Defining, Designing, and Evaluating Peripheral Displays: An Analysis Using Activity Theory

Tara Lynn Matthews
Tye Lawrence Rattenbury
Scott Alan Carter

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2005-20
http://www.eecs.berkeley.edu/Pubs/TechRpts/2005/EECS-2005-20.html

November 29, 2005

Defining, Designing, and Evaluating Peripheral Displays: An Analysis Using Activity Theory

Tara Matthews, Tye Rattenbury, and Scott Carter
EECS Department
UC Berkeley
Berkeley, CA 94720, USA
tmatthew, rattenbt, and sacarter@cs.berkeley.edu

ABSTRACT
Peripheral displays are an important class of ubiquitous computing applications. However, the field has suffered from a lack of clear, consistent terminology surrounding peripheral display research. We present an Activity Theory analysis of peripheral displays in order to establish common terminology and meaning for peripheral displays. We also present an Activity Theory-based approach for designing and evaluating peripheral displays.

KEYWORDS: Peripheral displays, Activity Theory, design, evaluation

INTRODUCTION

Everyday environments include many sources of information that we monitor with little to no conscious effort. Windows reveal the weather, the approximate time of day, and the level of business nearby. A first impression of a restaurant can be gleaned from aromas. Time is passively monitored using clocks; which help us manage when and how to switch between the various activities in our lives. At any given moment, most of the information in our environment is *peripheral* to our main focus of attention. It follows that in a ubiquitous computing environment involving more than one device, most devices and displays must be peripheral. *Peripheral displays* are an important class of ubiquitous computing applications that can allow a person to be aware of information without being overburdened by it [26].

While the general requisite of demanding little to no conscious attention provides a general description of peripheral displays, researchers have nonetheless employed a variety of refinements and generalization in their own definitions. The lacking consensus on a single definition contributes to the difficulty of evaluating peripheral displays, a known problem in the field [13]. Common terminology and meaning when discussing peripheral displays would enable better design and evaluation. One contribution of this paper is to present a way for researchers to understand and describe how peripheral displays are embedded in everyday contexts.

The descriptive theory of peripheral displays we propose and discuss in this paper is based on Activity Theory [11]. Our decision to use Activity Theory is rooted in the observation that any peripheral display may operate in different contexts (both socially and physically defined). Activity Theory provides a framework for describing context [20], and consequently provides a framework for describing how peripheral displays should interact in various situations. Specifically, a peripheral display is any information display that is (1) a *tool* in at least one activity of its user and (2) is used at the *operation* level (*i.e.,* usage requires relatively low cognitive cost[1]). *Tool* and *operation* are terms used in Activity Theory and will be defined in this paper.

Another contribution of this work is an Activity Theory-based approach for designing and evaluating peripheral displays. Our approach involves gaining an understanding of the activities for which you are designing, concentrating on design dimensions and evaluation metrics specific to peripheral displays, and adjusting how evaluation metrics are measured based on the activities supported. From our descriptive theory of peripheral displays we derive a small set of important design dimensions: *scope* (will the display support one or many of the target user's activities?), *class(es) of supported activities* (are the supported activities primary, secondary, or horizon activities for the target user?), and *criticality* (what is at stake if the user is not aware of the information in the display?). We then describe a set of evaluation metrics for peripheral displays – *appeal*, *learnability*, *awareness*, *effects of breakdowns*, and *distraction* – and present guidelines for selecting and measuring these metrics. Finally, we present two case studies as concrete examples of how these evaluation guidelines have been employed.

ACTIVITY THEORY AND PERIPHERAL DISPLAYS

Peripheral displays are often thought of in perceptual terms. In these definitions, peripheral displays are those that literally sit in the periphery of a person's field of vision or that issue subtle auditory cues. While these perceptual interpretations of peripheral displays are initially useful, they quickly break down. What happens if a person changes the angle of their gaze or starts attending to the auditory cue? Does the device cease to be a peripheral display? To fully analyze peripheral displays, one should also consider the "messy" scenarios when a peripheral display becomes the user's focus of attention. Although this type of interaction is well understood – falling into traditional

[1] By "low cognitive cost" we mean that *monitoring* the peripheral display does not distract from the user's focus of attention which, by definition, should not be the peripheral display.

interface design – we have not encountered a complete theory which can address these focal scenarios as well as the intended "peripheral" interaction scenarios.

Consequently, we sought out a descriptive theory of peripheral displays that would extend the perceptual framework by more robustly handling real, everyday usage scenarios. The theory we chose for this purpose is Activity Theory. Activity Theory is an expansive framework describing human behavior; but for our purposes there are four points that are particularly important:

1. **Activities, objects, and motives**. *Activities* correspond to long term (weeks, months, years) projects of a person or group of people. These projects are directed toward some *object* in the world like a new product design or a social relationship. An object may be physical or conceptual, but is always coupled with a *motive* – a driving force that seeks to satisfy some need or needs of the people in the activity [10,11].
2. **Actions and operations**. In terms of how activities are executed, they are composed of goal-directed *actions* which are themselves composed of environmentally contingent *operations*. Actions are equivalent to the classic HCI notion of tasks, operations to the operations from GOMS.
3. **Tool artifacts**. In terms of elements in the world, activities are composed of an object, the person or people involved, and the *tools* that the people use to carry out and support the operations and actions in the activity. Tools are socially constructed artifacts that both encode the operations they are used in and guide the user in formulating goals and actions for using them.
4. **Multiple, ongoing activities**. People have multiple, ongoing activities at any given time; however they are generally working on only a subset of these activities through their current actions and operations.

We discuss these points in the following sub-sections.

Activities, Objects, and Motives

An *activity* is a long term project of a person or group of people. People engage in activities to satisfy their needs – *e.g.*, the need to eat, the need to be social, the need to accomplish [11]. Although a person's needs have biological as well as social origins, their expression is largely determined by that person's socio-cultural milieu. These need expressions are captured in the object and motive of the activity. For example, in attempting to satisfy the needs listed above, a person might orient their actions toward the yield of a harvest, or a familial relationship, or a doctorate degree. Roughly, an object is the topic of an activity, the entity in the world that gives the activity temporal and spatial coherence as well as semantic grounding [2]. An object can be material (*e.g.*, a handcrafted chair), semi-tangible (*e.g.*, a plan for building a chair), or abstract (*e.g.*, a vague but developing design idea for a chair) [10,11].

The object of an activity is always associated with a motive. The motive of an activity summarizes the key aspects or transformations of the object that will satisfy the person's underlying needs. Hence, the motives of activities are generally referred to as life-forces. Lasting multiple days, months or years, these motives constitute a major part of every person's personality [11]. Common contemporary motives include seeking social companionship through the development of a relationship and achieving desired transformations of self through the learning of a new set of skills or the gathering of knowledge.

As we will discuss in this paper, peripheral displays are artifacts that should augment and support one or more of a user's activities. Practically, this means that the peripheral display should promote and empower the user in their efforts to satisfy the motives of these activities. Hence, peripheral display designers need to be aware of their target user's activities, and how a peripheral display could appropriately influence them.

