Graduate Theses and Dissertations   Graduate College

2009

Conceptualizing design affordances from a cognitive perspective

Jeremiah D. Still
Iowa State University

Conceptualizing design affordances from a cognitive perspective

by

Jeremiah Daniel Still

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Human-Computer Interaction

Program of Study Committee:
Veronica J. Dark, Major Professor
Ana Correia
Stephen Gilbert
Alison L. Morris
William S. Robinson

Iowa State University

Ames, Iowa

2009

TABLE OF CONTENTS

ABSTRACT

The term *affordance* carries multiple meanings for designers. Traditionally, affordances were discussed within a Gibsonian framework as arising from direct perception of physical constraints. However, some authors extended the meaning to include learned cultural constraints, leading to a theoretical debate about whether designers should draw a distinction between perceptual affordances and learned cultural conventions. In this dissertation it is suggested that a broader meaning of *affordance* is more appropriate for designers and that a unified account of affordances can be achieved using a cognitive conceptualization of *perceived* affordances. Within this cognitive framework, perceived affordances arise from automatic processes in the user that either are inborn or have developed over time as consistent interactions produce changes in long-term memory. Well learned conventions are examples of the latter. The cognitive mechanisms responsible for how perceived affordances arise and how they affect the cognitive system are presented along with a flowchart to help guide designer decisions.

Three experiments examining empirical differences between affordances and conventions are reported. The first experiment asked whether users have developed conventions in the absence of affordances. A simple task was used in which participants pressed buttons in response to directional cues. The results showed that affordances exist when the spatial configuration of the buttons is congruent with directional cues. In the absence of affordances, most participants demonstrated consistent button-to-action mapping that represented a convention. Behavioral differences between affordances and conventions were not found. The second experiment confirmed that in ambiguous tasks, conventions guide expectancies about button-to-action mappings. The cognitive attributes of affordances

and conventions were examined in the third experiment by manipulating working memory load and expected interaction congruency. Results indicated some behavioral differences between acting on affordances and acting on conventions. However, violating the button-to-action mappings defining either an affordance or a convention produced similar performance costs. Taken together, the results suggest that after the initial learning period, conventions play a critical role in the perception of a design's available actions, just as do perceptual affordances. Therefore, designers ought to employ perceptual affordances when possible and when that is not feasible reuse established conventions.

CHAPTER 1. GENERAL INTRODUCTION

Human interactions with computing technologies are going to rapidly change over the next decade. According to Kurzweil (2005), this change will be driven by an exponential growth of technology, which will quickly result in smaller and less expensive computing devices. These technologies will become increasingly embedded within everyday environments (e.g., multi-touch tables, mobile phones and digital audio players). This technological shift to a ubiquitous computing environment will require multidisciplinary investigations (involving psychology, engineering, design, computer science, business, etc) into the complex interactions between humans and devices (Weiser, 1993). The results of such multidisciplinary human-centered design investigations will provide deeper insight into how one transforms interactive devices into systems that naturally interface with users.

My research question is general, but important as society demands technology use within everyone's daily activities. Why do some products make our lives easier, while others make our lives harder? It is often the case that easy to use products clearly afford certain actions. The embedded affordances provide implicit guidance for users. Lengthy technical manuals or repeated demonstrations from an expert operator are not necessary. Users are able to expend minimal cognitive effort towards figuring out the available interactions needed to complete a current task goal.

My dissertation research focuses on the cognitive basis of affordances, a major concept in design. The term *affordance* as originally introduced by J. J. Gibson (1979) refers to the functions an object offers within an environment. In natural environments affordances arise from physical constraints during actor and object interactions (e.g., a metal plate on a

door affords pushing). Thus, one important question in the design literature is: How does a designer create an interface that affords certain actions? One method is to replicate naturally occurring characteristics within the interface's design. For example, 3-D buttons that have shadows as visual cues afford pushing because they have visual characteristics indicating that they can be pushed. But, this direct mapping between natural and virtual environments does not always best support the user's needs and may, in many cases, be impossible. According to Rogers (2004), there is currently no model of affordances that designers can use to predict what functions an interface will afford. And, St. Amant (1999) pointed out that there is only an informal relationship between the understanding of an affordance and its influence on design decisions. The result is that designers are forced to depend on their own intuitions about what actions an interface affords.

Currently there is a debate about whether the term "affordances" should be reserved for physical interactions only (e.g., pressing a physical mouse button) or whether it can be used to describe interactions within virtual environments (e.g., clicking and dragging a scroll bar). Norman (1999) suggested that designers should refer to "affordances" in interface interactions (i.e., interactions with software) as *conventions* because the action mappings in the interaction are "arbitrarily" decided by the designer. For example, in theory, the designer who first placed the scroll bar on a window arbitrarily made the decision about whether moving the bar up would move the corresponding page up *or* down. By drawing a distinction between affordances and conventions, Norman implies that they are fundamentally different. But, one can ask whether the user's expectation, due to past learning, of a convention based action is different from an affordance based action. According to McGrenere and Ho (2000), conventions in the software are learned by users, and therefore, are affording actions. Given

this debate, one can ask: Are learned convention interactions behaviorally different than physical affordance interactions? Further, it is unclear what the cognitive attributes are for an affordance or a convention based action. Are affordances and conventions behaviorally similar? Or, do they rely differently on cognitive resources? Answers to these questions should provide some diagnostic insight into whether designers ought to treat an affordance and a learned convention differently.

Dissertation Organization

This first chapter frames my general research interests within human-computer interaction and provides the motivation for this research. At the heart of this dissertation are three articles that examine affordance and convention based interactions. The second chapter, Conceptualizing Affordances from a Cognitive Viewpoint, theoretically explores the concept of affordances from a cognitive psychology viewpoint with the goal of clarifying for designers the meaning and cognitive attributes of affordances. The third chapter, An Empirical Investigation of Affordances and Conventions (Still & Dark, 2008), empirically examines the concept of affordances and conventions within a simple button pressing task to determine whether behavioral differences exists. The fourth chapter, Examining Working Memory Load and Congruency Effects on Affordances and Conventions, further empirically explores and compares the cognitive attributes of affordances and conventions. The fifth chapter provides general conclusions regarding these embedded three articles and describes my future research goals. I was the primary researcher and author of these articles. My major professor, Veronica Dark, is listed as a co-author for her valuable contributions to these articles.

References

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton
   Mifflin.

Kurzweil, R. (2005). *The singularity is near: When humans transcend biology*. New York:
   Viking Penguin Group.

McGrenere, J., & Ho, W. (2000). Affordances: Clarifying and evolving a concept.
   Proceedings of Graphics Interface, 1-8.

Norman, D. A. (1999, May). Affordance, conventions, and design. *Interactions*, 38-42.

Rogers, Y. (2004) New theoretical approaches for Human-Computer Interaction. *Annual
   Review of Information, Science and Technology, 38*, 87-143.

St. Amant, R. (1999). User interface affordance in a planning representation. *Human-
   Computer Interaction, 14*, 317-354.

Still, J. D. & Dark, V. J. (2008). An empirical investigation of affordances and conventions.
   In J. S. Gero and A. K. Goel (Eds.), *Design computing and cognition '08*, (pp.457-
   472). Netherlands: Springer.

Weiser, M. (1993). Some computer science issues in ubiquitous computing. *Communications
   of the ACM, 36*(7), 75-84.

CHAPTER 2. CONCEPTUALIZING AFFORDANCES FROM A COGNITIVE VIEWPOINT

This article was prepared for the *Journal of Human-Computer Interaction.*

Jeremiah D. Still & Veronica J. Dark

Abstract

The term *affordance* carries different meanings within design communities. Traditionally, affordances were discussed within a Gibsonian framework in which affordances arise from direct perception. Along these lines, some authors now describe affordances as being mostly perceptual while others describe them as being culturally bound. We suggest that both of these descriptions are correct and that they can be explained from a cognitive conceptualization of *perceived* affordances. From this cognitive perspective, we suggest that perceived affordances arise from automatic processes in the user that have developed over time through consistent interactions with the environment. Design consistency is critical for producing effortless usage, because interaction consistency facilitates the formation of long-term memory structures. These structures affect perception of an interface without any intent on the part of the user. We explore the underlying mechanisms that could explain how affordances arise and interface within the user's cognitive system.

Introduction

Overview of the Article

According to Card, Moran and Newell (1983), understanding the operation of the human information processing system is a critical step towards creating usable design. In this

article, we discuss automatization, a characteristic of the cognitive system that *supports* effortless experience of perceived affordances because it allows rapid activation of knowledge without intention. Automatic processing does not require the limited capacity working memory resources that define a bottleneck in awareness or controlled processing; the primary requirement for developing an automatic process is consistency of a stimulus-response mapping. We propose that when a design produces constrained interactions, the cognitive pattern recognition system learns to automatically identify the constraints. We are not interested in any specific type of constraint (e.g., physical, social, logical, cultural, etc.); rather, we are interested in the cognitive effect of a constraint's being in play. In this article, we describe a cognitive conceptualization of affordance, including the resource limitations of working memory, the role long-term memory plays in perception, and the difference between automatic and controlled processing; we show how this perspective provides valuable insight into affordances; and we conclude by demonstrating how the cognitive attributes of a perceived affordance relate to design outcomes. Part of our description is in the form of a flowchart depicting the automatization process and development of perceived affordances that designers may use when applying these concepts.

Why Consider Affordances?

The term *affordance* was first introduced to the Human-Computer Interaction (HCI) community through Don Norman's (1988) book *The Psychology of Everyday Things* (POET). He suggested that perceived affordances arise from physical, logical, and cultural constraints. The concept of affordances was widely adopted in the design literature, but Norman (1999) stated that the meaning of affordance had been overly and incorrectly used. As noted by McGrenere and Ho (2000), the term had taken on a number of different

meanings far from the original meaning proposed by J. J. Gibson (1979). The overuse of the term was a major concern for Norman (1999) who stated that, "Sloppy thinking about the concepts and tactics often leads to sloppiness in design. And sloppiness in design translates into confusion for users" (p. 41). Despite this "sloppiness", the recognition of affordances within the HCI community has led to better designs. For example, designers have used perceived affordances to provide intuitive visual instructions for the users and thereby avoided the need to explain how to use simple everyday things. According to Norman (1988), "[perceived] affordances result from the mental interpretations of things, based on our past knowledge and experience applied to our perception of the things about us" (p. 219). However, several important questions remain to be investigated; for example, how is an affordance integrated into and how does it operate within the user's cognitive system?

The Original Meanings of Affordance

Several meanings of the term affordance appear in the literature. J. J. Gibson originally conceived of the idea of an affordance and his conception of an affordance is still the most often referenced meaning. According to Gibson (1979), "The *affordances* of the environment are what it *offers* the animal, what it *provides* or *furnishes*, either for good or ill. The verb to afford is found in the dictionary, but the noun *affordance* is not. I have made it up. I mean by it something that refers both [to] the environment and the animal in a way that no existing term does. It implies the complementarity of the animal and the environment" (p. 127). Gibson's affordance represents a critical concept for the design community as it embodies the interaction between the user and the device based on the user's perception of the properties of the device. In fact, the design literature has emphasized that an important attribute of an affordance is how well it fits with the user. For example, Gaver (1991)

provided an excellent example of how the emergence of an affordance critically depends on the characteristics of the user when he said, "a cat-door affords passage to a cat but not to me, while a doorway may afford passage to me but not somebody taller" (p. 80). Affordances only emerge from the interaction between the user and the device. Some affordances may be good in that they help the user to correctly interact with the device, but some may be bad in that they interfere. Thus, the successful designer will consider what the device's interface affords the target user for good or bad.

Norman (1988) originally popularized the concept of an affordance within the HCI field when used it in a design context and expanded the type of interactions to which the term applied. While Gibson's affordances came from the direct perception of physical constraints, Norman suggested that *perceived* affordances arise from physical, logical, and cultural constraints.[1] Norman's view highlighted the fact that both physical and uniquely mental constraints directly affect the user's perception of available actions being offered within an environment. An example of a physical constraint might be a wall preventing a person from physically passing from one room to another. A logical constraint is one that strongly relies on a conceptual model of how a device operates. It is a logical constraint that stops people from attempting to open a door that they can see is locked. A cultural constraint is one that is formed through life experiences that are bound to local cultural norms. For example, to turn on a light in the US one flips the light switch up, but the opposite is true in the UK (Oshlyansky, Thimbleby & Cairns, 2004). Perceived affordances are by-products of

---

[1] Although the term "affordance", without a modifier, is sometimes reserved for the physical affordances described by Gibson (1979), Norman's (1988) term "perceived affordance" actually includes physical affordances. This produces the confusing situation in which the unmodified term is actually subordinate to the modified term.

constraints that are logical, physical, or cultural. According to Norman (1988), these constraints allow simple objects to be usable without extensive instructions. We are taking the next step in the process of understanding affordances in design by linking Gibson's concept of direct perception to the cognitive concept of automatic processing and examining how affordances are produced and affect the cognitive system.

