

# Reality-Based Interaction: A Framework for Post-WIMP Interfaces

Robert J.K. Jacob  
Orit Shaer

Audrey Girouard  
Erin Treacy Solovey

Leanne M. Hirshfield

Michael S. Horn

Jamie Zigelbaum†

Tufts University  
Department of Computer Science  
161 College Ave.  
Medford, MA 02155 USA  
{jacob, agirou01, lmille10, mhorn01,  
oshaer, etreac01}@cs.tufts.edu

†MIT Media Lab  
Tangible Media Group  
20 Ames St.  
Cambridge, MA 02139 USA  
zig@media.mit.edu

## ABSTRACT

We are in the midst of an explosion of emerging human-computer interaction techniques that redefine our understanding of both computers and interaction. We propose the notion of *Reality-Based Interaction* (RBI) as a unifying concept that ties together a large subset of these emerging interaction styles. Based on this concept of RBI, we provide a framework that can be used to understand, compare, and relate current paths of recent HCI research as well as to analyze specific interaction designs. We believe that viewing interaction through the lens of RBI provides insights for design and uncovers gaps or opportunities for future research.

## Author Keywords

Reality-Based Interaction, interaction styles, virtual reality, ubiquitous computing, tangible interfaces, next-generation, multimodal, context-aware, post-WIMP interfaces.

## ACM Classification Keywords

H.5.2 [Information Interfaces and Presentation]: User Interfaces

H.1.2 [Models and Principles]: User/Machine Systems — human factors;

## INTRODUCTION

Over the past two decades, HCI researchers have developed a broad range of new interfaces that diverge from the "window, icon, menu, pointing device" (WIMP) or Direct Manipulation interaction style. Development of this new generation of post-WIMP interfaces has been fueled by advances in computer technology and improved

understanding of human psychology. Defined by van Dam as interfaces "containing at least one interaction technique not dependent on classical 2D widgets such as menus and icons" [51], some examples of these post-WIMP interaction styles are: virtual, mixed and augmented reality, tangible interaction, ubiquitous and pervasive computing, context-aware computing, handheld, or mobile interaction, perceptual and affective computing as well as lightweight, tacit or passive interaction. Although some may see these interaction styles as disparate innovations proceeding on unrelated fronts, we propose that they share salient and important commonalities, which can help us understand, connect, and analyze them.

We believe that all of these new interaction styles draw strength by building on users' pre-existing knowledge of the everyday, non-digital world to a much greater extent than before. They employ themes of reality such as users' understanding of naïve physics, their own bodies, the surrounding environment, and other people. They thereby attempt to make computer interaction more like interacting with the real, non-digital world. By drawing upon these themes of reality, emerging interaction styles often reduce the gulf of execution [24], the gap between a user's goals for action and the means to execute those goals. We propose that these emerging interaction styles can be understood together as a new generation of HCI through the notion of *Reality-Based Interaction* (RBI). We believe that viewing interaction through the lens of RBI might provide insights for design and uncover gaps or opportunities for future research.

In this paper, we introduce a framework that unifies emerging interaction styles and present evidence of RBI in current research. We discuss its implications for the design of new interfaces and conclude by applying RBI to the analysis of four case studies.

## REALITY-BASED INTERACTION

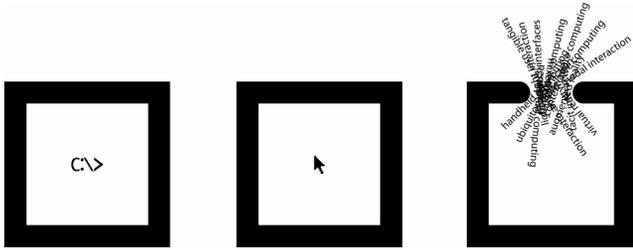
Interaction with computers has evolved from the first generation of Command Line, to the second generation of Direct Manipulation, to a new generation of emerging post-

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CHI 2008, April 5–10, 2008, Florence, Italy.

Copyright 2008 ACM 978-1-60558-011-1/08/04...\$5.00.

WIMP interaction styles. Direct Manipulation moved interfaces closer to *real world* interaction by allowing users to directly manipulate objects rather than instructing the computer to do so by typing commands. New interaction styles push interfaces further in this direction. They increase the realism of interface objects and allow users to interact even more directly with them—using actions that correspond to daily practices within the non-digital world.



**Figure 1. Generations of interaction: command line, direct manipulation, and diverse emerging interaction styles [26].**

### RBI Themes

We use the term “real world” to refer to aspects of the physical, non-digital world. However, the terms *real world* and *reality* are problematic and can have many additional, interpretations, including cultural and social reality. For that matter, many would also consider keyboards and mice to be as much a part of today's reality as any non-digital artifact. Thus, to clarify, our framework focuses specifically on the following four themes from the real world:

- *Naïve Physics*: people have common sense knowledge about the physical world.
- *Body Awareness & Skills*: people have an awareness of their own physical bodies and possess skills for controlling and coordinating their bodies.
- *Environment Awareness & Skills*: people have a sense of their surroundings and possess skills for negotiating, manipulating, and navigating within their environment.
- *Social Awareness & Skills*: people are generally aware of others in their environment and have skills for interacting with them.

To a greater extent than in previous generations, these four themes play a prominent role in emerging interaction styles. They provide a basis for interaction with computers that is markedly closer to our interaction with the non-digital world. While we believe these themes apply to most people and most cultures, they may not be entirely universal.

In the remainder of this section we describe these themes in more detail. In the *Case Studies* section we show how these four themes can be applied to analyze the design of current post-WIMP interfaces.