Actions and Operations

Activities are performed through multiple shorter-term (seconds, minutes, hours) processes called *actions*. Actions are similar to what HCI literature calls goal-directed tasks. For example, programming a module might be an action within the activity of completing a software project. In HCI literature regarding peripheral display use, the actions and goals a user is currently focusing her attention on is often called the *primary task* of the user. We will call this the *primary action*. It is important to stress that actions often service multiple activities. This statement includes, but is deeper than the claim that a generic action could service different activities at different times; *e.g.,* writing a passage of text could be a useful action in many different activities. The deeper claim is that a specific action simultaneously services multiple activities; *e.g.,* writing a passage of a research paper might service (a) a research project activity, (b) a developing understanding of the language in which the paper is written, and (c) providing a paycheck with which to support a family.

The activities that are serviced by the user's primary actions provide a context for understanding it [20]. For example, if a person is writing a passage of text, the context surrounding this action is dependent on whether (a) she is writing an email to a relative under the activity of maintaining a familial relationship, (b) she is a researcher writing a section of a research publication, (c) she is writing part of a fictional short story for fun, and so on.

Actions involve multiple *operations*, which are well defined habitual routines used during the performing of the action. Operations are directly influenced by the *conditions* of the environment in which they take place. An important distinction between the operation level and the action level for peripheral displays is that operations require low cognitive load to execute whereas actions require high cognitive load. This distinction is a consequence of the creative nature of setting and accomplishing action goals versus executing a rote habitual routine at the operation level. This distinction practically means that people are generally working on at most one action at a time (their primary action) due to limitations of cognitive resources like working memory. Conversely, a person can carry out many operations in appropriate environments, regardless of how related the operations are (*e.g.*, drinking and eating while having an intense discussion about a research project).

Thus, if a display is to be classified as peripheral it should be used primarily at the level of operations (and not at the action level). This does not mean, however, that a peripheral display will never reach the action level. In fact, a new peripheral display must be learned and appropriated by its user, which is generally done at the action level or even as a separate activity [11]. (Learning how to use a new peripheral display is one of the evaluation metrics that we discuss later in this paper.) For example, many peripheral displays abstract the information they present in some way. The user generally has to learn this abstraction to be able to easily (*i.e.* at the operation level) interpret the information presented.

Operations are sequenced to complete an action, but this sequencing is not always optimally efficient or constructive (*e.g.,* trial-and-error methods of problem solving). An important type of operation sequencing in the design of peripheral displays involves chaining together operations that do not build on one another. For example, most people will subconsciously glance at a clock or out a window while they are working on various actions. These glancing operations may not contribute to the completion of the user's primary action, but they occur nonetheless. Later in this paper we will define a class of peripheral displays that rely on operations that do not service the user's primary action.

Tool Artifacts

A fundamental pillar of Activity Theory is that people's interaction with the world is mediated by physical and psychological tools [11]. In relation to the other concepts in Activity Theory, tools are artifacts that enable people to act on the objects of their activities. In other words, these tools allow users to accomplish, understand, motivate, or see the future transformations of their activities.

Tools are socially constructed. Hence they are subject to the trends and fashions of cultures which are constantly evolving and transforming. Individually, tools are appropriated and adapted for various actions and operations. Activity Theory is vague when defining tools because it tries not to bias or limit the full variety of forms and functions that mediating artifacts might and do take.

Peripheral displays are tools. Moreover, the importance of a peripheral display is determined by its importance in the activity or activities that it supports. This could range from low (*e.g.,* a single clock in a room with multiple clocks) to high (*e.g.,* the altimeter in a cockpit).

Classification of a Person's Multiple Ongoing Activities

As noted above, a person's actions can service multiple activities simultaneously. Also, the sequence of operations performed by a person may include some operations that do not service the goal of their primary action. Relying on these two points, we posit a set of four classes of activities. Note that this classification is not standard in Activity Theory: we derive them here to categorize the types of activities peripheral displays are likely to support. Also, note that these classes of activities are not universal or static. They not only differ person to person, they change as the person's primary action changes (*e.g.,* a horizon activity may become a primary activity).

1. *Dormant activities*. This class includes activities that have been set aside by the user. Specifically, dormant activities are not serviced by any operation performed in the user's current sequence of operations. Since these activities are dormant, they are not serviceable by any peripheral display and hence are irrelevant for design and evaluation.
2. *Primary activities.* This class includes activities that are serviced by the user's primary action. In other words, it is the class of activities that are serviced by operations needed to complete the user's primary action. We will refer to tools that support this class of activities as primary tools.
3. *Secondary activities*. This class includes the activities that are serviced by operations that are in the user's primary action but do not promote the attainment of the primary action's goal. In other words, these activities are not the focus of the user's current action. Also, they are not likely to become primary in the near future.
4. *Horizon activities*. This final class of activities is similar to secondary activities with one important distinction. Both secondary and horizon activities are monitored by the user, but horizon activities are monitored with the intent that they will become the primary activity in the near future. In other words, a horizon activity was once a primary activity that is temporarily "on hold." The user is monitoring some aspect of the activity to decide when to start working on it again.

An Activity Theory Definition of Peripheral Displays

Based on the four characteristics of Activity Theory that we introduced above, we arrive at the following definition: a peripheral display is any information display that

1. is a tool in at least one activity of its user and
2. is used primarily at the operation level.

This definition is framed by an understanding that the peripheral display user has multiple, ongoing activities. Depending on how the peripheral display is used, it may be a primary tool in the user's primary activities or a non-primary tool supporting the user's secondary and/or horizon activities.

Notice in our definition of peripheral displays that they are primarily used at the operation level. We acknowledge an important class of displays that are designed to work at the action level instead. Following existing terminology, we refer to this class of displays as *notification displays.*

To illustrate the distinction between peripheral and notification displays, consider Figure 1a. This figure schematically illustrates two consecutive actions taken by a user (Jane). In this case, Jane has three activities. Two of these activities are primary (A and C) and one is a secondary activity (B). At the start of this diagram, Jane consciously selects an action to work that is related to activity A (*e.g.*, writing a paragraph in a research paper). While she performs the operations that compose action A, she also performs operations related to B (*e.g.*, monitoring a repetitive stress indicator). When Jane completes the action associated with activity A she consciously selects a new action related to activity C (*e.g.*, reading an e-mail from her child's daycare center). Again, while she performs this action she performs operations related to activity B. *The tool that allows Jane to complete operations on activity* B *while completing actions related to different primary activities is a peripheral display*.

Now consider Figure 1b. In this case, Jane receives a notification that pertains to activity D (*e.g.* her child's health) while she is completing an action related to activity A. This notification forces Jane to consciously switch her action. In fact, most notification displays present information that is intended to instigate action/task switches. Although the information presented in a peripheral display might result in the user switching actions, the peripheral display should be designed such that a user is not interrupted and can choose to finish her current action first. We illustrate this type of interaction in Figure 1c, where the peripheral display shows the new state of activity D, but Jane finishes her action servicing activity A before switching.

APPROACH FOR PERIPHERAL DISPLAY DESIGN AND EVALUATION

One of the central contributions of this work is an Activity Theory-based approach for designing and evaluating peripheral displays. Our approach involves: (1) preliminary investigation of the activities of target users; (2) selection of activities that the display will support and initial design of a display to supports those activities based on our design dimensions; (3) evaluation of displays based on our evaluation metrics and suggested methods; (4) deployment; (5) reassessment of activities that the display supports; and (6) design iteration.