Different Theoretical Views of Affordance

The concept of affordance has taken on a number of different meanings depending on the theoretical framework for which it is being adapted. The distributed cognition theory focuses on a combination of possible affordances arising from a number of different constraints (biological, physical, perceptual, cognitive, social) that emerge during collaboration of multiple users interacting with multiple devices (Zhang & Patel, 2006). The structuration theory points out the effect of the larger socio-cultural constraints on affordances (Vyas, Chisalita & Veer, 2006) and suggests that human actions are supported and discouraged by the users' social structures (see Giddens, 1984). Hartson (2003) used Norman's (1986) stages of action model to identify four different types of affordances: cognitive, physical, perceptual, and functional. He described how each type of affordance can play a role in creating a successful user interaction. These theories emphasize a number of different meanings of affordance but none provide an explicit description of how affordances are created or operate cognitively. This is problematic for a designer who needs to predict the existence of an affordance within an interface and understand its effect on the current activities being performed by the users.

As previously noted, Norman (1999) and others have suggested that the new views of affordance have confused the issue. According to McGrenere and Ho (2000), "As the

concept of affordance is used currently, it has marginal value because it lacks specific meaning" (p. 8). Similarly, Torenvliet (2003) suggested that "just as printing new money devalues existing money, the more new definitions the term *affordance* gains, the less value any one of them has" (p. 13). While we acknowledge that there are many types of affordances, we propose that a cognitive conceptualization of *perceived affordance* incorporates the characteristics of the concept that are useful within the context of design.

Our Purpose and Motivation

The purpose of this article is to propose a framework that describes the creation of perceived affordances and describes the attributes of an affordance within a cognitive system. Briefly, a cognitive system operates on the representations arising from structures in long-term memory. Our purpose is not to add confusion by introducing a new "affordance", but to clarify what an affordance is. Perceived – in perceived affordance – comes from the word perception, which describes the user's process of transforming sensory input into task usable representations. Thus, a perceived affordance is a result of an object's function becoming apparent to the user through its cognitive representation. Given this definition, even Gibson's mostly externally based view of affordance is a perceived affordance. Further, we propose that the critical distinctions between different types of affordances arise from the environment and the *cognitive* characteristics of the user. We believe that the cognitive perspective will benefit designers because it highlights the role that learning plays in creating perceived affordances.

According to Herbert Simon (1969/1981), "the designer, is concerned with how things ought to be – how they ought to be in order to attain goals, and to function" (p. 7). We adopt Simon's perspective in that designers create with the hope of producing objects that

help users meet their task goals. A created object is 'artificial'; it is "molded, by goals or purposes, to the environment in which it lives" (Simon, 1969/1981, p. iv). With the development of technology, the meaning of 'artificial' has become even more extreme as objects may only exist within a virtual space. Such an unconstrained space provides designers with the ability to create unworldly objects that meet previously unrecognized needs. For this reason it is critical to understand the cognitive properties of these designs (Visser, 2006).

Designers may find it beneficial to take advantage of natural-world constraints within artificial environments. In an unbound (artificial) environment this could be accomplished by replicating naturally occurring characteristics within the interface design (e.g., 3-D buttons that have shadows as visual cues). But, the direct mapping between the natural and virtual environments may not always best support the user's task needs and may in many cases not even be possible (Rogers, 2004). When natural mappings are not employed within a virtual object it is not clear whether actions may be afforded. This leaves designers to depend on their familiarity with what an interface *should* afford and how it ought to affect the cognitive system, making design decisions difficult.

You and Chen (2007) highlighted a particular benefit of affordances; "the concept of affordance challenges designers to avoid the reliance on symbols and cultural conventions in design. Instead, it encourages them to utilize possible intuitive actions that can serve as function in the process of user-product interaction" (p. 29). However, as suggested by Raskin (1994), the "intuition" experienced from an affordance could be reframed as the user's familiarity with the object's functions. The implication, which is that well learned cultural conventions could become just as "intuitive" as Gibson's affordances, fits well with a

cognitive view. Cultural conventions are the result of many consistent interactions that produce coherent structures in long-term memory.

Norman (1999) described conventions as arbitrary mappings between an action and a function that become common within some group. Conventions are often created within interface design when no perceptual affordances are available. We suggest that if a convention is used frequently, it becomes a perceived affordance. Thus, we agree with McGrenere and Ho (2000) who stated that software applications may afford specific interaction actions. These afforded actions emerge from learned conventions within a virtual interface. "The functions that are invokable by the user are the affordances in software… Norman claims that a scrollbar is a learned convention and implies that it is not an affordance. We disagree. The fact that the object affords scrolling is an affordance that is built into the software. The information that specifies this affordance is in fact a learned convention… [Acting on the bar will either move the page up or down]" (p. 6).

Although the desktop user is physically unconstrained to click anywhere within the visible screen, the past experiences of the user will make clicking the close-window button highly likely behavior when the desire is to end an interaction (see Figure 1). It is this type of constrained interface interaction (i.e., one that depends on learning) that defines Norman's (1999) conventions. Naturally occurring physical constraints are affordances (e.g., pushing a mouse button), while more arbitrary constraints (e.g., associating page movement with scroll bar direction) are conventions. When a physical affordance is present, the designer and user are likely to view the situation in a similar way. When a physical affordance is not present, the designer needs to consider the possibility that a convention or a perceived affordance exists. When we investigated whether users act on an affordance differently compared to a

convention in a very simple button experiment (Still & Dark, 2008; see Chapter 3), we found

that they did not. The convention and affordance interaction types did not differ from each

other as measured through response agreement or response consistency. This result suggests

that highly learned, consistent artificial interactions may become as effortless as naturally

occurring affordances.

← Affordance (virtual)

← Affordance (physical)
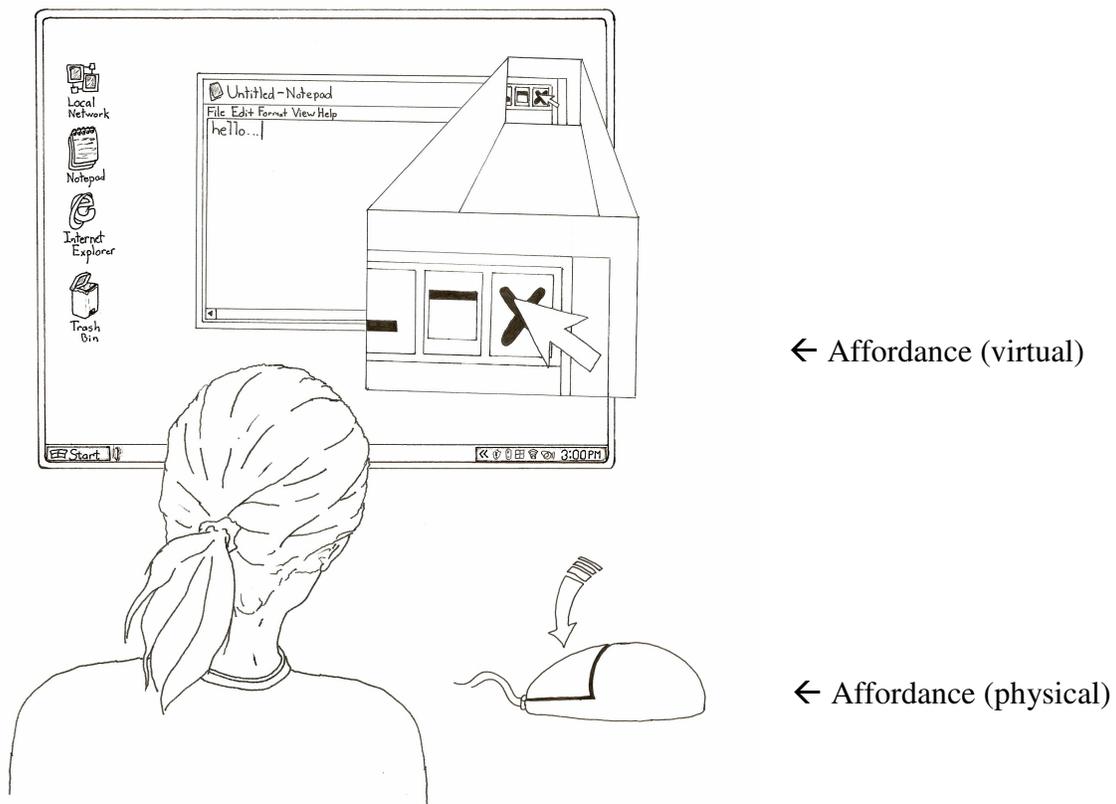
Artwork by Austin G. Still

Figure 1. Displays a user acting on both physical and virtual affordances to close a window. This figure shows how both the user's physical and cognitive attributes structure actions. Cognitive constraints guide a user's actions as much as physical constraints. When a user needs to perform a routine task like closing a window, the user does not use the mouse to randomly click around the virtual desktop but instead clicks only on the location likely to accomplish the task (i.e., the virtual button).

The proposed cognitive approach takes into account the user's underlying representation of a design whether it is constrained virtually or physically. Gibson's (1979) and Norman's (1988) illustrations of affordances focused on the physical constraints that arise within user and object interactions. In contrast, we propose that, for design purposes, it does not matter whether constraints arise through a physical or psychological means; for either case there is an affordance in play. Although there may be some historical difference in how the representation supporting a physical affordance and a virtual affordance become part of an cognitive system, once the representation is part of the system, its impact on behavior is similar. It produces effortless (direct) perception of a situation supporting a possible action.

A Cognitive System Perspective

As noted earlier, a cognitive system assumes that there are internal structures in long-term memory and that these structures can be activated by the information from the external world or by internal goals. Long-term memory changes constantly as a user experiences the world. The more frequently a situation is experienced, the stronger the long-term memory structure associated with that situation. Cognitive processing is constrained by the relatively limited amount of information that working memory is able to process at any time. Therefore, the system heavily relies on perceptual and interaction patterns that are stored in long-term memory to provide expectations about a situation. Both external environmental cues and internal higher order task goals contribute to the activation of long-term memory structures. If structures become activated, then the information processing represented by those structures is directly available, without consuming working memory resources. Thus, within the context of a cognitive system, a user's perception of an interface is a reflection of a representation that was molded by activated long-term memory structures through a

combination of environmental and cognitive contributions. In the sections to follow we expand upon components of the cognitive system, the process of automatization and the role of consistency in that process.

Working Memory

Working memory is associated with awareness. It is a temporary storage usually discussed as a mental workspace for controlled processing (e.g., Baddeley, 1986; Baddeley, 1992). For example, this work area could be used as calculation space for complex math problems or as an area in which visual information can be manipulated to produce an image in awareness. Current theories of working memory suggest that this work space may hold approximately four items[2] (Vogel, Woodman, & Luck, 2001). Thus working memory has a limited capacity, and, by extension, the amount of information available to awareness at any point in time is limited.

Working memory is also where information is assembled (encoded) for long-term memory storage. In addition to general world knowledge, long-term memory contains the type of information that most people associate with the term "memory"; that is, it contains knowledge and information about life events. Long-term memory is large, with no known storage limitations. Recently, Baddeley (2002) has incorporated long-term memory into his working memory model making it a more general model of memory. This adaptation is important to designers because it acknowledges the direct influence of past experience, through long-term memory, on *current* interface representations.

---

[2] This estimate of working memory capacity is for visual stimuli presented during articulatory suppression.

The amount of working memory resources currently being used must be considered in any task. The more working memory resources required by a task, the more effortful it is. Effortful processing is known to be slow and serial, however it is flexible. Technologies should avoid drawing on limited working memory resources as the user is often directing these resources to other tasks or goals. Technology should support the user's goals at the periphery of attention. In other words, the user should not be spending working memory resources trying to interact successfully with a familiar device. The relationship between acquiring knowledge and cognitive processing difficultly depends on the appropriateness of supportive structures in long-term memory. Typically interaction sequences and tasks must go through at least an initial heavy learning process involving limited capacity (i.e., working memory) resources before the process occurs without awareness. In addition to considering the amount of limited working memory resources required by a task, one can consider what attributes of the device could facilitate easy and consistent mapping of a task across a number of encounters.

The Impact of Long-term Memory on Working Memory

Current design consistency principles focus on minimizing the demands on working memory (e.g., Apple Computer, Inc., 1992; Nielsen, 2002; Shneiderman & Plaisant, 2005). We argue that in order to know how to minimize working memory demand, one must consider long-term memory's role in perception. The representations that are in working memory depend on knowledge and structures currently in long-term memory, i.e., what is already known. Although there are many specific models of memory, Cowan's (1988) model is particularly well suited to capturing the interaction between working memory and long-term memory. Cowan suggested that short term storage is actually highly activated

regions within long-term memory. This is a critical idea that has been missing from many designers' model of the user's cognitive system. The focus has been only on the limitations of working memory and not on the relationship between the contents of working memory and the contents of long-term memory. Long-term memory structures affect how representations are formed within working memory. Thus, by influencing the creation of representations, long-term memory directly affects the user's perception. Once this relationship is acknowledged, the designer has a new tool that may be used to circumvent the limitations of working memory (cf., Ericsson & Kintsch, 1995). Interactions with interfaces may be made easier for the user by providing information in such a way that is easily interpreted (represented) because the information relies heavily on previously consolidated representations available in long-term memory (i.e., more information can be included within a single representation in short-term memory).