### Naïve Physics (NP)

Naïve physics is the informal human perception of basic physical principles, or in other words, common sense

knowledge about the physical world. This includes concepts like gravity, friction, velocity, the persistence of objects, and relative scale. In the field of artificial intelligence, naive physics refers to an attempt to formally describe the world as most people (rather than physicists) think about it [21]. In the context of emerging interaction styles, user interfaces increasingly simulate or directly use properties of the physical world. For example, a tangible user interface (TUI) may employ physical constraints such as a rack or a slot to guide the way in which physical tokens can be manipulated (e.g. [48]). Emerging graphical user interfaces, such as the Apple iPhone (see case study below), employ physical metaphors that add the illusion of gravity, mass, rigidity, springiness, and inertia to graphical widgets.

### Body Awareness and Skills (BAS)

Body awareness refers to the familiarity and understanding that people have of their own bodies, independent of the environment. For example, a person is aware of the relative position of his or her limbs (proprioception), his or her range of motion, and the senses involved in perceiving certain phenomena. Early in life, most people also develop skills to coordinate movements of their limbs, head, eyes, and so on, in order to do things like crawl, walk, or kick a ball. Emerging interfaces support an increasingly rich set of input techniques based on these skills, including two-handed interaction and whole-body interaction. For example, many emerging virtual reality applications allow users to move from one place to another within a virtual environment simply by walking on a special track or treadmill (e.g. [34]).

### Environment Awareness and Skills (EAS)

In the real world, people have a physical presence in their spatial environment, surrounded by objects and landscape. Clues that are embedded in the natural and built environment facilitate our sense of orientation and spatial understanding. For example, a horizon gives a sense of directional information while atmospheric color, fog, lighting, and shadow provide depth cues [10]. People develop many skills for navigating within and altering their environment.

In the context of emerging interaction styles, many virtual reality (VR), mixed reality (MR), and augmented reality (AR) interfaces along the reality-virtuality continuum [33] use reference objects and artificial landmarks to provide users with clues about their virtual environment and simplify size and distance estimations in that environment [52]. Furthermore, by representing users' bodies in the virtual world, VR interfaces allow users to perform tasks relative to the body (egocentric). Context aware or sensing systems may compute users location and orientation, and display information that corresponds to the user's position in physical space [8, 9].

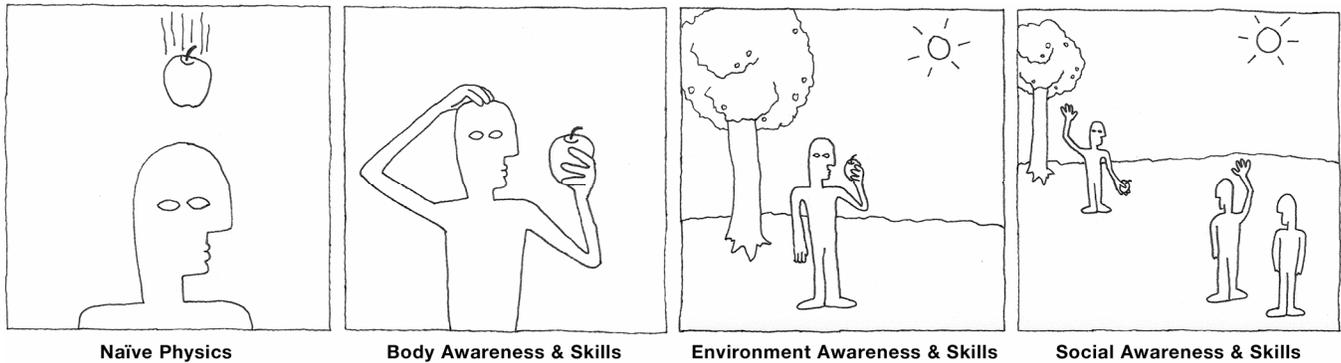


Figure 2. The four RBI themes.

People also develop skills to manipulate objects in their environment, such as picking up, positioning, altering, and arranging objects. Emerging interaction styles often draw upon users' object manipulation skills. For example, in VR and in TUIs, users often select an object by grasping it, either virtually or physically. Many lightweight or tacit interfaces also track manipulation of objects. (Some of these object manipulations draw from the naïve physics and body awareness and skills themes as well.)

#### Social Awareness and Skills (SAS)

People are generally aware of the presence of others and develop skills for social interaction. These include verbal and non-verbal communication, the ability to exchange physical objects, and the ability to work with others to collaborate on a task. Many emerging interaction styles encourage both social awareness and remote or co-located collaboration. For example, TUIs provide both the space and an array of input devices to support co-located collaboration. Virtual environments (e.g. Second Life [38]) exploit social awareness and skills by representing users' presence, by displaying their avatars, and by making the avatars' actions visible.

#### EVIDENCE FOR RBI IN CURRENT RESEARCH

Having introduced the four themes of RBI, we present evidence to support our claim that they can be seen in much post-WIMP interaction research. We examine published literature, discussions from a CHI 2006 workshop, and interviews conducted during an informal field study for evidence of the RBI framework.

#### Survey of Published Literature

With the view of RBI as a unifying thread for emerging research, one can go back through the literature and identify many examples of designers following RBI principles. We highlight some examples that we have found in recent papers. This is not a comprehensive list; there are numerous other examples along these lines. While the interface designs described below are diverse, they demonstrate design choices that implicitly adhere to the RBI themes.

Some papers comment on broader categories of emerging interaction styles and acknowledge the value of leveraging many aspects of the real world in developing new HCI systems. For example, Jackson and Fagan claim that by

leveraging our highly tuned perceptual abilities, VR attempts to "allow for direct manipulation of objects in the virtual environment (VE) using hands, body movement, or through virtual tools that allow participants to observe and interact within the VE as naturally as they would interact with objects in the real world" [25] (**examples of BAS, EAS**). Ullmer, Ishii, and Jacob note that tangible interactions are "observable with both visual and haptic modalities and draw on some of humans' most basic knowledge about the behavior of the physical world" [48] (**NP**). Discussing ubiquitous computing, Abowd explains, "it is the goal of these natural interfaces to support common forms of human expression and leverage more of our implicit actions in the world" [1] (**SAS, EAS**). Taking this a step further, he writes, "Humans speak, gesture, and use writing utensils to communicate with other humans and alter physical artifacts. These natural actions can and should be used as explicit or implicit input to ubicomp systems" (**SAS, EAS, BAS**). Streitz agrees that the real world around us should be the starting point for designing the human-computer interaction of the future [44].