In the first stage it is necessary to gauge the activities that will be important for the target users of the tool (*e.g.*, for an e-mail awareness display a designer should determine whether participants use their e-mail for work-related activities only or have a strong overlap with personal activities). This can be accomplished via interviews, surveys, ethnographies, or automated activity-clustering analysis.

In the second stage, it is important to establish the scope of the display (will it support one or many of the target user's activities?), the class of the activities that the display will support (will the supported activity be the target user's primary, secondary, or horizon activity?), and the criticality of the display (what is at stake if the user is not aware of the information in the display?). Once the designer has assessed how the proposed peripheral display relates to our three design dimensions, he or she should have enough information to start prototyping and building. We describe our design dimensions in more detail later in the paper.

The third stage involves evaluating the display. Evaluation can begin with a discount method, like heuristic evaluation [13]. However, more extensive evaluation methods should follow. However, prior work on peripheral displays has not provided clear guidelines for such evaluations. In this paper, we focus our discussion on guiding the selection of an extensive (*i.e.* non-discount) evaluation method to measure five metrics: appeal, learnability, awareness, effects of breakdowns, and distraction. The method of measuring and the importance of these evaluation metrics depends on the scope, activity class(es), and criticality of the display. We discuss evaluation metrics and methods later in the paper – with particular focus on how differences in each design dimension affect their application. To foreshadow this discussion, we highlight the following cases. For a critical display it is important to show that it fits into its environment, it is empirically easy to learn, it provides appropriate levels of notification, it is explicit about breakdowns, and is only as distracting as the importance of the information being shown. A display that is not critical should be aesthetically pleasing, need not be easily learnable, should convey information only according

Level	**Current activity**
activity:	AAAAAAAAAAAAAAACCCCCCCCCCCCCC
action:	AAAAAAAAAAAAAAACCCCCCCCCCCCCC
operation:	AAABAAABAAABAAACCBCCCCBCCCCBC
	Time ⟶

(a) Peripheral display supporting activity B.

Level	**Current activity**
activity:	AAAAAAAAAAAADDDDDAAAAAAAAAAAA
action:	AAAAAAAAAAAADDDDDAAAAAAAAAAAA
operation:	AAABAAABAAAADDDDDAAAABAAAABAA
	Time ⟶

(b) Notification to switch to activity D.

Level	**Current activity**
activity:	AAAAAAAAAAAAAAADDDDDDDDDAAAAA
action:	AAAAAAAAAAAAAAADDDDDDDDDAAAAA
operation:	AAABAAABAADDAAADDBDDDBDDAABAA
	Time ⟶

(c) Peripheral display supporting activities B and D.

Figure 1: Distinguishing peripheral and notification displays.

to a user's interest in the display, need not display breakdowns, and should be minimally distracting. Finally, a display that is supporting a primary or horizon activity is typically best evaluated via a mixture of qualitative methods and quantitative lab studies, while a display supporting secondary activities is typically best evaluated via a field study.

The remaining stages mirror the standard design cycle of prototype and deploy, test and evaluate, and design/re-design. Peripheral displays can be relatively generic tools, and users are likely to appropriate generic tools for unforeseen activities. Thus, it is important that following the deployment phase the designer evaluates the activities the display has evolved to support. Again, this can be accomplished via qualitative methods, automated analysis, or a combination of the two. Based on the findings of this second evaluation, the peripheral display design should be iterated to better support any new activities it has evolved, or been appropriated, to support.

With the same goal as our design approach, Mankoff and Dey [14] present a design process for peripheral displays. Their design process includes (1) interviewing and/or surveying people to find a location and type of information to display, (2) designing a display for that location and information, and (3) conducting a summative evaluation of the display. Our contribution goes beyond this work by grounding our design process in Activity Theory, describing peripheral display design dimensions, and presenting evaluation methods and metrics for each design dimension.

PERIPHERAL DISPLAY DESIGN DIMENSIONS

Building from the framework of Activity Theory, we identify three primary design dimensions for peripheral displays: scope of use, class(es) of supported activities, and criticality. We will describe each of these in turn, and then provide a list of example peripheral displays covering these dimensions.

Scope of Use

Peripheral displays are tools in their user's activities. The *scope of use* of a peripheral display refers to the number of activities the display is designed to support. Note that users will likely appropriate peripheral displays to support a variety of activities not originally intended by the designer. But this dimension refers specifically to the number of activities that the display is designed to support, rather than future activities that the display may support.

Designers should consider whether they are creating a display that is either (1) applicable to one specific activity or (2) applicable to more than one activity. The first category refers to displays that enable the user to perform operations that service only a single activity. If the user adopts a new activity in the future which can utilize the display, then the display can be re-designed and re-evaluated to better support this new activity. The second category includes peripheral displays whose associated operations service more than one activity. For example, a peripheral display that allows its user to monitor email (*e.g.*, the email orb [8]) can support both work and personal activities.

Note that we exclude displays that do not service any activity – *i.e.* those displays that do not promote or empower the user in their efforts to satisfy the motives of any of their activities. *E.g.*, very low fidelity monitoring devices, such as simple fire alarms that detect fires but do not provide the user with any additional information on how to deal with the fire, would categorically fall into this excluded group of displays. Ultimately, however, whether or not a display can service an activity depends on the user. For example, the altimeter in a cockpit is not a peripheral display for someone who does not pilot planes.

Class(es) of Supported Activities

People are generally involved in multiple ongoing activities but are focused on a small set of these through their primary action. The *class(es) of supported activities* design dimension captures whether the peripheral display supports primary activities, secondary activities, and/or horizon activities (see the description of these classes of activities above). Although it is natural to consider only two cases – peripheral displays that support primary activities and peripheral displays that do not – we further distinguish between displays that support secondary activities and displays that support horizon activities. Practically, this distinction allows the peripheral display designer to assess whether they need to design for transitions that enable a horizon activity to become a primary activity; or whether they only need to design for the monitoring of some aspect of a secondary activity. Scalable Fabric [24] is an example of a peripheral display designed to support horizon activities. It enables users to monitor the status of pending actions/tasks that have been temporarily set aside but could be resumed at any time. The display also supports transitioning between different activities, appropriately modifying its view when horizon activities become primary activities.

Even when a peripheral display does not support the user's primary activities, a designer should have some understanding of what the user's primary activities are and how the user's primary action supports them. For example, a peripheral display designer needs this understanding to choose an appropriate interaction modality. In some situations, interaction modalities carry different restrictions (*e.g.,* when driving a car, audio is more appropriate than visual). Without assessing the user's primary action, and hence the context of their current situation, a designer may not make appropriate or justifiable design decisions.

Criticality

The final design dimension is *criticality*; which refers to how critical or important the activities the display is designed to support are for the user. Though criticality represents a continuous range, we consider two categories for simplicity: non-critical and critical. Although these categorical distinctions are largely universal (*e.g.,* most people would consider activities involving life or death situations as critical), Activity Theory does provide a handle on a more subjective sense of criticality. Since each activity for a user has an associated motive, one can gauge criticality by assessing the importance of the appropriate motives for the user – generally speaking, the more important the motive the more critical the activity.