Ericsson and Kintsch (1995) described how well-developed long-term memory structures can ease working memory load. Basically, experts are able to overcome the capacity limitations of working memory by developing domain-specific perceived affordances. Novice chop stick users must direct cognitive effort, or working memory resources, towards navigating the chop sticks to allow eating noodles. Their actions must be focused on how to hold the chop sticks in a way that picks up the noodles. However, an expert user of chop sticks can eat quickly while simultaneously carrying on a complex discussion. The knowledge of how to use chop sticks to pick up food has already been stored for the expert user in long-term memory. It is directly accessible without cognitive effort; it is a perceived affordance.

Designers implicitly understand the relationship between learning and long-term memory. For example, designers do not consider a user's interactions with a mouse when the user is learning a new software program to be part of the learning process, because the designers assume that the user already knows how to use the mouse. This assumption reflects the fact that designers understand the impact of the user's long-term memory on the current interaction. While most physical affordances are apparent to users regardless of their specific past experiences because the affordances reflect structures that are the result of evolved history, perceived affordances reflect specific users with specific histories. Designers need to consider specific histories. The question is whether there was enough consistency in the past interface interactions that a perceived affordance exists.

According to Hannon (2008), "…our ability to learn and use new technologies is contingent upon our experience with prior technologies. On a computer, for instance, each time we learn a new interaction idiom such as drag-and-drop, or double clicking, or scrolling, we adopt new ways of understanding how software applications and hardware devices work" (p. 60). Designs that follow already learned consistencies take advantage of the perceived affordances in technology and as a result, the device interactions are processed effortlessly. This is similar to how the pattern recognition system processes the text you are reading without any awareness on your part of how you are processing each letter and compiling words.

Describing Automatic and Controlled Processing

All interface interactions affect long-term memory structures when interaction instances are stored in long-term memory. If consistency is detected across these instances, the interaction may become automatized, but consistency does not guarantee automatization.

By definition, a process is automatized when it does not require working memory resources for the target action to become apparent. Within the psychology literature controlled processing is associated with the awareness of on-line progress towards a product's production while automatic processing is associated with only being aware of the product produced and not the production process itself. Posner and Snyder (1975) offered diagnostic criteria that discriminated between these two types of cognitive processing (for a more recent account, see Moors & Houwer, 2006). Automatic *processing* is described as not being open to awareness, rendered without intention, carrying a light working memory load and leading to rapid responses to stimuli. Controlled *processing* is described as being open to awareness, resulting from intention, carrying a heavy working memory load and leading to slower responses to stimuli. It is important to note that both of these processes result in a product that is available to awareness.

One distinctive indicator of an automatic process is whether or not one has control over the process. The Stroop (1935) effect is a classic example of two automatic processes competing for a verbal response, occurring without intention, that happens to compete with the intended response. In the Stroop task participants are shown words such as, GREEN, BROWN, YELLOW, RED that are presented either in a visually congruent colored ink (GREEN is printed in green ink) or a visually incongruent colored ink (GREEN is printed in red ink). The participants' task is to report the ink color of the word ignoring the meaning of the word. Thus, this task instructs participants to only process a single stimulus dimension - color. The Stroop effect is the finding that participants have difficulty naming the ink color when the meaning of the word is incongruent. This difficulty arises because the response produced by the unintended automatic process of reading the word competes with the

intended output response of producing the ink color. Another characteristic of an automatic

process is that it occurs without awareness, thus it is not available for introspection. Hence,

people cannot actually describe what mental process they rendered to know how they read

the word *red* or what made them say *red*, only the product is available.

In contrast, the major distinguishing characteristic of a controlled process is its

demands on working memory resources. When people are aware of their processes, they are

able to regulate and describe their operations. For example, when participating in a

conversation one can reflect on what was said and why it was said. But, a controlled

rendering of a task consumes most of the available working memory resources, making it

difficult to successfully render two cognitively demanding tasks at once. For example, having

a serious phone conversion while driving in heavy traffic necessitates the continual switching

of resources between the two tasks. In fact, it has been shown that if resources are engaged in

completing the conversation task, the ability to successfully monitor the driving task will

significantly decrease (see Recarte & Nunes, 2003; Strayer, Drews & Johnston, 2003; Strayer

& William, 2001).

While both driving and conversing were just described as demanding tasks that need

a large amount of working memory resources to be rendered successfully, oftentimes

engaging in a conversion and driving a vehicle represent a mixture of controlled processing

and automatic processing. For example, some actions associated with operating a vehicle are

consistently performed the same way, such as turning the vehicle off or on or moving the

vehicle left or right with the steering wheel; once learned, these actions can be performed

automatically. These types of automatic processes consume few working memory resources

because the long-term memory structures are so well constructed that the whole structure acts

as a single unit, or schema. According to Norman and Shallice (1986), "…when the source schema for a task such as driving an automobile has been selected, all its component schemas become activated, including schemas for such acts as steering, stopping, accelerating, slowing, overtaking, and turning. Each of these component schemas in turn acts as a source schema, activating its own component schemas (braking, changing gear, signaling, and so on)" (p. 6). However, tasks that are dynamic, such as engaging in a serious conversation or driving in a busy environment, require significant working memory resources. Processes that are controlled consume the most working memory resources; therefore performing multiple controlled tasks demands limited resources for each task leading to a bottleneck in working memory processing.

Consistency in the Stimulus/Response Relationship is Critical

Shiffrin and Schneider (1977) explained how tasks are a mixture of both controlled processes and automatic processes. They demonstrated that after consistent, repetitive practice in a visual search task, participants were able to search through multiple distractors without the number of distractors modulating search times. Basically, they showed that with enough practice in a consistent mapping situation, a once mostly controlled task will become mostly automatic. However, if the situation is variable, or not predictable, the task will continue to be processed in a controlled fashion that requires working memory resources regardless of the amount of practice. For example, how could typing on a keyboard become automatic if the key positions periodically changed? The practiced task must be consistent to have the opportunity to become an automatic process.

Perceived affordances are produced by a consistent design, which generates consistent interactions between actions and functions. Further, for users to recognize

interaction consistency they must have a good mental model of how the device operates (see Norman, 1988). According to Norman, a good mental model facilitates rapid formation of structures within long-term memory, because the user is able to establish a clear cause and effect relationship between the user's actions and the device's responses. Consistent design is a reflection of the many constraints applied within the literature. Logically, something that is highly consistent is highly constrained. Thus, this provides some evidence for why the consistent design principle is important to consider within a cognitive framework.

Indeed, the design literature already highlights the importance of design consistency within their fundamental principles with the goal of improving a design's usability (e.g., Apple Computer, Inc., 1992; Nielsen, 2002; Shneiderman & Plaisant, 2005). Consistency in design has had a major impact on search behavior within the context of the internet. One may infer that users make fewer navigation errors if a website's formatting is consistent versus varied. An example of consistency in design is found within the web site programming community referred to as Cascading Style Sheets (CSS). CSS separates the formatting of style and content by having a single file determine the formatting for all the pages within a website (Bos, Celik, Hickson & Lie, 2006). CSS was developed so that web designers could modify content without having to recursively include style code. However, as revealed by considering the impact on the user of the consistency, the use of CSS actually minimizes the user's need to employ cognitive effort while searching through a website. For example, an underlined word would always indicate a hyperlink on all pages within a website. This consistent design leads to rapid learning and appropriate expectations of what items have hyperlinks.

Perceived Affordances

Building Perceived Affordances

Our description of perceived affordances focuses on the psychological interpretation of what an interface communicates and is not tied down to any specific type of constraint: logical, cultural, social or physical. Rather, perceived affordances are the result of continuous learning. The learning reflects changes in the user's cognitive system that capture the predictable (regular) characteristics present within a complex operating environment.

The cognitive system view provides evidence for how perceived affordances are acquired. The concept of perceived affordances proposed here is not directly associated with any recent descriptive account of affordance; rather, it is integrative. From a cognitive perspective, a perceived affordance is a reflection of previous knowledge rapidly and without awareness, influencing the pattern recognition system.

The designer must decide what aspects of the interaction can and should be automatized and then capitalize on prior knowledge that the user brings to the situation. Designers should not assume *tabula rasa*, but instead should consider the current foundation upon which the new task will be built, especially if perceived affordances may be in place. That is, the designer must recognize that previous knowledge may not be task neutral, it may interfere with or facilitate the user's interaction with the new task or technology.

The flowchart in Figure 2 depicts the process by which a perceived affordance can be created in a user. The flowchart represents the interplay between the user's cognitive outcomes associated with an interface's interaction, the actual design of an interface, and the design elements imparted by the designer. Further, it highlights a user's cognitive system and the flow of design outcomes; that is, it predicts the type of cognitive processing that occurs

after a user has some experience interacting with the device. The circles indicate the start or stop of the flowchart analysis. A diamond reflects a task evaluation. The response to those evaluations (yes or no) directs the flow through the chart in the direction of the arrows. The large rectangles represent the clarification of task processing (controlled or automatic process) and the cylinder represents a variable number of practice sessions required for sufficient experience, which depends on the complexity, familiarity and working memory load associated with the task.

This flowchart clarifies how perceived affordances are achieved from the designer perspective. When using the flowchart begin at the "start" circle and progress in the direction of the arrows. The first question, "Perceived Affordance?" refers to whether or not the task is already an automatic process within long-term memory. The second question, "Motivated?" is a motivational assessment. If the user will not put forth the effort to complete or select the task, the analysis must stop and the situation must be reexamined. The third question, "Designed to be Consistent?" asks if the present and past interactions between the user and device are consistent. If the interaction is identified as being a controlled process it will require working memory resources. Nevertheless, if the technology interaction is consistent, the novel process may become automatic for the user. This may take a lot of trials depending on the complexity of the pattern and learning conditions. If the user is not able to complete the consistent task effortlessly after numerous practice trials, the user may not be detecting the constraints. In other words, the user may not be recognizing the design's consistency. However, if the user recognizes the consistency, then more practice trials are required. The last question, "Effortless?" asks if enough practice sessions were completed for the action to become a perceived affordance (i.e., available through direct perception). If the user is not

effortlessly interacting with the design, the practice must halt and the interface's design

consistency needs to be reconsidered.
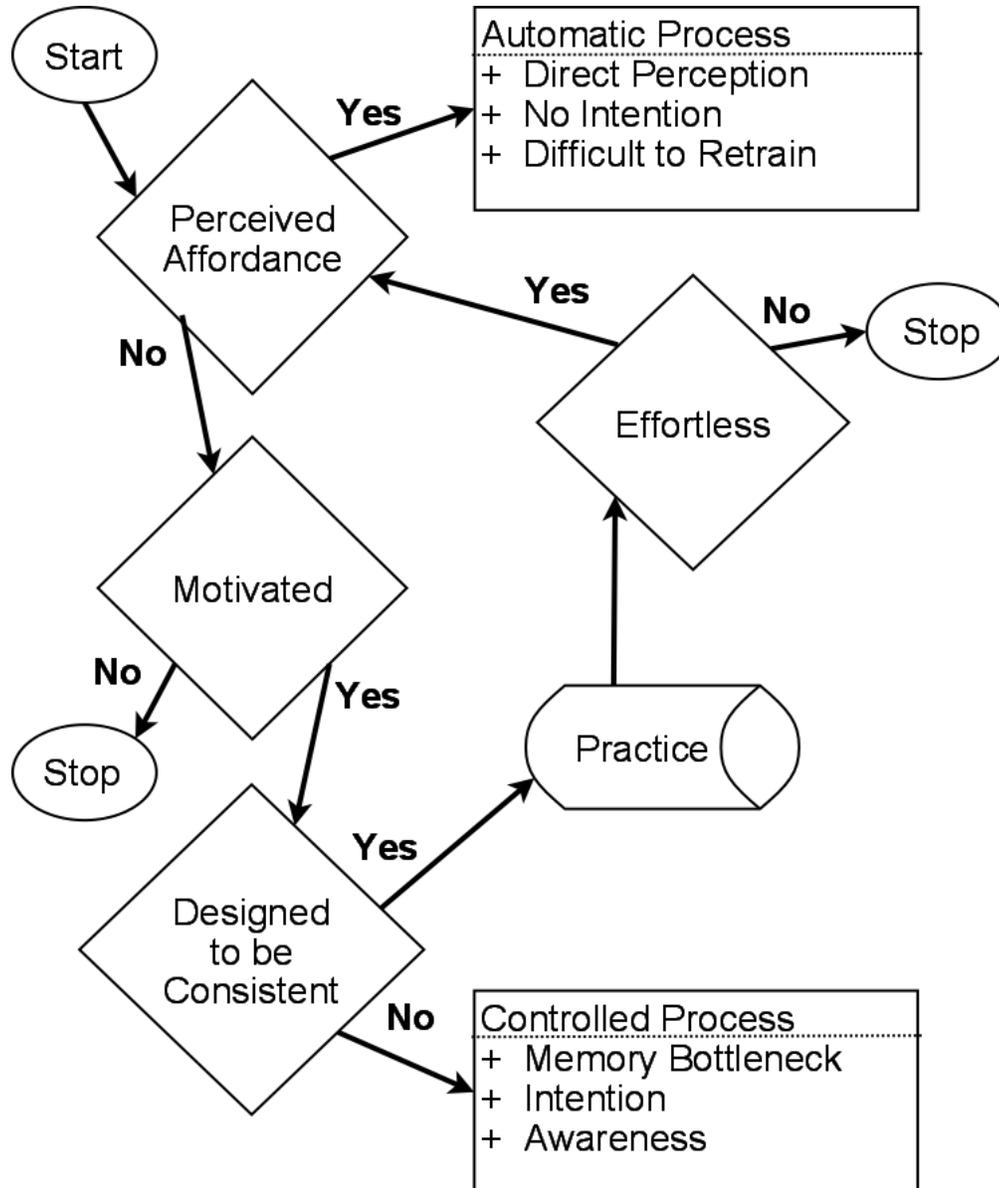
Figure 2. Evaluating user and design interactions within the context of automatic and controlled processing

From a cognitive system perspective the lack of transparency, natural mapping, or

appropriate feedback may cause increased time spent within the practice sequence (see

Figure 2). If the processing flowchart indicates that the device interaction task ought to be an automatic process, yet the task seems to be consuming a lot of resources after many practice sequences, designers should reconsider their interface and what it is communicating to the user (see Hutchins, Hollan & Norman, 1986; Norman, 1988, for more information). In other words, the designer may be predicting that the system interface is displaying consistency, but the user is not able to perceive the consistency. The user may not be able to determine the cause and effect relationship between their input and the devices output. This may be due to the user having difficulty finding a correct mental model of how the device operates, thus inhibiting the formation of a perceived affordance. Or, it may be that prior interaction structures are interfering with the desired outcome.

