation: Which Technique?” In *Proceedings of CHI 1996,* ACM, pp. 287-319.

[20] Kyng, M. “Scandinavian Design: Users in Product Development.” In *Proceedings of CHI 1994*, ACM, pp. 3-9.