List of Peripheral Displays

Our design dimensions have meaning only in relation to specific people and their sets of activities. Below we describe two simple personas. Then we present example displays for each of our three design dimensions and how they might be used. These examples serve to illustrate the connection between a peripheral display, its use, and the activities of the user.

Personas

Bob is a social services coordinator who exercises regularly, owns a home, and flies single-engine planes in his spare time. **Jane** is a computer science professor who is partially deaf, a single mother of a toddler, and a frequent user of public transportation.

Peripheral Display Examples

Scope of 1, Primary activity, Low Criticality: It is nearly time to head home and Jane is wrapping up loose ends and getting organized to leave work. Her office has an information ticker that indicates bus arrival and departure times. Since Jane is planning on riding the bus, she peripherally monitors the ticker to make sure she finishes her wrap-up in time to catch the next bus. Here, the ticker supports *one specific activity* for Jane: managing her work-life balance. This activity is Jane's primary activity and the information displayed in the ticker is non-critical (Jane could walk home if she misses the bus).

Scope of 1, Secondary activity, Low Criticality: Jane is concentrating on writing a research paper and a repetitive-stress monitoring program presents the length of time she has gone without taking a break. Here, there is one activity that the display supports (maintaining her health), the supported activity is secondary to her primary activity (writing a research paper), and the supported activity is relatively non-critical.

Scope of 1, Primary activity, High Criticality: Bob is monitoring an altimeter while flying a plane. Here, there is one activity that the display supports (flying a plane), the supported activity is Bob's primary activity, and the activity is critical (negligence of the activity would risk Bob's life).

Scope of 1, Secondary/Horizon activity, High Criticality: While Jane is cooking, she monitors a high fidelity visual display showing her baby playing in another room. In this case, the display supports one activity (maintaining the health of her baby) which is critical. This activity could be secondary if Jane's baby is relatively independent and Jane is not expecting to directly attend to her baby in the near future. Or, if Jane expects her baby to need her direct assistance in the near future, this could be a horizon activity.

Scope of 2, Primary and Secondary activities, Low Criticality: Jane is using a version of IC2Hear, an application that displays sound location and content information in a visual display, to monitor audience noise and to gain feedback of her own voice level while she teaches a class. While IC2Hear can support low- and high-criticality activities, here the display supports two low-criticality activities: managing a class and practicing public speaking. The former activity is primary while the latter is secondary.

Scope of 2, Horizon activities, Low Criticality: Jane is using Scalable Fabric to manage research projects that all have approaching deadlines. Her primary action is writing part of a research paper in activity A, but she is also waiting for an email to finish a grant proposal for activity B and for some data-processing algorithms to finish in activity C. Jane is using a Scalable Fabric to monitor the arrival of the email she needs and to determine when the data-processing has finished. Here, there are two supported activities, both of which are non-critical, horizon activities for Jane. The display also supports transitioning between different activities. For example, when the email arrives for activity B (Jane's grant proposal), Scalable Fabric shows that there is an email in Jane's inbox. But when Jane clicks on the inbox to read the email, activity B becomes the primary activity and Scalable Fabric appropriately modifies its view of activity B.

Scope of 2, Primary and Secondary activities, High and Low Criticality: Bob uses his heart monitor while at work to monitor his stress and fitness level. He also uses his heart monitor while flying to gauge his susceptibility to G-force blackouts. Here, the display supports two activities: health monitoring and flying. The former is non-primary and non-critical, the latter primary and critical.

Scope of Many, Secondary activities, High and Low Criticality: When Bob is busy talking with people over the phone or during meetings, he monitors a display that indicates the number of high priority e-mails remaining in his inbox. Here, the display supports more than one specific activity (Bob's high-priority e-mails relate to different work activities as well as some of his personal activities), which are secondary and of varying levels of criticality. Due to the types of actions Bob performs while using this display, the display is entirely visual (auditory cues would be disruptive to Bob's meetings and conversations).

EVALUATION METRICS

Traditional user interfaces that support a user's primary activity are evaluated on efficiency-related metrics related to direct interaction: efficiency or time to complete supporting actions, success rate of completing actions, number of errors, and quality of the resulting object. However, since peripheral displays are designed not to be directly interacted with, it is often harder to assess the influence of these displays on a person's overall work effectiveness. Hence, traditional evaluation metrics are not applicable for peripheral displays.

This section presents five evaluation metrics for peripheral displays that we derived from interviews with ten peripheral display creators presented in [15] and from our own design and evaluation experience. The metrics are *appeal*, *learnability*, *awareness*, *effects of breakdowns*, and *distraction*. First we define these metrics. Then we describe how to select and measure the metrics for various peripheral display designs. Our discussion on how to apply the evaluation metrics is based on Activity Theory.

Appeal (Usefulness, Aesthetics)

Appeal refers to how much the user qualitatively enjoys using the display and the way it looks. In other words, this metric represents their overall feelings about the display. This metric can be broken down into *usefulness* and *aesthetics*, which represent two very different aspects of users' qualitative feelings about a display.

Four of ten designers we interviewed mentioned usefulness and aesthetics as important when evaluating peripheral displays. They noted that adoption of a peripheral display depends on its appeal to users. In addition, this metric is informed by the ambient display heuristics presented in [13], several of which affect the appeal of a display: "aesthetic and pleasing design," "useful and relevant information," and "match between design of ambient display and environment."

As we discussed in the scope design dimension, designers will often create peripheral displays to augment and support specific activities of their target users. These activities are already ripe with other tools, social contexts, and semantic themes. By understanding these aspects of the user's activities the designer can conceptualize what displays might be useful and aesthetically pleasing. Of course this type of user profiling is common practice in many design methods, but we think by explicitly considering the target user's activities, designers will be better equipped in assessing the appeal of their displays.

Learnability

Learnability is the amount of time and effort required for users to learn to get information from the display at the operation level. This metric is often affected by the chosen mappings from inputted to displayed information; since peripheral displays do not always display input literally but instead show it in an abstract form.

The learnability of a display may influence its adoption. Users may be less likely to use displays that are difficult to learn, unless interpreting the display is meant to present a challenge and users expect this [5]. Again, the ambient display heuristics point to learnability as an important metric ("consistent and intuitive mapping"). This heuristic assumes that the display should be easy to learn and interpret. We agree that learnability is important, but we argue that the designer should determine how learnable a display should be based on the activities in which the display is used and on feedback from users. Every person will require a different amount of practice to use a new display peripherally (*i.e.* at the operation level). In Activity Theory, new tools are often traced from the activity level (when the tool is first introduced and the user is trying to learn about the tool as it relates to other tools) to the action level (when the user still has to direct their attention to the tool to use it appropriately) to the operation level (when the user can use the tool without high cognitive load [11]). The rate of this learning depends on the user as well as the situation; sometimes a challenging display is desirable [5].

Awareness

Awareness is the amount of information shown by the display that users are able to recall, understand, or use.

Since the purpose of a peripheral display is to convey some information, it follows that the user's awareness of that information can be used to judge the effectiveness of the display. Three of ten interview participants said user awareness of information is an important metric for evaluating peripheral displays. Our past studies of the Bus Mobile [13], hebb [1], sound displays for the deaf [7], and email displays [8] have all explored measuring awareness. Many studies in peripheral display literature have also measured awareness as an important metric [1,17,18,19, 21,22].