Negative Transfer from Prior Structures

The user's interaction experiences are always molded by the activated structures within long-term memory even if they are inappropriate. Consequently, if an interface's controls appear familiar but now have a different meaning, the designer may have created a phantom affordance. It is not the case that new information simply supplants old information. Rather, interference occurs when the old information competes with the learning of the new information (Keppel & Underwood, 1962). Psychologists use the term negative transfer to describe incorrect mapping of previous response knowledge onto different response situations (e.g., Besnard & Cacitti, 2005). For example, experienced users already have structures in long-term memory for what specific buttons do on an interface. So, although it might be cheaper and easier from an installation perspective to just keep a previous button in place and change its function, explicitly informing the user of the change or posting a sign stating that "now" the button does a different thing, this is likely to produce negative transfer.

Negative transfer could result in a potentially dangerous situation because an expert operator may without awareness respond to the traditional meaning of the button. Posting a flyer spatially near the interface that describes the change or verbally instructing the user of a function change is not enough. The designer needs to create a visually different button because it has a new system function. Providing different visual cues for new functions should decrease the likelihood of negative transfer or phantom affordances in which a known cue is associated with an unexpected response.

The effects of perceived affordances on task outcomes can be either positive or negative depending on the congruency of the current and previous interfaces. For example, a standard in web design is to identify hyperlink words by underlining them. For experienced users, underlined words conceptually afford clicking. But, what if the characteristic feature of underlining did not really represent a hyperlink on a web page? In this case, the designer has created a phantom affordance (e.g., underlining words now indicates a heading rather than the conventional hyperlink). Phantom affordances appear without the intention of the users, causing them frustration. These disconnects between the interface design and the user's previous knowledge may lead to unsuccessful interactions.

Chapter 4 empirically explores the effect of violating expectations in a button pressing task. We found that reversing the mapping between action and function produced an incongruency cost resulting in decreased task performance. Interestingly, the cost of reversing the action to function mapping was similar for physical affordances and conventions.

Selecting a Perceived Affordance

We have suggested that perceived affordances arise from a user's current interaction representation. There are often multiple types of these interactions available for selection within a single object. According to Brown and Blessing (2005), there are a very large set of possible object functions and these affordances are reduced through context-dependency. From our viewpoint this large pool of possible affordances arises from a large number of possible representations stemming from both physical and *mental* structures. However, although many affordances exist, it likely that a single affordance becomes apparent because of its relatedness to ongoing task needs, familiarity of interaction and environmental contingencies.

It is common for authors to propose several perceptual affordances for the same object depending on physical interaction conditions (i.e., stemming from the environment). For example, Michales (2003) pointed out how tool usage changes the user's possible selection of affordances. In addition, she writes that, "The next step is the spooky one: The implement that permits the action must be built, or designed and built: Object X affords lifting by a crane that can lift twice as much as any existing crane. The frozen pond affords skating when someone invents skates"(Michaels, 2003, p. 141). Ecological researchers often point out that once a user is using a tool, new affordances appear. It is clear from Michaels' quote that the way *affordance* has been used leads to the production of seemingly an endless number of affordances.

Why is it feasible to discuss tools of the world providing new affordances, but not cognitive tools (Norman, 1993) such as representations? Just as new perceptual affordances can emerge from interacting with a different physical tool, new perceived affordances arise

from using different mental tools, or representations. For example, humans often chose to represent information in different forms depending on their current task needs. When counting the number of people attending a conference presentation, it is easier to visualize and denote the total amount in terms of tally marks. However, if the task were to determine the total number of chairs in a room, it is easier to think in Arabic notation when multiplying the rows by columns. A variety of cognitive representations allow humans to successfully complete a diverse set of tasks. Dimensional reduction of a complex environment into a usable representation is a fundamental part of the cognitive system given our working memory limitations. For example, objects in a room activate an actor's source schema producing an internal context of their operating environment. This internal pattern of activation makes task-related affordances apparent with little cognitive effort.

According to You and Chen (2007), there may be a "…seemly infinite [number of] affordances, which will catch the user's attention most? Why is an action executed but not the others? Questions like these might be important to designers, but they are not within the scope of the theory of affordance" (p. 28). As just noted, however, from a perceived affordance perspective we are able to begin answering these questions. Given that a device activates multiple affordance representations depending on the user's current physical or knowledge characteristics, how is a specific representation selected over the others? According to the biased competition model of Desimone and Duncan (1995), the representation with the strongest activation wins. The strength of the representation is determined by a combination of the artifact's physical attributes and the user's internal representation - which is heavily influenced by higher order task goals - that most closely fits the system's current state.

Even having defined a method for selecting an affordance, the practical question of whether a physical affordance or a perceived affordance carries more activation remains. It might seem obvious to think that affordances stemming from external information carry more weight than affordances stemming from internal information, and that they are therefore more likely to be used, but recent research has shown that external information is not always preferred over information stored in memory. For example, users may actually act based on remembered information over graphically presented information to complete some tasks. According to Gray and Fu (2004), millisecond differences in time to execute a task matters to whether a user selects to retrieve knowledge-in-head or in-the-world. If it saves the user a marginal amount of time to potentially use a faulty memory over retrieving reliable information from an external source, users often will chose the internal memory (Gray & Fu, 2004). These results suggest that even when a physical affordance is presented in a perceptually salient and natural manner, the user may still choose to use a perceived affordance from long-term memory. More research will be needed to determine the effect these perceived affordances have within the context of usability.

Relationship to Gibson (1979)

Gibson (1979) stated that the two primary features of a perceptual affordance are that it is the result of direct perception and that it does not require excessive amounts of learning. Our view of perceived affordances is that direct perception can be conceptualized in terms of automatic processing, which is the result of activation of long-term memory structures. To say that a perceived affordance is the result of direct perception means that it is automatically processed; its impact on the pattern recognition system is rapid, occurs without awareness, and does not require any working memory resources.

As we have noted, every time users interact with an interface, they learn. That is, each instance of learning produces a small change within the organizational structure of a user's long-term memory (Cowan, 1988; Logan, 2002; Shiffrin & Schneider, 1977). The amount of learning required depends on whether or not an existing structure is in place to support the incoming information. For example, if there are not any existing structures in place, one can refer to the user as a novice, but if well-developed structures exist, the user is referred to as an expert. An expert user already has a number of perceived affordances in play within a complex interaction environment (cf., Ericsson & Kintsch, 1995). These perceived affordances build on each other, which allows the expert user to minimize learning requirements within complex interaction environments. Further, experts are able to render complex tasks without many errors or much effort. However, a novice user would have a very difficult time completing the same task due to the lack of long-term memory structures that produce necessary perceived affordances. Thus, minimal learning is continuously measured in reference to the current state of knowledge as learning is always building on previous experiences.

Conclusions

It has been suggested by a number of authors that affordances should include the observer's previous knowledge as indicated by Gibson's (1979) mailbox example in which he pointed out that direct perception allows mailboxes to afford inserting letters. However, others (e.g., Greeno, 1994) have pointed out that Gibson's mailbox example is atypical for his description of direct perception as it requires the activation of symbolic representations stored in memory containing the properties that provide identification of a mailbox. We suggest a human has direct perception of affordances in the external world because that is the

way the cognitive system evolved, but a human also has direct perception of aspects of the world as a function of learning (cf., Marcus, 2004). In other words, it does not matter whether affordances reflects perceptual or learned constraints; in both cases the interaction is experienced automatically. Direct perception provides effortless insight into what actions an object affords. Whether the direct perception arises from physical properties of the world and built-in to humans through evolution or is a result of automatic processing achieved through learning may be of interest to perception theorists, but it should be irrelevant from the designer's perspective. Designers should consider both kinds of direct perception, or the usability of their designs will suffer.

Previous knowledge in long-term memory plays a critical role in cognitive pattern recognition. Affordances emerge from constraints formed through interactions between an external object and the user's personal characteristics (Gaver, 1991). In other words, affordances are the result of an interaction between the external object provided physical constraints and the user's perception. The representation being used during task completion will drive the meaning of the specific affordance in play. Advanced knowledge, or expertise, is required to identify perceived affordances that exist within less physically constrained environments. The more degrees of freedom an interaction requires, the less likely a novice will detect the underlying structure or consistency. Highly constrained interactions allow novices to easily detect the consistency. However, with experience, users may learn to recognize complex patterns within less constrained interactions; they may develop perceived affordances.

The literature has yet to describe the cognitive process by which affordances are created. This article provides a framework that explicitly describes the process from which

perceived affordances arise and the cognitive attributes associated with such an outcome. Within the framework described, perceived affordances result from automatic processing of current information because structures within long-term memory guide the processing. Perceived affordances can reflect many different types of constraint. Automatic processing is defined as occurring without awareness, intention (through direct perception) and without any effort. These attributes are also used to describe affordances. Though a perceived affordance is not exactly what Gibson (1979) meant by affordance, it does have more application for designers. A number of authors have mapped affordances onto a number of different frameworks trying to achieve more applicability (see Zhang & Patel, 2006; Baerentsen & Trettvik, 2002; Vyas, Chisalita & Veer, 2006); we believe that our view unifies these different frameworks into one appropriate for design applications. The cognitive view of perceived affordance clarifies the aspects of an interaction that are important for design. It does not focus on a particular type of constraint, but rather on the possibility that a constraint may be in play. A constraint may be the result of consistent interactions, which lead to the formation of long-term memory structures that directly affect working memory representations without any intention or effort on the part of the user.

Future work needs to consider common affordances that arise within a group of experts (i.e., those who have had many consistent experiences within an interface). Applying the perceived affordances framework, for instance, the designer may become more sensitive to the fact that users who have been using a window, menu and pointer schema for years would directly perceive this platform much differently than would a novice. Consequently, developing future interfaces that are optimal for the novice user might not be optimal for the expert user. Designers working within the cognitive system framework would be more likely

to recognize that previously adopted platforms should have heavy influence on future platforms' usability. According to Hannon (2008), designers often modify their products' 'contingencies' without any clear progression towards a more usable product. Instead of focusing on interaction design challenges, such as creating interaction conventions as a community, each company often creates its unique solution. This independence may lead users to become frustrated with technology as they are required to interact with more and more devices to complete their daily activities.

It is our hope that the cognitive system framework will help designers create interactions that take advantage of perceived affordances when appropriate. This approach may encourage designers to produce devices that create minimal cognitive cost. Decreasing the user's cognitive load will create smarter users (Sternberg, 1986). And, designers always want smarter users!

Acknowledgments

References

Apple Computer, Inc. (1992). *Macintosh human interface guidelines*. Reading, MA: Addison–Wesley.

Baddeley, A. D. (2002). Is working memory still working. *European Psychologist*, *7(2)*, 85-97.

Baddeley, A. D. (1992). Working memory. *Science*, *255*, 556-559.

Baddeley, A. D. (1986). *Working memory*. Oxford, England: Oxford University Press.

Baerentsen, K. B., & Trettvik, J. (2002). An activity theory approach to affordance.

    Proceedings of 2nd *Nordic conference on Human-Computer interaction* . *31*,51-60.

Besnard, D. & Cacitti L. (2005). Interface changes causing accidents. An empirical study of

    negative transfer. *International Journal of Human-Computer Studies, 62(1),* 105-125.

Bos, B., Celik, T., Hickson, I. & Lie, H. H. (Working Draft November 6, 2006). Cascading

    Style Sheets, level 2 revision 1 CSS 2.1 Specification. *World Wide Web Consortium*,

    Retrieved December 8, 2006, from http://www.w3.org/TR/CSS21/.