Other recent papers focus on specific new devices or interaction styles and note inspirations from a particular aspect of the real world, corresponding with the RBI themes mentioned above.

**NP** concepts were included in several recent papers. Poupyrev, Newton-Dunn and Bau use the metaphor of air bubbles when working with a multifaceted display. "Imagine a glass ball filled with water with a small bubble of air inside. As we rotate the ball the air bubble moves to the top. If the bubble is a pointer and there are images on the surface of the ball, i.e. icons, we can select one by aligning the bubble with the icon" [36]. Forlines and Shen explain that: "by moving their two fingers apart diagonally, the user controls the zoom level of the lens visualization . . . The amount of zoom is calculated to give the appearance that the tabletop is stretching under the user's fingers. There is an illusion of a pliable rubber surface" [20]. Apted explains the affordances of the Tabletop system in terms of **NP**: "Some of these affordances are derived from equivalent, purely physical interactions that occur with printed photographs . . . To maintain the link with the

physical world, users interact only with photographs—there are no buttons, menus or toolbars to be navigated” [4].

Many other systems draw upon **BAS**. For example, Ångeslevä, et al. conceived an information storage system for portable devices “designed to ease cognitive load by relying on our proprioceptive sense and the use of the body image of the user as a mnemonic frame of reference” [3]. Buur, Jensen, and Djajadiningrat proposed two innovative design methods for tangible interaction that employ a greater range of human actions. They claim that: “Currently, the actions required by electronic products are limited to pushing, sliding and rotating. Yet humans are capable of far more complex actions: Human dexterity is highly refined. This focus on actions requires a reconsideration of the design process.” [11].

New interfaces also take advantage of **EAS**. For example, a group working with a new graspable handle with a transparent glove noted that the “graspable handle enables the user to perform a holding action naturally—the most basic action when physically handling a curved shape in the real world” [5]. Vogel and Balakrishnan specifically mention the benefits of reality: “When a display surface can sense touch, selecting items by tapping with your finger or a pen is immediately appealing, as it mimics real world interaction” [53].

Finally, many emerging interfaces were designed with **SAS** concepts in mind. Dickie, et al. explain that: “In eyeLook, we modeled our design strategy on the most striking metaphor available: that of human group communication ... By incorporating eye contact sensing into mobile devices, we give them the ability to recognize and act upon innate human nonverbal turn taking cues” [15]. Smith, Vertegaal, and Sohn use similar justification for their design: “users are also very familiar with the use of their eyes as a means for selecting the target of their commands, as they use eye contact to regulate their communications with others” [43]. Apted, Kay, and Quigley also employ **SAS** in their design: “The nature of a tabletop interface makes it very natural to use in a social setting with two or more people” [4].

In summary, researchers often leverage users’ knowledge and skills of interaction with the real world. We observed that this knowledge includes naïve physics, as well as body, environment and social awareness and skills. Although the researchers did not explicitly refer to these reality-based themes, they made design choices reflecting the same principles.

#### **CHI 2006 Workshop**

We found another source of supporting evidence for the RBI concept from a workshop we conducted at the CHI 2006 conference. Entitled “What is the Next Generation of Human-Computer Interaction?”, the workshop brought together researchers from a range of emerging areas in HCI to look for common ground and a common understanding of a next generation human-computer interaction style [27]. It began with the same basic questions that we are trying to solve here, but left the answers open to input from the

participants. In reviewing the discussions and breakout groups, we looked for ideas that support or contradict our notion of reality-based interaction.

We observed general agreement that the focus is shifting away from the desktop and that technology is moving into new domains. We also found that many of the commonalities that the breakout groups identified were closely connected to reality-based interaction, for example, exploiting users’ existing knowledge about different materials and forms to communicate syntax. In a wrap-up session discussing RBI, we found good support for the reality-based interaction concept but expressed in a variety of different terminologies, and with some new ideas and dimensions added to it [27].

#### **Informal Field Study**

We also interviewed a few researchers about their design processes in an informal field study. The interviews were done with a dozen graduate students from the Media Lab at MIT. None of the interviewees had been exposed to the RBI concepts before the interview. Two interesting examples are discussed below.

James Teng from the Ambient Intelligence Group discussed the Engine-Info project [32], an educational car engine with transponders at different key places. The user hears descriptions and explanations through a Bluetooth audio earpiece based on the inferred direction that the user is looking. James described the interaction of the user as being “more intuitive” since the user already knows how to indicate parts of interest (by simply focusing one’s gaze) (**BAS**). This work takes “advantage of the physical world” since you can walk around, understand the scale, and directly see how various components are connected (**EAS**).

From the Object-Based Media Group, Jeevan Kalanithi described the design rationale behind Connectibles [28], a tangible social networking interface. He noticed that people have meaningful social behaviors established in the real world, such as gift giving. Connectibles is an “imitation of social reality”—people must physically give the object (a customizable interactive keepsake) to another person for that person to collect it and interact with it (**SAS**). He noted that this contrasts with many online social networking interfaces, in that the physical objects are more scarce and expensive than digital ones, perhaps resulting in Connectibles displaying a better representation of a person’s close relationships.

#### **IMPLICATIONS FOR DESIGN**

Up to this point, we have claimed and presented some evidence that the themes of RBI are a good characterization of key commonalities among emerging interaction styles. In considering the implications of the RBI framework for design, we further suggest that the trend towards increasing reality-based interaction is a positive one. Basing interaction on pre-existing real world knowledge and skills may reduce the mental effort required to operate a system because users already possess the skills needed. For casual use, this reduction might speed learning. For use in

situations involving information overload, time pressure, or stress, this reduction of overhead effort may improve performance. Applying RBI concepts such as naïve physics to an interface design may also encourage improvisation and exploration because users do not need to learn interface-specific skills.