Like appeal and learnability, awareness will depend on the user's activities. The designer should consider how the information in the peripheral display relates to the other information in the user's activities. Understanding this relationship will help the designer assess how aware the user needs to be about information in the display, and also how often it should be updated and how often the user should be monitoring the display.

Effects of Breakdowns

The effects of breakdowns refer to how apparent breakdowns are to users and how easily users can recover from them. The design goal should be to make breakdowns obvious and recovery easy.

The ambient heuristics point to the effects of breakdowns, saying that "error prevention" is an important design consideration, since "users should be able to distinguish between an inactive display and broken display." In an evaluation of the Bus Mobile the visibility of breakdowns was shown to be a problem [13]. The state signified by the bus tokens still underneath the white skirt (see Figure 4) had two possible meanings: no buses are scheduled or a motor is broken. Users could not tell the difference. One of our interview participants told a story of a peripheral display breakdown causing mild panic in the lab group. The display showed the activity on a main server. At one point, an error in the display caused it to freeze. Users interpreted this to mean that the main server had frozen and they frantically searched for problems with it. In one of our past field studies of the Email Orb [8], the Orb was not displaying anything for half a day before users noticed. These examples show the importance of making breakdowns apparent and recovery easy.

As with the earlier metrics, understanding the effects of breakdowns depends on the user's activities. If the designer understands what aspects of the user's activity are influenced by the peripheral display, and how these aspects relate to the activity as a whole, the designer can make more appropriate decisions on how to expose breakdowns and how to support recovery. Understanding the criticality of the activity and the scope of the display are also important. Breakdowns in the gauges in a cockpit could represent a serious concern which needs to be communicated to the user (pilot) without inducing unnecessary stress or cognitive load. Likewise, if the display supports multiple activities (scope greater than one) the designer should consider the inter-activity interactions that could influence how to communicate the breakdown in the display and how the user can most efficiently recover from it.

Distraction

Distraction is the amount of attention the display attracts away from a user's primary action.

Designers should measure distraction since it affects the user's ability to carry out their primary action and will likely influence their qualitative reactions to using the display. Three of ten interview participants named distraction as an important metric, saying that a crucial measure of success for *peripheral* displays is that they *be peripheral* and not unnecessarily distract the user. Our past studies of the Bus Mobile [13], hebb [1], sound displays for the deaf [7], and email displays [8] have all measured distraction, often (but not always) with a goal of minimizing it.

Distraction is a natural metric given our Activity Theory analysis. Since the display is monitored at the same time as the user is performing an action, it will necessarily require some user attention. This does not necessarily result in the user being less efficient. For example, the display could provide information that is useful and/or important to the user's action, enabling them to perform better with the display than without.

In the next section we describe how Activity Theory and our design dimensions constrain how these evaluation metrics should be applied and measured.

GUIDELINES FOR DESIGNING AND EVALUATING DISPLAYS IN EACH DESIGN DIMENSION

How a peripheral display should be evaluated will vary based on our Activity Theory-derived design dimensions (scope, class(es) of supported activities, and criticality). In particular, metrics will vary in importance depending on a display's relative position along each design dimension (*e.g.,* for displays with *high criticality, awareness* is more important and *appeal* is less important). In this section we discuss the evaluation of peripheral displays relative to each dimension.

For scope, we distinguish between displays that are associated with one activity versus displays that are associated with more than one activity. For classes of supported activities, we discuss peripheral displays that support the user's primary activity (*e.g.,* the altimeter in a cockpit), displays that support secondary activities (*e.g.,* a bus schedule display), and displays supporting horizon activities (*e.g.*, a pending task display). For criticality, we discuss peripheral displays that are associated with critical activities (*e.g.,* the altimeter) and displays associated with non-critical activities (*e.g.,* a bus schedule display).

In our experience, scope, class(es) of supported activities, and criticality can be treated as independent dimensions. Accordingly, we present how each dimension influences peripheral display design separately. Evaluators should combine the evaluation methods presented below depending on where their display falls within these dimensions.

Scope

In general, changes in scope do not result in changes in the relative importance of our evaluation metrics – *e.g.*, knowing that a peripheral display supports one activity instead of two does not make appeal any more or less important. However, as scope increases, evaluating each metric can become significantly more complicated. In this section, we describe the additional complications that evaluators should consider.

- *Appeal:* As scope increases, the designer needs to consider the impact of the peripheral display in more situations and contexts. Consequently, the appeal of the display will have more constraints. For example, if a peripheral display supports both work and personal activities, the designer will need to consider the potentially different aesthetics for each of these domains of the user's life and find a design that harmonizes with both.
- *Learnability:* As scope increases, the designer must consider how the user will learn to use the peripheral display in each of the activities that are supported. In some cases, a single method of learning is sufficient. However, there are situations where learning to use the display would differ depending on the user's situation. For example, in the earlier scenario of Bob using his heart monitor, we considered two situations: monitoring his health while at work and using the heart monitor while flying. In these situations, Bob's use of the heart monitor could be significantly different requiring a new learning process.
- *Awareness:* As scope increases the number of situations and contexts that the peripheral display is used in also increases. Thus, measuring a user's awareness could become more complicated. For

example, if an email monitor is designed to support both work and personal activities, the ability for the user to maintain awareness of information in the display may change as the relative importance of the supported activities changes. In this case, the designer should consider how different situations (*e.g.*, upcoming work deadlines, important personal matters, lifestyle changes) alter the user's awareness of the display.

- *Effects of breakdowns:* As scope increases the number of potential usage situations also increases. To assess the effects of breakdowns, the evaluator not only needs to consider how a breakdown influences each supported activity individually, but they also need to assess how the breakdown influences combinations of activities and moments of transition between activities. For example, in the earlier scenario of Jane using IC2Hear, if the display malfunctions and indicates that she is talking far too quietly, Jane's reaction will not only influence her primary activity of managing her class but it will also influence her public speaking activity.
- *Distraction:* Generally, as scope increases, assessing distraction becomes more complicated. If a peripheral display conveys information about many activities, it also potentially conveys more information than a display that only supports one activity. This means there are more opportunities for distractions. Designers need to assess not only activity-specific distractions (*e.g.*, IC2Hear distracting Jane from managing her class), but cross-activity distractions like an important personal email causing Bob to step out of a meeting.

Lab studies can be particularly useful in examining support for a single activity in depth. But, as scope increases, controlling lab experiments to handle all the possible interactions of the supporting activities becomes unmanageable. Thus, as scope increases we recommend increasing the use of field evaluations.

Also, for displays that support transitioning between multiple primary, secondary, and horizon activities, it is important to evaluate support for the transitions themselves. In particular, a display supporting horizon activities should change its representation of an activity when that activity transitions from being a horizon activity to being the user's primary activity. We discuss this case in more detail below. Of course, if a high scope display supports the user's primary activities in addition to horizon activities, it should treat information for the primary activities in a different way since this information is relevant to the user's current focus of attention.

In the following two sub-sections we discuss how the *criticality* design dimension influences our evaluation metrics.

High Criticality Displays

Peripheral displays associated with a critical activity should present information in the most readable form possible. Hence, they should provide excellent information awareness. These displays are not tools for opportunistic information, curiosity, or aesthetic appeal. This means metrics like appeal will be less important than learnability, awareness, effects of breakdowns, and distraction. Below we describe how each metric should be adjusted for displays that support activities that are highly critical.