Brown, D. C., & Blessing, L. (2005). The Relationship between function and affordance.

    Proceedings of ASME Design Theory and Methodology Conference, Paper #

    DETC2005-85017.

Card, S., Moran, T. & Newell, A. (1983). *The Psychology of Human-Computer Interaction.*

    Hillsdale, N.J: Erlbaum.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their

    mutual constraints within the human information-processing system. *Psychological*

    Bulletin, *104*, 163-191.

Desimone, R. & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual*

    Review of Neuroscience, 18, 193-222.

Ericsson, K. A. & Kintsch, W. (1995). Long-term working memory. *Psychological Review,*

    102, 211-245.

Gaver, W. W. (1991). Technology affordances. *Computer-Human Interaction Conference*

    Proceedings, 79-83.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton

    Mifflin.

Giddens, A. (1984). The constitution of society: Outline of the theory of structuration. University of California Press.

Gray, W. D. & Fu, W. (2004). Soft constraints in interactive behavior: the case of ignoring perfect knowledge in-the-world for imperfect knowledge in-the-head. *Cognitive Science*, *28*, 359-382.

Greeno, J. G. (1994). Gibson's affordances. *Psychological Review, 101*(2), 336-342.

Hannon, C. (2008, September). Mental and conceptual models, and the problem of contingency. *Interactions*, 58-64.

Hartson, H. R. (2003). Cognitive, physical, sensory, and functional affordances in interaction design. *Behaviour & Information Technology*, *22*(5), 315-338.

Hutchins, E. L., Hollan, J. D., & Norman, D. A. (1986). Direct manipulation interfaces. In D. A. Norman & S. W. Draper (Eds.), *User centered system design: New perspectives on human-computer interaction* (pp. 87 - 124). Hillsdale, NJ: Erlbaum.

Keppel, G. & Underwood, B. J. (1962). Proactive inhibition in short-term retention of single items. *Journal of Verbal Learning and Verbal Behavior*, *1*, 153-161.

Logan, G. D. (2002). An instance theory of attention and memory. *Psychological Review, 109*(2), 376-400.

Marcus, G. (2004). *The birth of the mind: How a tiny number of genes creates the complexities of human thought.* New York: Basic Books.

McGrenere, J. & Ho, W. (2000). Affordances: Clarifying and evolving a concept. *Proceedings of Graphics Interface*, 1-8.

Michaels, C. (2003). Affordances: Four points of debate. *Ecological Psychology*, *15*(2), 135-148.

Moors, A. & Houwer, J. D. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin, 132*(2), 297-326.

Nielsen, J. (Ed.), (2002). *Coordinating user interface for consistency*. Morgan Kaufmann: New York.

Norman, D. A. (1986). Cognitive engineering. In D. A. Norman & S. W. Draper (Eds.), *User centered system design: New perspectives on human-computer interaction* (pp. 31 - 61). Hillsdale, NJ: Erlbaum.

Norman, D. A. (1993). *Things that make us smart: Defending human attributes in the age of the machine.* Cambridge, MA: Perseus Books.

Norman, D. A. (1999, May). Affordance, conventions, and design. *Interactions*, 38-42.

Norman, D. A. (1988). *The psychology of everyday things*. New York: Basic Books.

Norman, D. A.& Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, and D. Shapiro (Eds.), Consciousness and self-regulation: Advances in research and theory, vol. 4 (pp. 1-18). New York: Plenum Press.

Oshlyansky, L., Thimbleby, H. & Cairns, P. (2004). Breaking affordance: Culture as context. *Proceedings of the third Nordic conference on Human-computer interaction*, 81-84.

Posner, M. I. & Snyder, C. R. R. (1975). Facilitation and inhibition in the processing of signals. In P.M.A. Rabbitt & S. Dornic (Eds.), *Attention and performance* V (pp. 699-682). New York: Academic Press.

Raskin, J. (1994, September). Intuitive equals familiar. *Communications of the ACM, 37*(9), 17.

Recarte, M. A. & Nunes, L. M. (2003). Mental workload while driving: Effects on visual search, discrimination, and decision making. *Journal of Experimental Psychology: Applied*, *2*(9), 119-137.

Rogers, Y. (2004) New theoretical approaches for Human-Computer Interaction. *Annual Review of Information, Science and Technology, 38*, 87-143.

Shiffrin, R. N. & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic, attending, and a general theory. *Psychological Review*, *84*, 127-190.

Shneiderman, B., & Plaisant, C. (2005). *Designing the user interface (Fourth Edition): Strategies for effective human-computer interaction*. Boston: Pearson Addison Wesley.

Simon, H. A. (1969/1981). *The sciences of the artificial* (2nd ed.). Cambridge, MA: MIT Press.

Sternberg, R. J. (1986). *Intelligence applied: Understanding and increasing your intellectual skills*. San Diego: Harcourt, Brace & Jovanovich.

Still, J. D. & Dark, V. J. (2008). An empirical investigation of affordances and conventions. In J. S. Gero and A. K. Goel (Eds.), *Design computing and cognition '08*, (pp. 457-472). Netherlands: Springer.

Strayer, D. L. & William, J. A. (2001). Driven to distraction: Dual-Task studies of simulated driving and conversing on a cellular telephone. *Psychological Science*, *6*(12), 462-466.

Strayer, D. L., Drews, F. A. & Johnston, W. A. (2003). Cell phone-induced failures of visual

    attention during simulated driving. *Journal of Experimental Psychology: Applied*,

    1(9), 23-32.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of*

    Experimental Psychology, *18*, 643-662.

Torenvliet, G. (July, 2003). We can't afford it!: The devaluation of a usability term.

    Interactions, 12-17.

Visser, W. (2006). *The cognitive artifacts of designing*. Mahwah, NJ: Erlbaum Associates.

Vogel, E. K., Woodman, G. F. & Luck, S. J. (2001). Storage of features, conjunctions, and

    objects in visual working memory. *Journal of Experimental Psychology: Human*

    Perception and Performance, 27, 92-114.

Vyas, D., Chisalita, C. M. & Veer, G. C. (2006). Affordance in interaction. *ACM Proceeding*

    of the 13th European conference on Cognitive ergonomics:Trust and control in

    complex socio-technical systems, *250*, 92-99.

You, H. & Chen, K. (2007). Applications of affordance and semantics in product design.

    Design Studies, *28,* 23-38.

Zhang, J., Patel, V. L. (2006). Distributed cognition, representation, and affordance.

    Cognition & Pragmatics, 14, 333-341.

CHAPTER 3. AN EMPIRICAL INVESTIGATION OF AFFORDANCES AND CONVENTIONS

This article was published in the *Proceedings of the Third International Conference on Design Computing and Cognition.* [1]

Jeremiah D. Still & Veronica J. Dark

Abstract

There is a debate in the literature concerning whether a distinction between affordances and cultural conventions ought to be drawn. It is possible that in the absence of affordances users develop conventions to resolve interaction ambiguity. We explored whether a difference between affordances and conventions existed through a button pressing task. Our results show that affordances exist when the spatial button configuration is congruent with directional cues. When affordances were not available, most participants demonstrated consistent button-to-action mapping that sometimes represented a convention. Additionally, there was no difference in response time in the affordance and convention conditions.

Introduction

A question of concern to designers is why some designs cause repeated failure while others facilitate error free use. That is, what aspects of a design lead the user to more easily comprehend the design's functions? In his book the Psychology of Everyday Things (POET), Norman (1988) suggested that good designs provide implicit visual instructions that

---

[1] This article may be republished by ProQuest Information and Learning through its UMI Dissertation Publishing business, given full citation to the original, with kind permission of Springer Science and Business Media - they maintain all rights.

constrain how an object should be used [1]. The constraints can be physical, logical, and/or cultural in nature. Because these implicit visual instructions are simply understood and do not demand much cognitive effort (limited working memory resources) to figure out the proper interaction to meet the user's task goals, Norman referred to them as *perceived affordances* in order to distinguish the concept from the more narrow concept of affordances as developed by J. J. Gibson (1970). Gibson coined the term to capture the idea that the physical structure of some object allows observers to directly perceive potential interactions [2].

In commenting on the many different uses of the term affordance in the design literature, Norman (1999) distinguished between perceived affordances in general and a subclass of culturally constrained perceived affordances that he called conventions [3]. Users cannot ignore affordances because they physically constrain interactions. However, conventions can be ignored because the constraint is learned, rather than physical. Thus, Norman suggested that there is a qualitative difference between conventions, or situations in which the mapping of visual cues to actions is an arbitrary choice of the designer, and affordances, or situations in which the mapping is a natural consequence of the physical world. He suggested that most computer systems are limited to a small number of built-in affordances (e.g., keyboard, pointing device, mouse buttons, etc.), and that these affordances are of little use for facilitating software application interactions. For example, because the user is not physically constrained to click on a specific icon, Norman cautioned that designers should not conclude that an icon within a desktop metaphor affords clicking. To do so would be an incorrect use of Gibson's affordance.

Although the desktop user is physically unconstrained to click on an icon or anywhere else within the visible screen, past experiences of the user will make the clicking on the icon a likely behavior. It is this type of culturally constrained interface interaction (i.e., one that depends on learning) that defines a convention. Another example of a convention is having learned that moving a scroll bar up moves the page up while viewing the document. The relationship between the scroll bar and the page movement is not transparent until the user moves the bar and receives visual feedback [3]. According to Norman, "This is what the interface designer should care about: Does the user perceive that clicking on that object is a meaningful, useful action, with a known outcome?"(p.40).

Implied by Norman's distinction between affordance and convention is that they should be treated differently by the designer. Others, however, disagree with Norman's conceptualization of conventions as distinct from affordances. According to McGrenere and Ho (2000), software application interfaces may afford specific interactions [4]. "The functions that are invokable by the user are the affordances in software… Norman claims that a scrollbar is a learned convention and implies that it is not an affordance. We disagree. The fact that the object affords scrolling is an affordance that is built into the software. The information that specifies this affordance is in fact a learned convention… [Acting on the bar will either move the page up or down]" (p. 6). Additionally, McGrenere and Ho described how users often customize their interfaces making affordances more efficient to undertake. They suggested that affordances are efficient when the execution of it is rapid, comfortable, and reduces exertion (e.g., hot keys or short-cuts).

Oshlyansky, Thimbleby and Cairns (2004) noted the lack of empirical work associated with perceived affordances in relationship with design [5]. They were interested in

how perceived affordances (conventions) vary across similar cultures. To empirically examine a potential cultural difference, Oshlyansky et al. surveyed US and UK students. The participants were asked to determine the current state of a bulb ("ON" or "OFF") based on an image of a light switch. The participants were not told if their responses were correct. The results showed that most participants from the UK reported a light switch in the down position as indicating that a light bulb was "ON" while most of the US participants reported a light switch in the down position as indicating that a light bulb was "OFF". The study provided strong evidence that students from different cultures interpreted the same very simple interaction (i.e., move a switch up or down) in opposite ways. Thus, Oshlyansky et al. stressed that designers need to consider previous knowledge when designing interfaces.

The light switch study assessed participant *knowledge* without providing feedback. The results suggested the existence of a different convention in the US and the UK regarding how to interact with a light switch. Our investigation focuses on how users *act* in a very simple situation that sometimes offers an affordance and sometimes does not. Our question is whether, when no affordance is available, there will be evidence of a convention. We could find no formal definition of a convention in the literature. Therefore, we defined a convention as 80% or more of the users responding in the same way, suggesting that they have learned a meaningful action based on past experience.

We empirically examined conventions (arbitrary button-to-action mappings) and affordances (natural button-to-action mappings) associated with the spatial configuration of two buttons (or keys) on when the user was acting on directional cues. We asked whether users would adopt consistent button-to-action mappings that might be the result of similar interface interactions within their everyday environments. Basically, the task for the user was

an analog of a person being asked to navigate a menu with two buttons on a controller. The questions concerned: 1) whether performance would reflect affordances when they were available, 2) whether conventions exist when affordances are not available, and 3) whether affordances and conventions would differ in their impact on users.

We explored the possibility that conventions exist within the context of a button pressing task. In this task two buttons were available and the participant was instructed to try to "move" in one of four specific directions (i.e., up, down, left, and right). We used the numeric keypad of a standard keyboard in which the buttons closest to the user's body were slightly lower than buttons farther from the user. The spatial configuration of the buttons was manipulated; participants were presented with two buttons arranged vertically, horizontally, or diagonally. The configurations are shown in Figure 1. We assumed that an affordance would be present when the spatial button configuration was congruent with the directional cue (affordance condition). Thus, the vertical buttons (panel A) afforded two directions (up, down) because one button was slightly higher than the other on the keyboard, the horizontal buttons (panel B) afforded two directions (left, right), and the diagonal buttons (panel C) afforded all four directions because the buttons were both above/below and to the side of each other.