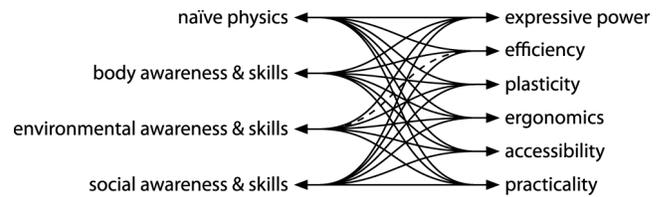
However, simply making an interface as reality-based as possible is not sufficient. A useful interface will rarely entirely mimic the real world, but will necessarily include some unrealistic or artificial features and commands. In fact, much of the power of using computers comes from this multiplier effect—the ability to go beyond a precise imitation of the real world. For example, in a GUI, one might want to go beyond realistically pointing to and dragging individual files to more abstract commands like *Archive all files older than 180 days* or *Delete all files that contain the text string “reality-based”* [37, 41]. Designers must strike a balance between the power of their interface and its level of reality. Many designers make these decisions implicitly in their work (seen in the review of current research above). The RBI framework makes these design tradeoffs explicit and provides explanatory power for understanding the costs and benefits of such decisions.

### Tradeoffs

As noted above, mimicking reality alone is not enough; there are times when RBI principles should be traded against other considerations. We propose that the goal is to give up reality only explicitly and only in return for other desired qualities, such as:

- *Expressive Power*: i.e., users can perform a variety of tasks within the application domain
- *Efficiency*: users can perform a task rapidly
- *Versatility*: users can perform many tasks from different application domains
- *Ergonomics*: users can perform a task without physical injury or fatigue
- *Accessibility*: users with a variety of abilities can perform a task
- *Practicality*: the system is practical to develop and produce

These qualities are discussed below as tradeoffs and they are key to our analysis of the case studies in the next section. Figure 3, further illustrates these tradeoffs. Note that while the RBI framework explicitly highlights design tradeoffs, it does not provide a structured methodology for discussing these tradeoffs. Rather, employing the Design Space Analysis [30] methodology while discussing these tradeoffs can help with structuring the discussion and comparing alternative options. In terms of the Questions Options and Criteria (QOC) notation used for representing a Design Space Analysis, the principles we discuss below can be used to form questions.



**Figure 3. RBI design tradeoffs. As a specific example, the dotted line shows the tradeoff between the reality EAS theme and efficiency, which occurs in Case Study 4.**

### Reality vs. Expressive Power

The expressive power, or functionality, of a system is often seen as its most important contribution, although it is notable that more features do not always result in a better system—feature creep can make an interface too difficult, complex, or unwieldy to use [31]. In some cases it is better to privilege the expressive power of a system; in other cases it is better to limit functionality in favor of realism. For example, in the Bumptop system [2] documents and files are arranged in a virtual, three-dimensional space. These objects can be stacked, shelved, crumpled, and tossed around in a virtual room, but they cannot be placed in a complex tree structure of folders. This places a limit on the expressive power of the system, giving up possible functionality in order to maintain the clear virtual 3D space.

### Reality vs. Efficiency

While walking is not usually as fast as driving, sometimes it is preferable to use skills that are as easy as walking rather than privileging efficiency. This tradeoff is clear when examining the differences between systems designed for expert and novice. An expert video editor will often heavily rely on hotkeys (keyboard shortcuts for common features). For experts, it is important that the interface allow them to access commands very quickly. For a novice video editor, an interface with a reality-based design, such as the Tangible Video Editor [56] that allows video clips to be put together like a jigsaw puzzle, may be preferable.

### Reality vs. Versatility

A single GUI based system can be used to perform a variety of tasks such as editing films, writing code, or chatting with friends. On the other hand, a TUI system, such as the Tangible Video Editor [56] only lets you complete a single task, edit video clips, while allowing for a higher degree of realism.

### Reality vs. Ergonomics

Repetitive stress injuries and fatigue can be caused by some interfaces [54]. In designing a new interface it is important to consider ergonomics and these considerations may oppose RBI goals.

### Reality vs. Accessibility

Realistic actions may not be ideal for the disabled. There are many cases when reliance on strict realism can prevent some users from interaction, making the interface less accessible. In this case, the use of less realistic abstractions and tools may be preferable.

### Reality vs. Practicality

Practical matters such as cost, technological limitations, space, size, durability, power consumption, and environmental impact are also relevant for designers and may be traded off against realism.

### Summary

Figure 3 displays some of the tradeoffs that may be considered throughout the design of an RBI interface. It shows the RBI themes, on the left side, traded against the qualities described above, on the right side. We propose a view that identifies some fraction of a user interface as based on the RBI themes plus some other fraction that provides computer-only functionality that is not realistic. As a design approach or metric, the goal would be to make the first category as large as possible and use the second only as necessary, highlighting the tradeoff explicitly [45].

For example, consider the character Superman. He walks around and behaves in many ways like any human. He has some additional functions for which there is no analogy in humans, such as flying and X-ray vision. When doing realistic actions, he uses his real world commands—walking, moving his head, and looking around. But he still needs some additional non real world commands for flying and X-ray vision, which allow him to perform tasks in a more efficient way, just like a computer provides extra power. In the design of a reality-based interface we can go a step further and ask that these non real world commands be analogous to some realistic counterpart. For example, in a virtual reality interface, a system might track users' eye movements, using intense focus on an object as the command for X-ray vision [46].

We might further divide the non-realistic part of the interface into degrees of realism (x-ray by focus vs. by menu pick). The designer's goal should be to allow the user to perform realistic tasks realistically, to provide additional non real-world functionality, and to use analogies for these commands whenever possible. We should use a conventional walking gesture to walk—unless using a less natural command would provide extra power (e.g. speed, automatic route finding). The designer should not give up the reality of the walking command lightly, not without gaining some added efficiency.