- *Appeal:* It is more important to measure how well the display fits with its environment than to measure its aesthetic appeal. For example, it is not important to measure how beautiful a cockpit gauge is independently, but it is important that the gauge not unduly stand out in its environment, making it likely to garner more attention than it deserves.
- *Learnability:* It is important to empirically demonstrate that the display is readily learnable. This is important because critical activity displays are related to an important motive and should not distract the user from focusing on the associated activity, even during the learning period. For example, the altimeter gauge in a cockpit is meant to support quick access to important information. An extended learning curve on this display would be inappropriate and dangerous.
- *Awareness:* It is important to measure how well the display communicates its information. For the altimeter example, a well-designed gauge might allow pilots to be aware of normal information as they deem appropriate, but alert the pilot when the plane's altitude is abnormal.
- *Effects of breakdowns:* It is vital to reveal breakdowns to the user since they may depend on the information in the display to support their critical activity. The notification of the breakdown should be conveyed quickly and clearly. Recovery should be easy for the user. For example, a pilot would need to know immediately if the altimeter broke and would need a way to remedy the problem.
- *Distraction:* Any distractions caused by a critical activity display should correspond to the importance of new information being shown (*i.e.,* the more important the information, the more distracting the display could/should be). The display may need to indicate to the user that they should abandon their current action for a different one. For example, a pilot might be preparing the landing gear when the altimeter indicates that the plane's altitude is too low. This information should be shown in a way that distracts the pilot from her primary action in order to gain her attention. With the new information, the pilot may abandon her current action and start a new one like steering the plane upward.

Low Criticality Displays

A non-critical activity display is a tool for maintaining awareness related to a relatively unimportant activity. For example the Bus Mobile [13] provides bus schedule information for an activity of managing a work-life balance. Non-critical activity displays are often characterized as being artifacts for opportunistic information, curiosity, or aesthetic appeal, rather than for productivity. This means metrics like appeal and distraction are more important than learnability, awareness, and breakdowns. Below we describe how each metric should be adjusted for displays that support activities that are less critical.

- *Appeal:* This is an important metric since the display is non-critical and likely used by choice rather than necessity. Users will not adopt a display unless it is aesthetically pleasing or useful. The level of harmony between the display and its environment is a matter of aesthetics (see, for example, Informative Art [23]).
- *Learnability:* A non-critical display need not necessarily be readily learnable, but the learnability of the display should match user expectations. The designer should expose the model by which a user is to learn the meaning of the display: a designer may create a display that is by design difficult to comprehend, but she should expose to the user the furtive nature of the display's meaning. For example, Slow Technology [5] is designed to be slow to learn and understand in order to give people time to reflect.
- *Awareness:* It is important to measure awareness only in relation to a user's interest in the displayed information. That is, it is important to know the extent to which a user retains information only when they have some interest in that information. For example, users who have minimal investments will likely consider information conveyed by a hallway stock display irrelevant and their awareness of the conveyed information should not be tested.
- *Effects of breakdowns:* It is less important to notify users that a breakdown has occurred. It is still important to design and verify a simple, obvious recovery, method, but it is less crucial than it is for critical activity displays.
- *Distraction:* Any distractions caused by a display should correspond to the importance of new information being shown. For example, if the display shows information of very low importance, it should not distract the user from her primary action.

In the following three sub-sections we discuss how the *class(es) of supported activities* design dimension influences our evaluation metrics.

Displays Supporting Primary Activities

Some displays are designed to support the same activity, or activities, that the user's primary action supports. For example, the altimeter in an airplane supports the activity of safely flying. The pilot is assumed to be performing some action related to this activity at all times. Therefore, monitoring the altimeter will always be done simultaneously with a primary action that is servicing the flying safely activity.

A primary-activity display supports the user's current activity and may influence the user's primary action. This means that the display is related to a specific context and a relatively stable set of actions that support the user's primary activity. Therefore, evaluations can be conducted in a lab setting where the activity and context of use are simulated. Field studies may also be conducted and could provide valuable qualitative results, though lab studies have several benefits. First, they allow better control over experimental variables, allowing for quantitative measures. Field studies, on the other hand, make quantitative measures difficult. Second, lab studies can be easier to conduct since they are shorter-term and allow for the use of Wizard of Oz or partially implemented prototypes. It is important that the context of use is realistically simulated in the lab, because peripheral display use depends on the environment and the user's primary action. In such a lab study, each metric can be measured as follows:

- *Appeal:* This metric should be measured qualitatively via interviews following use.
- *Learnability and awareness:* These metrics can be measured quantitatively, which is beneficial for situations when they are very important (*i.e.,* for critical activity displays). A user's awareness of information can be tested with knowledge questions regarding the displayed information. Gathering awareness data at various stages in the lab study gives an indication of how quickly and effectively the user learned to get information from the display.
- *Effects of breakdowns:* This metric can be measured in the lab by simulating breakdowns and observing a user's response.
- *Distraction:* This metric can be measured in a lab study via holistic speed and success (*e.g.,* in the activity of flying safely, the user's current goal might be to land the plane; the altimeter gauge might indicate that the plane is approaching safely but distract the user so that she makes a mistake and crashes or forgets an important step like communicating with the radar tower).

Displays Supporting Secondary Activities

Since secondary activities do not typically become primary activities, progress is made by monitoring their status. For example, the Bus Mobile [13] shows bus schedule information, related to the activity of managing a work-life balance by providing users with help in deciding when to leave work and head home. While waiting for a bus to be close enough, a user could be working on another activity, such as writing a research paper.

Since the user's primary action and primary activities generally define their context [20], a peripheral display that only supports secondary activities could be used in a variety of contexts which may vary and change. It follows that the context in which the peripheral display is used is difficult to realistically simulate, making lab studies difficult. Since it is important that peripheral displays be evaluated in the context of realistic use, field studies should be used for peripheral displays supporting secondary activities.

A field study uses primarily qualitative measures of evaluation metrics. For this reason, a display should be deployed for an extended period of time (*i.e.,* weeks) so users can develop accurate qualitative feedback. In such a field study, each metric can be measured via qualitative feedback from users during and after deployment.

Secondary activity displays do not impose any priority on evaluation metrics. Rather, designers should determine the display's other design dimensions before assigning importance to metrics. For example, ensuring that a display is very easy to learn is more important for high criticality displays than for low criticality displays – the fact that the display supports secondary activities does not affect this metric.

Displays Supporting Horizon Activities

Horizon activity displays are designed to allow users to monitor non-primary activities that may become primary in the near future. This type of monitoring does not involve making progress on the activity; rather, the user switches to the activity to make progress on it. For example, Scalable Fabric [24] enables users to monitor the status of actions that have been temporarily set aside but could be resumed at any time. For example, users might set aside activities while waiting for a file to download, or for a co-worker to provide some information, or for a paper to finish printing.

In-lab evaluation is possible for horizon activities. Switches between a primary activity and the horizon activity can be simulated in a semi-realistic way. For example, a lab study could involve users monitoring a peripheral display of email updates so that they can switch to an email activity when certain messages arrive, while editing a document as a primary activity. This simulated scenario is semi-realistic because it is fairly well defined (*i.e.,* though the specific email a user would wait for is not clearly defined and the primary activity could vary, this a realistic situation for many users).