We predicted that participants would make the afforded response when it was possible. Of interest was the nature of responses when affordances were not available (i.e., for left/right cues with a vertical orientation of buttons and for up/down cues with a horizontal orientation). In these instances, individual participants could respond in one of two-ways. The participant could use a consistent button-to-action mapping (e.g., a participant might consistently choose the right button when given the up cue and the left

button when given the down cue) or the participant could select buttons so that there is no

consistent pattern. If participants individually show consistency in their responses, then one

can ask if there is consistency between participants in the nature of the button-to-action

mapping. If a convention exists, then most participants should use the same consistent

button-to-action mapping.

A.   B.   C.
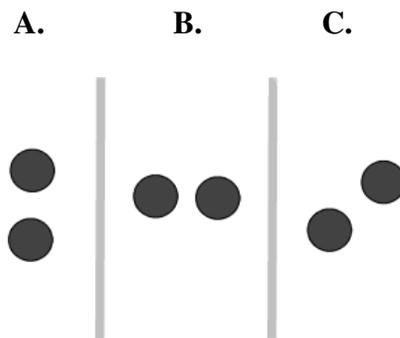
Figure 1. Each panel represents one of the two-button configurations used in the current
research.

Method

Participants

The university institutional review board approved all experimental procedures.

Thirty undergraduate volunteers (28 right handed, 14 females) were recruited to participate in

exchange for course research credit in an introductory psychology course. Each participant

had normal or corrected to normal vision.

Stimuli and Apparatus

The trials were presented on an HP Pentium 4 Windows XP machine with a 17 inch

monitor. On each trial, one of four directional cue words (Up, Down, Left, or Right) was

presented centrally in size 48 Arial font. Responses were collected through a PS/2

keyboard's numeric keypad. Only three pairs of keys (six buttons) were used and those keys

were covered with the appropriately colored stickers as is shown in Figure 2. Red stickers

represented the diagonal configuration by covering keys 5 and 9. Blue stickers represented

the horizontal configuration by covering keys 1 and 2. Yellow stickers represented the

vertical configuration by covering key 3 and 6. This study was created and executed within

E-prime experimental presentation software (Psychology Software Tools, Inc.,

www.pstnet.com).
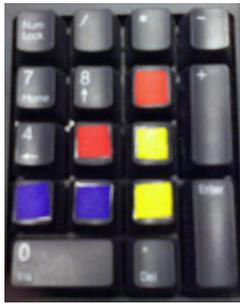
Figure 2. Photograph of the three configurations of two-button pairs as implemented on the keyboard.

Participants were run at individual computer stations in groups of up to four. They

wore sound deadening earmuffs and there were dividers between computer stations so that

participants could not see each other.

Procedure

Participants faced an ambiguous situation in which they were not instructed *how to*

act on the given button configuration, nor was any feedback given. Participants were told,

"You will be instructed to place your fingers onto the specified color. Then you will be given

a directional cue to either move up, down, left or right. You should move in the direction

indicated to the best of your ability given the available button configuration. Your response

times are going to be recorded. Please, respond as accurately and as quickly as possible!" At

the beginning of each color block they were also given these instructions, "Please place your fingers on the [Red, Yellow, or Blue] buttons. Respond as accurately and as quickly as possible. Press one of the [Red, Yellow, or Blue] buttons to continue." These instructions were intended to encourage participants to make speeded responses and to indicate that one button press was in fact correct for each directional cue. The instructions also ensured that participants' fingers were on the correct buttons before the onset of a block of trials.

Spatial configuration of buttons was manipulated across blocks in a within-subject factorial design with three spatial configurations of the buttons (vertical, horizontal, diagonal) and four directional cues (up, down, left, right). Only two buttons were available for response within each block. There were three blocks: vertical (yellow buttons), horizontal (blue buttons), and diagonal (red buttons). Block order was counterbalanced across participants such that there were three possible block orders (yellow, blue, red; blue, red, yellow; red, yellow, blue) and each participant was assigned to one order. Each block contained 80 trials, 20 with each directional cue. The order of cues within a block was randomized for each participant.

After the instructions for a block were presented, the participant would view a 'Get Ready' slide for 2000 msec which was directly followed by the presentation of a randomly selected directional cue (Up, Down, Left, or Right). Upon seeing the directional cue, the participant would press one of the possible buttons and the next trial would begin. The procedure is illustrated in Figure 3. Participants completed the experiment in approximately 20 minutes and were fully debriefed at the end.
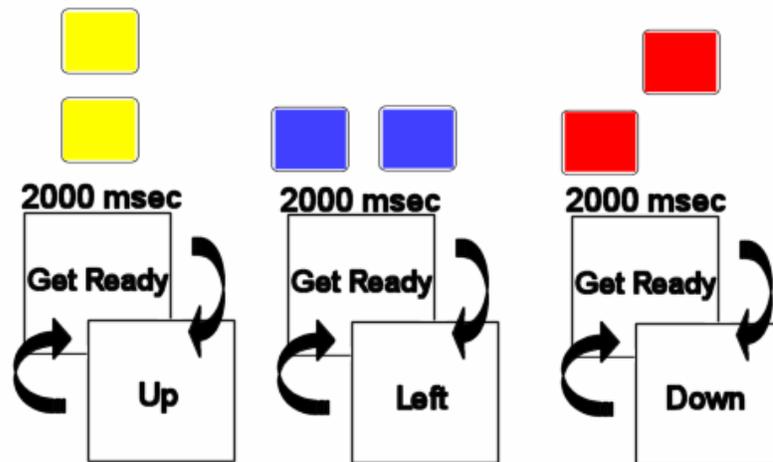
Figure 3. Schematic representation of the experimental procedure. The arrows indicate the change of displays. The three colors and button configurations represent the three blocks of trials. For each button configuration, a participant was presented 20 trials, each of which began with 'Get Ready' for 2000 msec and then a directional cue (up, down, left, or right) would appear until the participant responded with a button press. The 'Get Ready' slide then appeared, indicating the onset of the next trial.

Results

All statistical tests used an alpha level of .05. Error bars in the figures represent the mean standard error. In addition to recording which button was pressed, response times were recorded. Responses on specific trials were excluded from the analysis if reaction times were less than 200 msec or greater than 2000 msec. This filtering only removed 2% (150 outliers) of the total data. Three dependent variables were examined: response choice, consistency of choice, and response time.

Response Choice

Participants were faced with a binary response decision on every trial. There were three different sets of binary responses (i.e., three spatial configuration of buttons) and four different cues. For each combination of spatial configuration and cue, we calculated the proportion of times that each response was chosen. Because of the binary nature of the

response, a complete picture of the results can be derived from consideration of the proportion of trials of which one of the two responses is made. We had predicted that affordances were present for all cues in the diagonal condition and that an affordance was present only for the up/down cues with the vertical configuration and only for the left/right cues with the horizontal configuration. The proportions of trials on which the afforded response was made in these conditions are presented in solid bars in Figure 4. For the nonaffordance conditions (i.e., left/right cue with the vertical orientation and up/down cues with the horizontal orientation), the response yielding the highest group proportion is displayed in a patterned bar. The associated response is indicated below the bars.
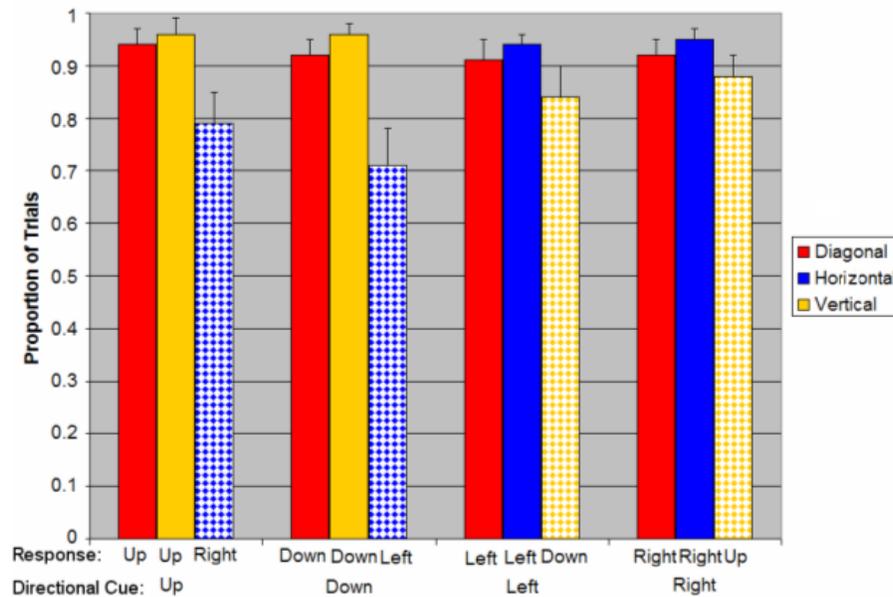

Figure 4. Proportion of trials on which the group dominant response was made as a function of each combination of button configuration and directional cue. The response under each bar is the dominant choice for that combination. The solid bars represent affordance conditions and the patterned bars represent the nonaffordance conditions.

The first thing to notice is that the proportion of what we had assumed were afforded responses (solid bars) is above 0.90 for all button configurations, suggesting that we were

correct in our assumption that affordances were present. The second thing to notice is that although the proportions in the nonaffordance conditions (patterned bars) were well above 0.50 (chance), suggesting regularity in how participants were responding in these situations, the proportions appeared to be lower than in the affordance conditions. The third thing to notice is that the difference in proportions between the affordance and nonaffordance conditions appears to be stronger for the up/down cues than the left/right cues.

A 4 (cue) x 3 (type of affordance: diagonal affordance, horizontal/vertical affordance, no affordance) within-subjects Analysis of Variance (ANOVA) confirmed that the apparent patterns were real. There was a main effect of type of affordance, $F(2,58) = 9.175$, $MSE = .079$, $p < .001$, reflecting lower proportions in the no affordance conditions, however, the Type of affordance x Cue interaction also was significant, $F(6, 174) = 2.27$, $MSE = 0.026$, $p = .039$. A separate ANOVA on just the up/down cues showed a main effect of type of affordance, $F(2, 58) = 8.64$, $MSE = 0.086$, $p = .001$. Comparisons among the means showed that performance was lower in the no affordance conditions ($M = 0.746$) than in the diagonal affordance ($M = 0.925$) or vertical affordance conditions ($M = 0.952$), which did not differ. A separate ANOVA on just the left/right cues showed only a marginally significant main effect of type of affordance, $F(2, 58) = 2.64$, $MSE = 0.045$, $p = .08$, indicating only a trend for lower performance in the nonaffordance condition.

Consistency of Choice

As expected, participants were influenced by the affordances when they were available. They also appeared to be responding in a consistent manner in the nonaffordance conditions. As noted earlier, the observed consistency could reflect either a general tendency within each individual to make one particular response to each cue or it could reflect a

difference in the number of individuals consistently choosing one response with some

individuals consistently choosing the other response. In order to determine which might be

going on in our data, participants were placed into one of three categories for each

combination of cue and configuration. We arbitrarily defined "consistent" responding as

making the same response to a given cue 80-100% of the time. Thus, participants could be

consistent in making the group dominant response (which is labeled > 79% in the figure), or

consistent in making the opposite response (which is labeled < 21%), or they could be not

consistent (which is labeled NC). The number of individuals falling into each category for

each cue and configuration combination is shown in Figure 5. These categorization data

show that even in nonaffordance conditions, individuals are performing in a consistent

fashion.
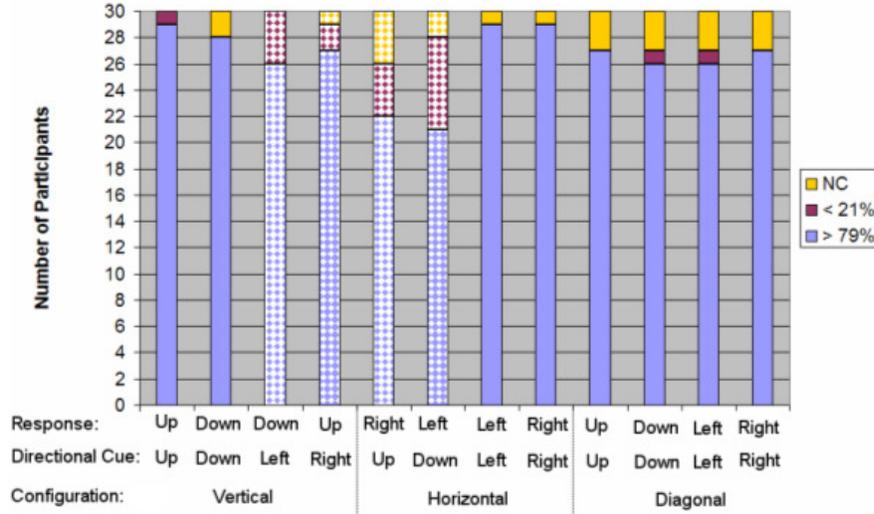
Figure 5. Distribution of individuals falling into each of the three consistency categories as a function of each combination of button configuration, cue, and response. Solid bars represent affordance conditions and patterned bars represent nonaffordance conditions.