**Figure 4. Superman walks normally, but uses additional non real-world commands to provide extra functionality.**

## CASE STUDIES

In the following case studies, we show examples of designs that both adhere to each of our RBI themes and stray from these themes to make a tradeoff. Our analysis makes explicit the idea that giving up some reality may be worthwhile only when some other value is gained.

### Case Study 1: URP

URP [49] is a tangible user interface (TUI) for urban planning that allows users to place models of buildings on an interactive surface collaboratively (Figure 5). URP responds by overlaying digital information onto the surface such as a representation of the shadows that buildings will cast at different times of day, the pattern of wind as it flows through the urban space, and the glare caused by the sun reflecting off different building surface materials. URP also provides a collection of physical tools for manipulating environmental conditions such as time of day and wind speed. We have selected URP as a case study because it is one of the most fully developed and widely known TUI systems. It also serves as an example for a large class of TUIs that support the spatial manipulation of discrete physical objects on an interactive surface.

### RBI Themes

URP's defining feature is that it represents the application domain (urban architecture and planning) with an actual physical model rather than an image on a computer screen. The basic interaction techniques (add, remove, and position models) build directly on users' knowledge of naïve physics (**NP**) and physical space (**EAS**). To add a building to the system, a user simply picks up a model and places it on the surface. There is no menu system, no series of clicks and drags, and no indirect input device to negotiate. A secondary interaction technique involves inspecting the urban model from different points of view (e.g. birds-eye view, street-level view). This task is notoriously difficult with graphical user interfaces. With URP, the user simply moves his or her body to change viewpoints. The inspection task directly leverages users' knowledge of their own bodies and their ability to move their bodies to different positions in an environment (**BAS**, **EAS**), and it relies on users' understanding of relative scale (**NP**). Furthermore, the physical model created by users is persistent in the real world (**NP**) and continues to exist even when the system is turned off. Finally, URP encourages collaboration between multiple co-located users. Users do not need to share a single input device or crowd around a computer screen. In this way, URP allows users to draw more directly on their existing social interaction skills (**SAS**).

### Tradeoffs

To enhance URP's functionality, its designers added tools to adjust parameters such as time of day, wind direction, and building material. For example, to change a building's material, users must touch the model with a special material wand. This feature adds expressive power to the system, but it also violates the simple physical metaphor of positioning

building models. In this sense, the designers made a tradeoff between *reality and expressive power*.

Furthermore, while it is easy for users to do some things with URP (such as position buildings), it is difficult or impossible for users to do other things (like changing the shape of a building on the fly or adding a new building to the system that was not part of the initial set of models). In this sense, the designers made a tradeoff between *reality and practicality* and between *reality and expressive power*, this time in favor of reality.

### Case Study 2: Apple iPhone

The iPhone illustrates how commercial designers are evolving interfaces by incorporating insights and techniques gleaned from HCI research.

#### RBI Themes

One of the iPhone's features is a multitouch display (Figure 5). A technology that has been around for decades in research circles [12], multitouch sensing is used to create applications that are based on naïve physics (**NP**). In the photograph viewing application, zoom functions that would traditionally be accessed through combo boxes, button presses, or other widgets are instead activated by pinching and stretching the iPhone's display with two fingers using the illusion of a pliable rubber surface [20] (**NP**).

While viewing a photo in full screen mode, the user flicks the screen to the left or right to see the next image in the list rather than pressing a directional pad or using a jog wheel. This uses environmental awareness and skills (**EAS**) via a spatial metaphor—all objects in the real world have spatial relationships between them. Similar use of **EAS** is also found in iPhone applications such as *Weather*, *iPod*, and *Safari*.

The iPhone applications of *iPod*, *Safari*, *Phone* and *Photos* also use body awareness and skills (**BAS**) in their interaction design. When the user puts the phone close to his or her face, it shuts off the screen to prevent accidental button presses. The other three applications use inertial sensing to orient displayed content so that when the iPhone is placed in landscape or portrait view, the image is always right side up (**NP**).

**NP** in the form of inertia and springiness is found across almost all of the iPhone's applications. When scrolling to the bottom of an email, the window appears connected to the bottom of the screen as if by springs. When flicking through the contact list, a fast flick will keep the contacts

scrolling after the user's finger has been removed, as if the list itself had mass.

#### Tradeoffs

The designers of the iPhone utilize reality-based interaction themes throughout the phone's interface, sometimes favoring reality and sometimes favoring other interaction designs.

Rather than using the hardware keyboards that have become ubiquitous on smartphones, the designers stretched the iPhone's display to cover almost the entire front surface, allowing for the use of a graphical keyboard to input text. Here the designers have decided to favor *versatility over reality*—reality in this case being the persistence of physical buttons (**NP**) for text entry. This tradeoff allows the iPhone to have a much larger screen but sacrifices the passive haptics associated with hard buttons and the persistence that can allow for efficiency gains and lower cognitive load.

Another tradeoff is found in the *Phone* application. On many mobile phones, a user can enter the first few letters of a contact's name and press send to place calls. This feature is missing on the iPhone. A similar tradeoff is made by the removal of search functionality for contacts, emails, and songs, artists, or albums. The designers have favored *reality over efficiency*—here they have sacrificed speed gains to strengthen the application's reliance on **NP**. While strengthening the reality-based interaction design, these decisions may also be influenced by the lack of a hard keyboard.

In the *Safari* browser, web pages are displayed in their full form rather than reformatted in a column view as found in most other phones with browsers. This makes it more difficult for people with poor eyesight to read. This is a tradeoff of *reality over accessibility* in favor of reality. The designers have favored a reliance on direct representation of reality applying **NP** and **EAS**.