Like secondary activity displays, horizon activity displays do not impose any importance level on evaluation metrics. Designers must examine the other design dimensions (criticality and scope) in order to determine the priority of each metric. However, since horizon displays may support switching to a monitored activity, evaluators should evaluate the learnability of this mechanism. Ideally, it would be easy to learn how to switch between the primary and horizon activities.

In the following section, we show how the design dimensions and evaluation metrics we have presented are used in case studies.

CASE STUDIES

Here we present two case studies in evaluating displays using our Activity Theory based design dimensions and evaluation metrics. The first is IC2Hear, a visual display of sound for the deaf. In our deployment of IC2Hear, it was a high scope, high criticality, primary, secondary, and horizon activity display. The second is the Bus Mobile, a physical display of bus arrivals. In our deployment of the Bus Mobile it was a low scope, low criticality, secondary activity display.

IC2Hear

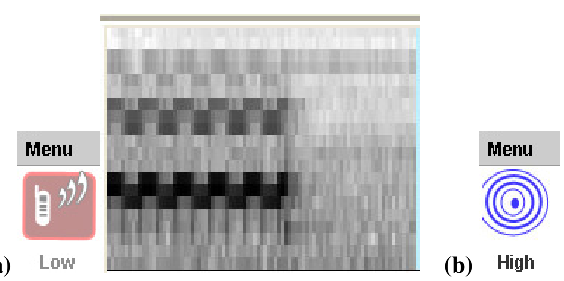
IC2Hear is a peripheral, visual display to help deaf people maintain an awareness of sounds in their environment [7,16]. We discuss results from [16], since we conducted this research and had control over the evaluations.

To begin, we interviewed 8 participants who are deaf to understand how they would use a visual representation of sound in their daily lives, and what design and functional preferences they had. This enabled us to understand what design dimensions (scope, class of use, and criticality) we should target in our prototypes.

Participants told us they would value increased awareness of sounds at home, work, and when mobile. They listed many sounds, including emergency alarms, activities of other people, phones ringing, doorbells, household appliances, traffic, dogs barking, and so on. Such diverse use meant that our IC2Hear prototype would have a high scope.

Also, participants listed sounds indicating events of both high (*e.g.,* emergency alarms) and low criticality (*e.g.,* activities of others). This meant our IC2Hear prototype would be both a high and low criticality display.

Finally, participants indicated that they would want to use the display to monitor sounds that could provide them with a better awareness of their surroundings, the people around them, and potentially dangerous situations. All of these are *secondary activities*. Participants also described scenarios involving monitoring their usage of household and office appliances (*e.g.,* knowing when water boils in a kettle or when the microwave timer chimes). These are examples of *horizon activities*, which users are monitoring in order to resume later. Participants did not list any *primary activities* for which they wanted increased sound awareness, but we



(a) (b)

Figure 2: (a) Implemented version of Spectrograph with Icon showing a phone ringing. This initial prototype places the spectrograph and icon windows side-by-side. (b) Single Icon showing a loud unrecognized sound with low frequency.

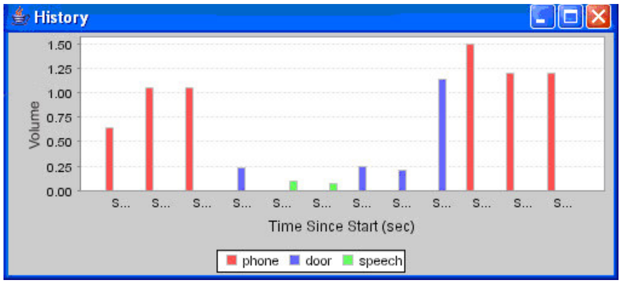


Figure 3: History Display shows each sound recognized as a different colored bar along a time axis. The height of the bar indicates the relative volume of the sound.

decided to make our design flexible enough to support them (*e.g.,* our persona Jane monitors audience response while giving a lecture).

We implemented two IC2Hear prototypes that utilize sound recognition technology to identify sounds. Figure 2b shows Single Icon, which indicates the current sound with an icon. Here it shows many large rings, indicating a loud, unrecognized sound. Figure 2a shows Spectrograph with Icon, a combination of Single Icon and a standard spectrograph, a visualization of amplitude (darkness), frequency (y-axis patterns), over time (x-axis). Here it shows a cell phone ringing. Figure 3 shows the History Display, a bar graph of past sounds where color identifies the sound, height of the bar indicates the sound's volume (*e.g.,* loud sounds make taller bars), and position along the x-axis indicates time. It shows recognized sounds only.

Since IC2Hear has a high scope, supporting many primary, secondary, and horizon activities, we conducted an uncontrolled lab evaluation to gather qualitative feedback on any office activity. We narrowed the scope by simulating an office setting and training our system to recognize office sounds for the evaluation (phone, voices, door opening/closing and door knocking). During the study, users monitored the peripheral display while checking email in a relaxed setting without any particular tasks. The feedback gathered from this session was qualitative in nature and enabled us to explore situations of importance to each user.

IC2Hear can support primary activities. If we had a Primary usage situation that we particularly wanted to study, we would conduct a dual-task lab study to isolate learning, awareness, distraction, and effects of breakdowns metrics.

IC2Hear shows information ranging from low to high criticality. The highly critical information is prioritized in our design and in user feedback. Users emphasized that information should be readily available since it could be very important: they need to "identify WHAT noise is being made and if it's an important noise… I would prefer some sort of identification of the noise that is being made." Also, users indicated many more critical sounds of which they wanted greater awareness than they did non-critical sounds (*e.g.,* they emphasized emergency alarms, intruders, traffic, people approaching from behind). Thus, our designs favor making users easily *aware* of sounds over being *aesthetically* pleasing. Our evaluation results show that all of our displays enabled easy identification of sounds, satisfying this requirement.

It is important that critical information displays appropriately *distract* users. Since the criticality of information shown on IC2Hear varies, users made it clear that they wanted overt notifications for very critical sounds and non-distracting updates for less critical sounds. They suggested that the icon flash when the sound was "really important," and be more "visual[ly] quiet" for less important sounds. We found that Single Icon and Spectrograph with Icon did not adequately distinguish between sounds of varying criticality. For this reason, users favored the History Display, which emphasized loud sounds (which users told us were more important) by using larger bars.

Displays of critical information should be easy to *learn*. Users were skeptical about using the Spectrograph because they thought it was too difficult to learn. Users thought Single Icon and the History Display were easy to learn, preferring these two displays over Spectrograph with Icon.

Again, displays of highly critical information should make *breakdowns* easy to discover. Single Icon and Spectrograph with Icon make breakdowns apparent since they are constantly changing with new sounds. If unsure, users can easily test the display by clapping their hands loudly and watching the display change. The History Display, however, only shows recognized sounds (phone, voices, door opening/closing, door knocking) which happen less often. This could impede users from noticing a system breakdown. One possible solution would be to display unrecognized sounds on the History Display as well.

Bus Mobile

The Bus Mobile is a physical display of bus arrival information (see Figure 4) [13]. It includes 6 physical tokens, each representing a bus. A token stays under a white curtain until its bus is 25 minutes from the bus stop, at which time it lowers 25 inches. With each minute, the token moves up 1 inch, indicating that the bus is getting closer to the stop. When the token reaches the white curtain, the bus has arrived.