The data presented in Figure 5 confirm the interpretation offered for the response

choice data. Affordances exist in the spatial configuration of the buttons. In an ambiguous

situation in which individuals are told to make a response indicating a movement, a button

that matches the direction of the movement affords a response. We believe that the data also

support the existence of conventions. A convention is present when the majority of

participants consistently choose the same response when no affordance is present. We

operationally defined 80% agreement over participants as indicating the existence of a

convention. By this definition, there is a convention in place for mapping right to an up

response and left to a down response. The trend for mapping up to a right response and

down to a left response is weaker and does not constitute a convention by our definition.

Response Time

Affordances reflect physical information and, therefore, responses to affordances

might be faster than responses reflecting a convention (right/left responses in the vertical

configuration) or a nonaffordance based consistent response (up/down responses in the

horizontal configuration). A difference in response time would support Norman's (1999)

suggestion that designers treat affordances and conventions differently. Our final analysis

compared the response times in the different cue and configuration combinations for those

individuals who consistently gave the most frequent group response in the nonaffordance

conditions. Seven participants were excluded from the analysis of response time to the

up/down cues and four were excluded from the analysis of the right/left cues because their

responses were inconsistent or did not follow the pattern shown by the group as a whole. The

mean response times are shown in Figure 6. Although for both the up/down cues and the

right/left cues, the responses were numerically faster in the affordance conditions than the

nonaffordance conditions, the differences were not statistically reliable. A within-subjects

ANOVA on the response times to the up/down cues with cue direction and type of

affordance as variables showed a significant main effect of cue direction, $F (1, 22) = 8.28$,

$MSE = 10,626$, $p < .001$, in which responses to the up cue were faster ($M = 560$ msec) than

responses to the down cue ($M = 580$ msec). Neither the main effect of type of affordance nor

the interaction effect were significant.

A similar ANOVA on the response times to the right/left cues showed no significant

differences. We defined the responses to right/left cues in the vertical configuration as

reflecting a convention. Although one must be cautious in drawing conclusions from null

effects, the similar response times in the convention and affordance conditions in conjunction

with the fact that the response times were rather quick, suggests that the choices in both cases

were made quickly without much conscious effort.
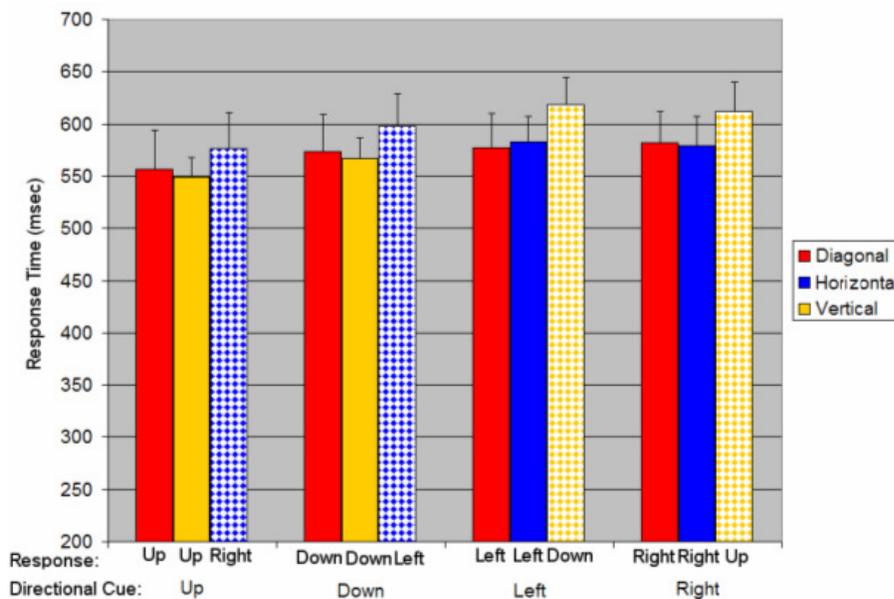
Figure 6. Response time for participants making the group dominant response as a function of each combination of type of affordance and directional cue. The response under each bar is the dominant choice for that combination. The solid bars represent affordance conditions and the patterned bars represents the nonaffordance conditions. Seven participants were omitted from the analysis of the up/down cues and four participants were omitted from the analysis of the left/right cues.

Discussion

We explored the influence of affordances and conventions, which reflect previous knowledge, on a task in which participants made a speeded response to "move" in a specific direction using two buttons at different spatial configurations. The experiment was designed to answer several questions. One question was whether participants would use an affordance when it was available. The answer is yes. Each button configuration had at least one pair of affordances and in each case, the afforded response was the one made most consistently by participants. We note that performance with the diagonal configuration was just as high for *all* four directional cues as it was when the buttons only afforded two directions (up and down or left and right). Such data indicate that the diagonal configuration is an optimal button configuration to convey four directions.

In addition to showing that participants acted on the provided spatial affordances, the data reflected the fact that most participants' button-to-action mappings were consistent. It is likely that this consistency arose from logical constraints imposed by the binary task. Because the four directions consisted of pairs of opposites, a logical constraint emerges when a participant selects one button to use in responding to one half of the pair, then the other button should logically be used for the other half of the pair. So, for example, when a participant chooses the top button of the vertical button set in response to right directional cue, by logic the bottom button should be used for left. In addition to the within-subject consistency in responses, we found consistency within the group. That is, most participants chose the same button to reflect the same cues. These data strongly suggest that the actions reflected learning that may have resulted from similar experiences interacting with common technologies.

Finding no guidelines in the literature as to how much within-group consistency is needed to define the existence of a convention, we arbitrarily used an 80% criterion. Response to only one of two spatial configurations of buttons satisfied this criterion for the existence of a convention in the nonaffordance conditions – the vertical button configuration. Given the directional cue to move right, the conventional response was to press the top button when using the vertical button configuration. We conjecture that this conventional mapping may be a reflection of common usage. For example, in US automobiles, the turn signal is typically to the left of the steering wheel and is designed so that an up action signals a right turn and a down action signals a left turn. Experience driving leads to an association in long-term memory between the directions and the actions. Figure 7 illustrates another instance of this strong nonaffordance mapping (i.e., this convention) within a remote control interface design (see Figure 7). Thus, previous experience with button-to-action mappings like this could explain the existence of our convention.

Figure 7. Example of a remote control device using a two-button configuration for four directions. Notice that the mapping of direction to the button configurations is congruent with our results.

Basically, the idea is that based on such experiences, when faced with the ambiguous situation of pressing one of two vertically arranged buttons to indicate left, the participant

"knows" that the appropriate answer is to press the down button. The nonaffordance conditions associated with the horizontal button configuration reflected high within-group agreement, but the agreement was not sufficiently strong for us to conclude that a convention was in place. The lower within group agreement may reflect either that competing mappings exist in daily activities or that the mapping is not common.

The message for designers from our data is that users are far from being blank slates. As noted by Vyas, Chisalita, and Van der Veer (2006), "During the technology use, users continuously interpret and reconstruct the meanings related to the technology, …" (p. 93) [6]. In even simple actions like button pressing, they are biased to prefer one of two binary responses. In some situations, the bias may be widespread, indicating the presence of a convention. Designers need to know the nature of conventions that might have an impact on their designs. In other situations, the bias may be less pervasive, but even then each individual may have a bias; even if users are not following a convention, they may still be acting consistently. In other words, if users are performing poorly on a design that uses popular conventions it could be a result of their following a minority consistent strategy. Regardless, when users act according to a consistent strategy, that strategy may have been previously learned and could therefore be very difficult to retrain. In addition, violating the users' previous knowledge could cause accidents [7]. Finally, we note that if conventions are not stable across device interactions, it could result in a continuous battle between interface designs leading to negative transfer for both designs (reflected by numerous errors for both designs). Therefore, it is important to know the users' current understanding of any conventions in place for there to be successful design outcomes.

As already noted, our prediction that diagonally offsetting the buttons would produce a situation in which there was an affordance for all four directions was supported. No difference was found between the diagonal configuration and any of the other affordances conditions. Thus, according to Fitt's Law, the diagonal button configuration is the best design choice for moving in the direction of up, down, left, and right. In spite of this, we do recognize that this is only optimal when only one of the directions is possible at a time (i.e., down/up or left/right). Another unique action would have to be included in order to discriminate between user's wanting to move up rather than right or vice versa.

In summary, we have demonstrated in a simple interaction that conventions may exist. We would like to stress the importance of understanding the users' previous experiences when there is the possibility of a convention based interaction. Ignoring the users' previous knowledge may result in repeated errors. Further, the previous experience may have developed into a convention. Research has shown that such highly practiced consistent mapping situations are difficult to retain because they are no longer under conscious control [8].

Our results do not provide evidence that affordances allow faster responses, at least in the simple task we examined. Prior learning of a button-to-response mapping removes ambiguity from the interaction. The learned mapping is stored in long-term memory, so when a similar instance of interaction arises again, relevant information in long-term memory becomes active and this leads to more efficient processing of the situation and the response. Neisser (1976) referred to this as the perceptual cycle in which past experience drives current perception [9]. Our data suggest that even though the designer may not see a visually constrained interface, the user may. When an affordance is present, the designer is likely to

view the situation just as the user does. When an affordance is absent, the extent to which the designer and user view the situation in a similar fashion will depend on the extent to which they have had similar past interactions that lead to a similar perceptual cycle. In these situations, the designer needs to consider the possibility that a convention exists, a point also made by Oshlyansky, Thimbleby and Cairns (2004).

From a user-centered design perspective, there are two major reasons for recognizing the concurrent conventions in play. First, requiring the user to repeatedly learn new interface conventions consumes the users' working memory which, in return, slows their information processing system. Good interface design rapidly becomes pervasive in supporting the users' task goals (i.e., success in the interaction does not need to itself become a goal). Second, there may be operating conditions (e.g., non-optimal arousal, working memory is filled with primary task calculations, etc.) in which a user does not have full access to working memory. Under such conditions, an interaction may be slow or fail. Thus, a good interface design should take advantage of knowledge stored in long-term memory so that the user's interaction with the interface is more effortless and more successful. In short, considering users' conventions during the design process will lead to increased usability by minimizing interference with existing applications.

Acknowledgments

Original Springer publication citation

Still, J. D. & Dark, V. J. (2008). An empirical investigation of affordances and conventions. In J. S. Gero and A. K. Goel (Eds.), Design Computing and Cognition '08, (pp.457-472). Netherlands: Springer.

References

1. Norman DA (1988) The psychology of everyday things. Basic Books: New York

2. Gibson JJ (1979) The ecological approach to visual perception. Houghton Mifflin: Boston

3. Norman DA (1999) Affordance, conventions, and design. Interactions, May issue, pp. 38-42

4. McGrenere J, Ho W (2000) Affordances: clarifying and Evolving a Concept. Conference Proceedings of Graphics Interface 1-8

5. Oshlyansky L, Thimbleby H, Cairns P (2004) Breaking affordance: culture as context. Conference Proceedings of the third Nordic conference on Human-Computer Interaction 81-84

6. Vyas D, Chisalita CM, Van der Veer GC (2006) Affordance in Interaction. ACM Proceeding of the 13th European conference on Cognitive ergonomics: Trust and control in complex socio-technical systems 250: 92-99

7. Besnard D, Cacitti L (2005) Interface changes causing accidents: an empirical study of negative transfer. International J of Human-Computer Studies 62(1): 105-125

8. Shiffrin RN, Schneider W (1977) Controlled and automatic human information processing: II perceptual learning, automatic, attending, and a general theory. Psychological Review 84: 127-190

9. Neisser U (1976) Cognition and reality. San Francisco: W.H. Freeman

CHAPTER 4. EXAMINING WORKING MEMORY LOAD AND CONGRUENCY EFFECTS ON AFFORDANCES AND CONVENTIONS

This article was prepared for the *International Journal of Human-Computer Studies.*

Jeremiah D. Still & Veronica J. Dark

Abstract

It has been debated whether designers should draw a distinction between perceptual affordances and cultural conventions. There is little behavioral evidence for either side of the debate. We investigated the cognitive attributes of acting on affordances and conventions through manipulating working memory load and expected button-to-action mapping congruency. The findings suggest both sides of the debate are correct. There was a behavioral difference between acting on affordances and acting on conventions. However, like affordances, learned conventions were found to structure responses towards expected actions. Therefore, designers ought to employ perceptual affordances when possible and when not feasible they ought to reuse established conventions. Additionally, evidence is presented that violating expected affordance-based and convention-based button-to-action mappings caused a similar performance cost. We believe that after the initial learning period, conventions play a critical role in the perception of a design's available actions just as do perceptual affordances.

Introduction

Background

A question of concern to designers is why certain interfaces lead users to effortless interactions and others to frustration. According to Norman (1988), perceived affordances

facilitate this effortless experience through implicit guidance towards a design's available actions. The concept of perceptual affordance was originally introduced by J. J. Gibson (1979), who coined the term affordance to refer to the function an object clearly offers within an environment. In natural environments, affordances emerge from physical constraints during user and interface interactions (e.g., a key on the keyboard affords pressing). A designer is able to take advantage of affordances within in a virtual world by maintaining naturally occurring characteristics within that world (e.g., including 3-D cues, like shadows, on a virtual keyboard). However, it is often the case that characteristics associated with a perceptual affordance are not available within an interface. When perceptual affordances are not available, designers are required to use conventions.