### Case Study 3: Electronic Tourist Guide

The GSSystem [7] (GSS) is an electronic tourist guide system that displays pictures of buildings surrounding the user on a hand-held device. As a context-aware example, GSS represents an interface that is location (**EAS**) and orientation (**BAS**) aware. It “exploit[s] knowledge about the physical structure of the real world” [7] to compute what information is displayed to the user. More specifically, the system can calculate the location of a user (through a GPS

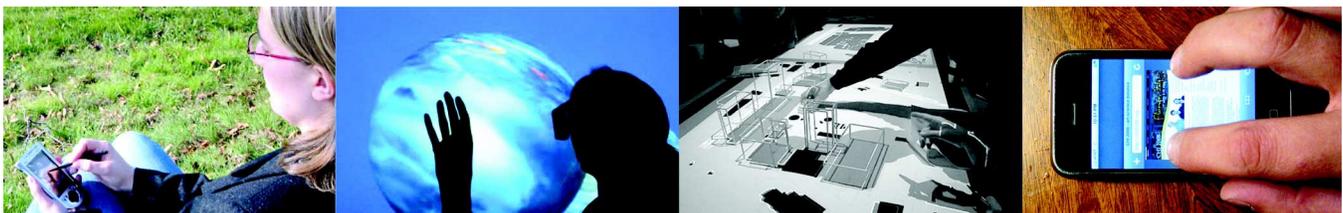


Figure 5. Examples of systems that use RBI themes. Left to right: a handheld device, a virtual reality system, URP [49], iPhone.

receiver if available), and determine the orientation of the user. To exploit the knowledge about the real world, GSS determines the surrounding buildings' geometry and calculates what elements are visible to the user. Outputs of the system include photos of buildings and objects the user may see from his or her location.

#### **RBI Themes**

This system exploits **EAS** and **BAS**: it uses the physical reality of the user, i.e. his or her current environment, to improve interaction with the interface. The user has an innate knowledge of how to input his or her location: he or she simply needs to go there.

#### **Tradeoffs**

This system has two modes, depending on how the location is specified: via GPS or button presses. The latter mode can lead to previewing an area. This is an example of the *reality vs. expressive power* tradeoff, since this mode breaks direct coupling to the local environment—you can now see what is not there. It maintains the strong reality metaphor only as long as the GPS mode is used.

#### **Case Study 4: Visual-Cliff Virtual Environment**

We examine Slater's visual-cliff virtual environment [42, 50] because it illustrates the core components of immersive VR. The system allows users to navigate through three rooms that contain furniture, windows, and doors as in a real world house.

#### **RBI Themes**

Interaction techniques in VR systems take advantage of reality-based concepts. Many systems enable users to pick up objects and place them in new locations (**NP**, **EAS**). The command for doing this is the same as in the real world, and the results are the same as in the real world (though often without haptic feedback). In addition, almost all systems incorporate body awareness (**BAS**) and environment awareness (**EAS**), some to a greater extent than others. Immersive virtual environments track head movements and use this information to update graphics in the system based on the user's new viewpoint. Many systems track other information such as hand movements, eye movements, and even full body movements. Based on this, the user can make real world motions and gestures that affect the virtual environment as expected.

#### **Tradeoffs**

Early methods for locomotion in virtual reality were hand gestures or leaning in the direction of movement. In the visual-cliff, several methods for locomotion were compared. In one study [42], Slater, Usoh and Steed found that walking in place was a better mode of locomotion than push-button-fly. This is an example of favoring *reality over efficiency*, as shown in Figure 3. In a subsequent study [50], they added real walking as a condition (even more realistic), and found that it was a further improvement over the other two methods.

These studies show that realistic interaction worked better in the virtual environment than interaction commands that

must be learned. However, there is a *reality versus practicality* tradeoff to these methods. Although the study showed that real walking performed better, it may be necessary to use the walking in place system if space is constrained.

#### **Summary**

These four case studies illustrate emerging interfaces that apply the RBI themes. Each system had inspirations from the real world, but also gave up reality when necessary to achieve additional design goals.

#### **RELATED TAXONOMIES AND FRAMEWORKS**

To date, work that attempts to explain or organize emerging styles of interaction has focused more on individual classes of interfaces than on ideas that unify several classes [16, 18, 19, 22, 35, 47]. Some work has focused more generally on new issues that are not present in interactions with traditional WIMP interfaces [8, 9, 14, 17, 29]. For example, Coutrix and Nigay[14] as well as Dubois and Gray[17] have developed interaction models for guiding designers of mixed reality systems (such as augmented reality, tangible systems, and ubiquitous applications). Other work has focused on specific new interaction styles [6, 23, 37, 55]. While previous work focuses on subsets of interaction styles, our RBI framework applies to a wider range of emerging interaction styles.

Finally, the work that helped define the GUI generation was an inspiration for our work. Shneiderman took a variety of what, at the time, seemed disparate new user interface inventions and brought them together by noting their common characteristics and defining them as a new generation of user interfaces [39]. Hutchins, Hollan and Norman went on to explain the power and success of these interfaces with a theoretical framework that provided a basic understanding of the new generation in human terms [24]. Our hope is to take the first step in that direction for the emerging generation of interaction styles.

#### **CONCLUSION**

We hope to advance the study of emerging interaction styles with a unifying framework that can be used to understand, compare and relate these new interaction styles. The reality-based interaction (RBI) framework characterizes a large subset of seemingly divergent research areas. The framework consists of four themes: naïve physics, body awareness and skills, environment awareness and skills, and social awareness and skills. Based on these themes, we show implications for the design and analysis of new interfaces.

Our framework is primarily a descriptive one. Viewing the emerging generation of interfaces through the lens of reality-based interaction provides researchers with explanatory power. It enables researchers to analyze and compare alternative designs [13], bridge gaps between seemingly unrelated research areas, and apply lessons learned from the development of one interaction style to another. It can also have a generative role [40] by

suggesting new directions for research, such as incorporating RBI themes in the design of interfaces for different user populations (e.g. children or expert users) or studying the effects of different degrees of RBI themes in an interface.