Figure 4: Bus Mobile.

In terms of our design dimensions, the Bus Mobile is an opposite of IC2Hear. The Bus Mobile has a scope of 1 since it is meant to support one activity:

determining when to leave to catch the bus. Bus information is non-critical. Monitoring bus arrivals is secondary, since few primary activities require bus schedule information.

Because it has a low scope, the Bus Mobile was evaluated for the one secondary activity it supports: monitoring bus arrivals to know when to catch the bus. As such, a field study was conducted in order to observe Bus Mobile usage simultaneously with a wide range of primary activities.

The Bus Mobile was deployed in an undergraduate computer lab for two and a half weeks. Computer login/out data was collected to determine if user departures corresponded to the bus schedules. Also, 60 students were surveyed before, during, and after the deployment. Results enabled researchers to better understand the impact of the Bus Mobile on user behavior in a realistic setting.

The Bus Mobile shows non-critical information. Therefore, its *appeal* to users is important to their adoption. Surveying users after real usage enabled them to give grounded feedback. (Bus riders liked the display and missed it after the deployment.) Metrics like awareness and learnability should match user expectations, information that was gathered in both formative and summative qualitative surveys. Since bus arrival information is non-critical, the display should not be distracting. Results showed that the display was not distracting, but was difficult to learn. Though informing users of breakdowns is less critical for a display of non-critical information, it is still an issue for designers to support. The Bus Mobile often does not adequately inform users of breakdowns. For example, if several buses are not scheduled on weekends their tokens will never leave the white curtain. Are the tokens under the white curtain always not scheduled, or are the motors that lower them broken?

CONCLUSION

In this paper, we introduced a descriptive theory of peripheral displays based on Activity Theory. Specifically, a peripheral display is any information display that is (1) a *tool* in at least one activity of its user and (2) is used at the *operation* level. We also described an approach for designing and evaluating peripheral displays. From our descriptive theory of peripheral displays we derived a small set of important peripheral display design dimensions: *scope*, *class(es) of supported activities*, and *criticality*. We then described a set of evaluation metrics for peripheral displays – *appeal*, *learnability*, *awareness*, *effects of breakdowns*, and *distraction* – and how to select and measure these metrics. Finally, we presented two case studies as concrete examples of how these evaluation metrics might be employed. We hope that this work contributes toward a common language, design approach, and evaluation process amongst peripheral display designers.

REFERENCES

1. Carter, S., Mankoff, J., Goddi, P. “hebb: an infrastructure for conveying shared interests.” Technical Report No. CSD-03-1259, EECS Department, University of California, Berkeley, 2003.
2. Collins, P., Shukla, S., Redmiles, D. “Activity theory and system design: a view from the trenches.” In *Proceedings of CSCW,* 11. pp. 55-80, 2002.
3. Engeström, Y., Miettinen, R., Punamäki, R-L. (eds.) *Perspectives on Activity Theory*. Cambridge University Press, 1999.
4. Gaver, W.W., Smith, R.B., O’Shea, T. “Effective sounds in complex systems: The ARKola simulation.” In *Proceedings of CHI*, pp.85-90, 1991.
5. Hallnäs, L. and Redström, J. “Slow technology – designing for reflection.” In *Journal of Personal Ubiquitous Computing*, 5. 3:201-212, 2001.
6. Halverson, C. “Activity theory and distributed cognition: Or what does CSCW need to DO with theories?” In *Proceedings of CSCW,* 11. pp. 243-267, 2002.
7. Ho-Ching, F.W., Mankoff, J., Landay, J.A. “Can you see what I hear? The design and evaluation of a peripheral sound display for the deaf.” In *Proceedings of CHI*, pp. 161-168, 2003.
8. Hsieh, G. and Mankoff, J. “A Comparison of two peripheral displays for monitoring email: Measuring usability, awareness, and distraction.” Technical Report No. CSD-03-1286, EECS Department, University of California, Berkeley, 2003.
9. Kaye, J. “Symbolic olfactory display.” Master’s Thesis, MIT Media Arts and Sciences, 2001.
10. Kuutti, K. “Activity theory as a potential framework for human-computer interaction research.” In *Context and Consciousness: Activity Theory and Human-Computer Interaction*. MIT Press, pp. 17-44, 1995.
11. Leontiev, Aleksie N. *Activity, Consciousness, and Personality*. Prentice-Hall, 1978.
12. Maglio, P.P. and Campbell, C.S. “Tradeoffs in displaying peripheral information.” In *Proceedings of CHI*. pp. 241-248, 2000.
13. Mankoff, J.. Dey, A.K., Hsieh, G., Kientz, J., Lederer, S., Ames, M. “Heuristic evaluation of ambient displays.” In *Proceedings of CHI*, pp. 169-176, 2003.
14. Mankoff, J. and Dey, A.K. “From conception to design: A practical guide to designing and evaluating ambient displays.” In O'Hara, K. and Churchill, E. (eds.) *Public and Situated Displays*. Kluwer, 2003.
15. Matthews, T. “Interviewing peripheral display designers and developers.” Technical Report No.

EECS-2005-19, EECS Department, University of California, Berkeley, 2005.

16. Matthews, T., Fong, J., Mankoff, J. "Visualizing non-speech sounds for the deaf." In *Proceedings of ASSETS*, pp. 52-59, 2005.

17. McCrickard, D.S. and Zhao, Q.A. "Supporting information awareness using animated widgets." In *USENIX Technical Program*, pp. 117-127, 2000.

18. Mynatt, E.D., Back, M., Want, R., Baer, M., Ellis, J.B. "Designing audio aura." In *Proceedings of CHI*, pp. 566-573, 1998.

19. Mynatt, E.D., Rowan, J., Craighill, S., Jacobs, A. "Digital family portraits: Providing peace of mind for extended family members." In *Proceedings of CHI*, pp. 333-340, 2001.

20. Nardi, B. (ed.) *Context and Consciousness: Activity Theory and Human-Computer Interaction*. MIT Press, 1996.

21. Pedersen, E. R., and Sokoler, T. "AROMA: Abstract representation of presence supporting mutual awareness." In *Proceedings of CHI*, pp. 51-58, 1997.

22. Plaue, C., Miller, T., Stasko, T. "Is a picture worth a thousand words? An Evaluation of Information Awareness Displays." To appear in *Proceedings of Graphics Interface*, pp. 117-126, 2004.

23. Redström, J., Skog, T., Hallnäs, L. "Informative art: using amplified artworks as information displays." In *Proceedings of DARE*, pp. 103-114, 2000.

24. Robertson, G. Horvitz, E., Czerwinski, M., Baudisch, P., Hutchings, D.R., Meyers, B., Robbins, D., Smith, G. "Scalable Fabric: flexible task management." In *Proceedings of AVI*, pp. 85-89, 2004.

25. Star, S. L. and Ruhleder, K. "Steps towards an ecology of infrastructure: complex problems in design and access for large-scale collaborative systems," In *Proceedings of CSCW*, pp. 253-264, 1994.

26. Weiser, M. and Brown, J.S. "Designing calm technology." *PowerGrid Journal, 1*, 1, July 1996.