#### ACKNOWLEDGMENTS

We thank our collaborators Andrew Afram, Eric Bahna, Tia Bash, Georgios Christou, Michael Poor, Andrew Pokrovski, and Larissa Supnik in the HCI group at Tufts, as well as Caroline Cao and Holly Taylor of Tufts, Leonidas Deligiannidis of the University of Georgia, Hiroshi Ishii of the MIT Media Lab and the students in his Tangible Interfaces class, Sile O'Modhrain of Queen's University Belfast, and Frank Ritter of Pennsylvania State University.

We also thank the participants in our CHI 2006 workshop on "What is the Next Generation of Human-Computer Interaction?" for their thoughts and discussion about this area, which have helped us refine our work, and Ben Shneiderman in particular for discussions on this topic. And we thank Jeevan Kalanithi and James Teng of the MIT Media Lab for participating in our field study.

Finally, we thank the National Science Foundation for support of this research (NSF Grant Nos. IIS-0414389 and IIS-0713506) and the Natural Sciences and Engineering Research Council of Canada for support of this research. Any opinions, findings, and conclusions or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of these organizations.

#### REFERENCES

1. Abowd, G.D. and Dix, A.J., Integrating Status and Event Phenomena in Formal Specifications of Interactive Systems. in *Proc. SIGSOFT 1994*, (1994), Addison-Wesley/ACM Press.
2. Agarawala, A. and Balakrishnan, R. Keepin' it real: pushing the desktop metaphor with physics, piles and the pen *Proc. CHI2006*, ACM Press, 2006.
3. Ångeslevä, J., Oakley, I., Hughes, S. and O'Modhrain, S., Body Mnemonics Portable device interaction design concept. in *UIST*, (2003).
4. Apted, T., Kay, J. and Quigley, A. Tabletop sharing of digital photographs for the elderly *Proc. of SIGCHI*, ACM Press, 2006.
5. Bae, S.-H., Kobayash, T., Kijima, R. and Kim, W.-S., Tangible NURBS-curve manipulation techniques using graspable handles on a large display. in *Proc. UIST 2004*, (2004), ACM Press, 81-90.
6. Beaudouin-Lafon, M. Instrumental Interaction: An Interaction Model for Designing Post-WIMP User Interfaces *Proc. CHI 2000*, Addison-Wesley/ACM Press, 2000, 446-453.
7. Beeharee, A. and Steed, A. Exploiting real world knowledge in ubiquitous applications. *Personal and Ubiquitous Computing*, 11. 429-437.
8. Bellotti, V., Back, M., Edwards, W.K., Grinter, R.E., Henderson, A. and Lopes, C. Making Sense of Sensing Systems: Five Questions for Designers and Researchers *Proc. CHI 2002*, ACM Press, 2002, 415-422.
9. Benford, S., Schnadelbach, H., Koleva, B., Anastasi, R., Greenhalgh, C., Rodden, T., Green, J., Ghali, A., Pridmore, T., Gaver, B., Boucher, A., Walker, B., Pennington, S., Schmidt, A., Gellersen, H. and Steed, A. Expected, sensed, and desired: A framework for designing sensing-based interaction. *ACM Transactions Computer-Human Interaction*, 12 (1). 3-30.
10. Bowman, D.A., Kruijff, E., LaViola, J.J. and Poupyrev, I. *3D User Interfaces: Theory and Practice*. Addison-Wesley/Pearson Education, 2004.
11. Buur, J., Jensen, M.V. and Djajadiningrat, T., Please touch tangible UIs: Hands-only scenarios and video action walls: novel methods for tangible user interaction design. in *Proc. DIS'04*, (2004).
12. Buxton, B.  
<http://www.billbuxton.com/multitouchOverview.html>, 2007.
13. Christou, G. Towards a new method for the evaluation of reality based interaction *CHI '07 extended abstracts*, ACM, San Jose, CA, USA, 2007.
14. Coutrix, C. and Nigay, L. Mixed reality: a model of mixed interaction. *Proceedings of the working conference on Advanced visual interfaces*.
15. Dickie, C., Vertegaal, R., Sohn, C. and Cheng, D. eyeLook: using attention to facilitate mobile media consumption *Proc. UIST 2005*, ACM Press, 2005.
16. Dourish, P. *Where The Action Is: The Foundations of Embodied Interaction*, MIT Press, Cambridge, Mass., 2001.
17. Dubois, E. and Gray, P. A Design-Oriented Information-Flow Refinement of the ASUR Interaction Model. *Proceedings of EHCI 2007*.
18. Fishkin, K.P. A Taxonomy for and Analysis of Tangible Interfaces. *Personal and Ubiquitous Computing*, 8 (5). 347 - 358.
19. Fishkin, K.P., Moran, T.P. and Harrison, B.L. Embodied User Interfaces: Toward Invisible User Interfaces *Proc. of EHCI'98*, 1998.
20. Forlines, C. and Shen, C., Touch: DTLens: multi-user tabletop spatial data exploration. in *Proc. UIST 2005*, (2005).
21. Hayes, P.J. *The second naive physics manifesto Cognitive Science Technical Report URCS-10*, University of Rochester, 1983.
22. Hornecker, E. and Buur, J., Getting a Grip on Tangible Interaction: A Framework on Physical Space and Social Interaction. in *Proc. of CHI 2006*, (2006), ACM, 437-446.
23. Hurtienne, J. and Israel, J.H. Image schemas and their metaphorical extensions: intuitive patterns for tangible interaction *Proc. TEI 2007*, ACM Press, 2007.
24. Hutchins, E.L., Hollan, J.D. and Norman, D.A. *Direct Manipulation Interfaces*. Draper, D.A.N.a.S.W. ed.

- User Centered System Design: New Perspectives on Human-computer Interaction*, Lawrence Erlbaum, Hillsdale, N.J., 1986, 87-124.
25. Jackson, R.L. and Fagan, E., Collaboration and learning within immersive virtual reality. in *Proc. Collaborative Virtual Environments*, (2000), ACM Press, 83 - 92.
  26. Jacob, R.J.K., Girouard, A., Hirshfield, L.M., Horn, M.S., Shaer, O., Solovey, E.T. and Zigelbaum, J., Reality-Based Interaction: Unifying the New Generation of Interaction Styles (Work in Progress paper). in *CHI 2007*, (2007), 2465-2470.
  27. Jacob, R.J.K., Girouard, A., Horn, M., Miller, L., Shaer, O., Treacy, E. and Zigelbaum, J. What Is the Next Generation of Human-Computer Interaction? *interactions*, 2007.
  28. Kalanithi, J. Connectibles: Tangible Social Networking, MS Thesis, MIT Media Lab, Cambridge, 2007, 132pp.
  29. Klemmer, S.R., Hartmann, B. and Takayama, L., How bodies matter: five themes for interaction design. in *Proc. DIS 2006*, (2006), ACM Press, 140-149.
  30. MacLean, A., Young, R.M., Bellotti, V.M.E. and Moran, T.P. Questions, Options, and Criteria: Elements of Design Space Analysis. *Human-Computer Interaction*, 6. 201-250.
  31. Maeda, J. *The Laws of Simplicity (Simplicity: Design, Technology, Business, Life)*. The MIT Press, 2006.
  32. Merrill, D. and Maes, P., Augmenting Looking, Pointing and Reaching Gestures to Enhance the Searching and Browsing of Physical Objects. in *Proc. Pervasive'07*, (2007).
  33. Milgram, P. and Kishino, F., Augmented reality: A class of displays on the reality-virtuality continuum. in *SPIE*, (1994), 282-292.
  34. Mohler, B.J., Thompson, W.B., Creem-Regehr, S.H., Willemsen, P., Herbert L. Pick, Jr. and Rieser, J.J. Calibration of locomotion resulting from visual motion in a treadmill-based virtual environment. *ACM Trans. Appl. Percept.*, 4 (1). 4.
  35. Nielsen, J. Noncommand User Interfaces *Comm. ACM*, 1993, 83-99.
  36. Poupyrev, I., Newton-Dunn, H. and Bau, O. D20: interaction with multifaceted display devices *CHI '06 extended abstracts*, ACM Press, 2006.
  37. Rohrer, T. *Metaphors We Compute By: Bringing Magic into Interface Design*, Center for the Cognitive Science of Metaphor, Philosophy Department, University of Oregon, 1995.
  38. Second Life. <http://secondlife.com/>.
  39. Shneiderman, B. Direct Manipulation. A Step Beyond Programming Languages. *IEEE Transactions on Computers*, 16 (8). 57.
  40. Shneiderman, B. *Leonardo's Laptop: Human Needs and the New Computing Technologies*, MIT Press, Cambridge, Mass., 2002.
  41. Shneiderman, B. Why Not Make Interfaces Better Than 3D Reality? *IEEE Computer Graphics and Applications*, 2003, 12-15.
  42. Slater, M., Usoh, M. and Steed, A. Taking steps: the influence of a walking technique on presence in virtual reality. *ACM Trans. Comput.-Hum. Interact.*, 2 (3). 201-219.
  43. Smith, J.D., Vertegaal, R. and Sohn, C., ViewPointer: lightweight calibration-free eye tracking for ubiquitous handsfree deixis. in *Proc. UIST 2005*, (2005).
  44. Streitz, N.A., Tandler, P., Muller-Tomfelde, C. and Konomi, S. Roomware: Toward the Next Generation of Human-Computer Interaction Based on an Integrated Design of Real and Virtual Worlds. Carroll, J.M. ed. *Human-Computer Interaction in the New Millenium*, Addison-Wesley/ACM Press, Reading, Mass., 2001.
  45. Tanriverdi, V. A Virtual Reality Interface Design (VRID) Model and Methodology, Tufts University, 2001.
  46. Tanriverdi, V. and Jacob, R.J.K., Interacting with Eye Movements in Virtual Environments. in *Proc. CHI 2000*, (2000), Addison-Wesley/ACM Press, 265-272.
  47. Ullmer, B. and Ishii, H. Emerging Frameworks for Tangible User Interfaces. Carroll, J.M. ed. *Human-Computer Interaction in the New Millenium*, Addison-Wesley/ACM Press, Reading, Mass., 2001.
  48. Ullmer, B., Ishii, H. and Jacob, R.J.K. Token+Constraint Systems for Tangible Interaction with Digital Information *ACM TOCHI*, 2005, 81-118.
  49. Underkoffler, J. and Ishii, H. Urp: A Luminous-Tangible Workbench for Urban Planning and Design *Proc. CHI'99*, Addison-Wesley/ACM Press, 1999, 386-393.
  50. Usoh, M., Arthur, K., Whitton, M.C., Bastos, R., Steed, A., Slater, M. and Brooks, F.P. Walking > walking-in-place > flying, in virtual environments *Proc. SIGGRAPH'99*, Addison-Wesley/ACM Press, 1999, 359-364.
  51. Van Dam, A. Post-WIMP user interfaces. *Commun. ACM*, 40 (2). 63-67.
  52. Vinson, N.G. Design guidelines for landmarks to support navigation in virtual environments *Proc CHI*, ACM Press, 1999, 278-285.
  53. Vogel, D. and Balakrishnan, R., Distant freehand pointing and clicking on very large, high resolution displays. in *Proc. UIST 2005*, (2005).
  54. Warren, J. A Wii Workout: When Videogames Hurt *Wall Street Journal Online*, 2006.
  55. Weiser, M. The Computer for the Twenty-first Century *Scientific American*, 1991, 94-104.
  56. Zigelbaum, J., Horn, M., Shaer, O. and Jacob, R., The tangible video editor: collaborative video editing with active tokens. in *Proc. TEI '07*, (2007), ACM, 43-46.