Conventions are initially arbitrary mappings between an action and a function within an interface that become relatively common within a group (Norman, 1999). A point of contention in the affordance literature is whether or not designers should be concerned with the distinction between perceptual affordances, which are presumed to be unlearned, and conventions, which are presumed to be learned (McGrenere & Ho, 2000; Norman, 1999; Oshlyansky, Thimbleby & Cairns, 2004). Some authors claim that perceptual affordances are unique and not comparable to conventions. However, there have been few empirical studies contrasting affordance and convention interactions (see Still & Dark, 2008, for one such study). We believe that after the initial learning period, conventions play a critical role in the perception of a design's available actions, just as do perceptual affordances.

Conventions usually develop within a design that is constrained in such a way that embedding a perceptual affordance is impractical or impossible. It is reasonable to assume that in the absence of perceptual affordances, users would learn conventions to resolve

interaction ambiguities. Basically, a convention is an interaction that emerges when multiple users each form a learned association between an action and function in the absence of a perceptual affordance. The associative strength between the action and function will vary as a function of interaction frequency. When the association begins to influence user expectations, it can be considered a convention. In fact, in our view, the associative strength between an action and function may grow so strong that it becomes a perceived affordance. For example, clicking an underlined-word hyperlink affords the presentation of another web page. Once the user learns the interface convention through enough interaction instances, the association between action and function becomes expected. The user in the hyperlink example strongly expects that underlined items on a new web page will be clickable hyperlinks.

Designers employ perceptual affordances because the available actions appear to be immediately perceived by the user. In other words, a user acting on an affordance does not have to consciously think about what action is available; it is just apparent. In cognitive terms, the action is directly available without mental effort; it is available without the need to expend limited-capacity, working memory resources. The current study was designed to examine the effect of variations in working memory resources on actions associated with affordances and conventions. Participants engaged in button interactions while holding varying amounts of information in working memory. Examining the effect of working memory load on affordance-based and convention-based interactions will provide insight into the cognitive attributes of perceptual affordances and conventions.

Designers respect perceptual affordances within their designs and do not violate affordance-based visual cue-to-action mappings because they know that design interactions

would suffer. Less is known about the cost to performance of violating convention-based expectancies in comparison to the cost of violating affordance-based expectancies. Knowing whether or not there are differences will inform designers about whether they need to consider the impact of violating conventions as much as they consider the impact of violating affordances.

Another question examined in this article is the ease with which participants are able to overcome interference when they engage in interactions that violate affordance-based expectancies versus convention-based expectancies. Learning new information often requires deliberate processing, which requires limited working memory resources. In this article, we compared accuracy and response time of interactions related to affordances and conventions when both working memory load and expectancy were manipulated within a spatial target selection task. Our goal was to gain insight into the cognitive attributes of using a perceptual affordance in contrast to a learned convention within a design (Visser, 2006).

Affordances vs. Conventions

The affordance versus convention debate is fueled by the introduction of virtual interfaces that lack the traditional physical characteristics of affordances. Such virtual interfaces have led to the question of whether the term affordance should be reserved for only physical objects that provide a natural mapping between their design and available functions. According to Norman (1999), the virtual desktop does not physically constrain the user's interaction to specific actions, like clicking only on 'available' buttons. He stressed that designers should refer to the learned constraint of actions (e.g., clicking on virtual buttons) as conventions and that the term affordance should be reserved for naturally occurring physical

constraints like clicking a mouse button. Norman's distinction suggests a qualitative difference for designers between conventions and affordances.

In contrast, others believe that the term affordance may be adopted within the context of virtual interfaces. According to McGrenere and Ho (2000), affordances in the software are learned by users and affordance-based and convention-based interactions are not qualitatively different. "The functions that are invokable by the user are the affordances in software… Norman claims that a scrollbar is a learned convention and implies that it is not an affordance. We disagree. The fact that the object affords scrolling is an affordance that is built into the software. The information that specifies this affordance is in fact a learned convention… [Acting on the bar will either move the page up or down]" (p. 6) It is clear by these authors' statements that they believe conventions, once learned, afford actions.

On the one hand, if McGrenere and Ho (2000) are right, there may be little difference between interactions based on physical affordances and those based on conventions. On the other hand, if Norman (1999) is correct, there may be important differences between how interactions based on affordances and conventions are processed by users. We intend to shed some light on whether or not there is behavioral support for the distinction between these two design concepts.

One can ask the question whether conventions, once learned, are behaviorally the same as affordances. We first investigated this question using a simple button task (Still & Dark, 2008). In this task, user responses were constrained to two buttons, but users were required to respond to four different direction words defining two different direction pairs: up - down and left - right. Our research goal was to capture *current* expectations of button-to-action mappings, so feedback on button pressing was not given. The spatial layout of the

buttons was such that each spatial layout "afforded" responses that matched either one or both of the direction pairs. The vertical configuration "afforded" actions to the up-down pair. The horizontal configuration "afforded" actions to the left-right pair. The diagonal configuration "afforded" actions to both pairs (i.e., to all four directions). For example, given the vertical button configuration, the top button afforded the direction up and the bottom button afforded the direction down but there were no affordances for left and right.

We found that even when no affordance was available, participants demonstrated consistent button-to-action responses. There were two types of consistency. First, individual participants were logically consistent within a direction pair in that they mapped one button to one direction within each pair. So, left was mapped to one button and right was mapped to the other. Similarly, up was mapped to one button and down was mapped to the other. Second, the nature of this mapping was consistent across participants. The mappings for the affordance conditions were as expected (e.g., the left button was mapped to left) and were the same for almost all the participants. Of more interest was the finding that there was consistency across participants in how the nonaffordance conditions were mapped. Over 80% of the participants mapped the top button in the vertical configuration to the direction right (and the bottom button to the direction left), and 70% of the participants mapped the right button in the horizontal configuration, to the direction up (and the left button to the direction down).

In our simple button experiment, no difference in participant performance (response agreement, response consistency, response time) was found between conventions and affordances (Still & Dark, 2008). It appeared that the vertical top button afforded moving right as much as it afforded moving up. From these results we suggested that affordances and

conventions may be behaviorally indistinguishable. Thus, from a design standpoint, the results showed that although the decision to represent four directions with either a horizontal or vertical button configuration would seem to be arbitrary, one choice better fit the users' current expectations.

Overview of Experiments 1 & 2

In this paper, we report data collected in one session from one set of participants, but participants engaged in two versions of a button-pressing task, one with and one without feedback. We will describe these two different participant experiences as Experiment 1 and Experiment 2.

Still and Dark (2008) was one of the few studies to empirically compare performance with affordances and conventions. In Experiment 1, we replicated the basic procedure of this earlier study to see whether the results would replicate. That is, we wanted to test whether participants began the experiment with direction-to-action mapping expectancies (e.g., top mapped to right) already in place. As before, participants faced an ambiguous situation in which they were not instructed how to act on the given button configuration, nor was any feedback given. To foreshadow the results, we did replicate our original findings, supporting the existence of conventions. In Experiment 2, we explored the effect of working memory load and mapping congruency on participants' ability to perform actions based on affordances and conventions. We achieved this by modifying our simple button pressing experiment so that there was a correct answer and feedback was given. To foreshadow the results, we found differences in how these variables affected affordances and conventions.

Experiment 1

Introduction

Still and Dark (2008) demonstrated that a convention existed in a simple button-pressing situation. For example, as already described, when asked to press a top or bottom button in response to the direction cue "right", 80% of participants pressed the top button. Participants in the current experiment were faced with the same button task employed by Still and Dark, however, for practical reasons the number of trials for each direction cue under each button configuration was reduced from 60 to 12. Our intent was to capture participants' extant expectations concerning button-to-action mappings.

Method

Participants

Twenty-seven undergraduate volunteers participated[1] in exchange for course research credit (females = 16). Only one participant reported being left hand dominant. All participants reported having normal or corrected to normal vision.

Stimuli and Apparatus

The experiment was presented on an HP Pentium 4 Windows XP machine with a 17 inch monitor. Responses were collected through a PS/2 keyboard's numeric keypad. Only three pairs of keys (six buttons) were used and these keys were covered with colored stickers. The same experimental setup used in the Still and Dark (2008) button pressing experiment

---

[1] Data from an additional three participants were excluded from all analyses because their target selection performance was two standard deviations under mean participant performance in the congruent affordance conditions (i.e., 80% target selection accuracy) of Experiment 2.

was used. Red stickers covered keys five and nine representing the diagonal configuration. Blue stickers covered keys one and two representing the horizontal configuration. Yellow stickers covered keys three and six representing the vertical configuration. The stimuli were created and executed within E-prime experimental presentation software (Psychology Software Tools, Inc., www.pstnet.com).

Participants were run at individual computer stations in groups of up to three. They wore sound deadening earmuffs and there were dividers between computer stations preventing the participants from viewing or hearing each other.

Procedure

Participants were presented with an ambiguous situation in which they were not provided with any feedback on their button pressing actions. Participants were instructed at the beginning of each block to place their fingers on the appropriate buttons (i.e., vertical, horizontal or diagonal). Further, they were told to use these buttons to move in the cued direction to the best of their ability given the current spatial button configuration. Participants were randomly presented a directional cue: up, down, left or right that remained on the screen until a response was given (see Figure 1). They were encouraged to respond as accurately and as quickly as possible. These instructions were intended to encourage participants to make speeded responses and to indicate that one button press was in fact correct for each directional cue. Participants acted on 12 directional cues per button configuration block, yielding a total of 36 trials. Block order was counterbalanced across participants. Experiment 1 took approximately 3 minutes to complete.
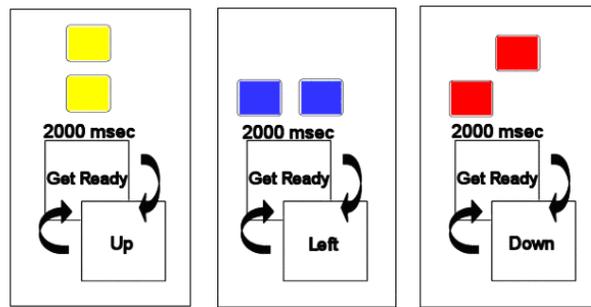
Figure 1. Schematic of the experimental procedure. The arrows indicate the flow of displays. The three colors and button configurations represent the three blocks of trials. For each button configuration, a participant was presented 12 trials, each of which began with 'Get Ready' for 2000 msec and then a directional cue (up, down, left, or right). The cue was displayed until the participant responded with a button press. The 'Get Ready' slide then appeared, indicating the start of the next trial.

Results & Discussion

Examination of which button was chosen provided a measure of the participants' expectations of button-to-action (direction) mappings given vertical, horizontal and diagonal button configurations. The mappings were obtained without feedback and therefore should reflect prior expectancies. We classified the response patterns for each direction pair (up/down, left/right) with each button configuration into one of three button-to-action types (congruent, inconsistent or incongruent) based on our previous findings (Still & Dark, 2008). The 12 trials with each configuration were first examined for logical consistency. Logically consistent responses were those in which one button was mapped to each of the opposite directions in a pair (i.e., one button was associated with up and the other with down; one button was associated with left and the other with right.) For a given configuration, the responses were considered logically consistent if at least two out of three responses for each direction in a direction-pair were mapped to the same button and logically inconsistent otherwise. The logically consistent sets of responses were further classified as congruent if

the button-to-action mapping matched the majority pattern found in the earlier work and incongruent if the mapping matched the opposite pattern. Table 1 displays the number of participants falling into each category as a function of button spatial configuration, direction pair, and button-to-action mapping.

Table 1. Number of participants in each category as a function of button configuration and direction pair.

Category	Spatial Button Configuration and Direction Pair					
	Vertical		Horizontal		Diagonal	
	Up/Down (Affordance)	Left/Right (Convention)	Left/Right (Affordance)	Up/Down (Convention/Bias)	Up/Down (Affordance)	Left/Right (Affordance)
Congruent	26	20	26	17	25	26
Incongruent	0	4	1	8	1	0
Inconsistent	1	3	0	2	1	1

The affordance button-to-action mappings were highly agreed upon across button configurations. This is probably due to the perceptual matching between up/down or left/right and the physical buttons' spatial locations. Although the nonaffordance conditions lacked the aid of perceptual cues, there still was a high level of participant agreement. The fact that very few people showed an inconsistent response pattern indicates that there was a logical constraint influencing participants' button-to-action mappings, just as was found in our previous work (Still & Dark, 2008). Also, as shown by Still and Dark, a large majority of participants agreed that the top vertical button should be mapped for moving right and the bottom button for moving left and a smaller majority of participants agreed that for left horizontal button should used for moving down and the right button for moving up.

In both our current and previous results, participants appear to have expectancies about which button to press in the nonaffordance conditions and the expectancy across participants is stronger for the vertical than the horizontal button configurations. In Still and

Dark (2008), we arbitrarily chose 80% agreement across participants as the definition of a convention. But, just as in the earlier work, we found that even with the horizontal configuration, the choice was not arbitrary. Clearly, participants are not responding randomly in the nonaffordance conditions with either configuration. We now believe that choosing an arbitrary point, like 80% agreement, to define a convention is establishing a dichotomy when none exists. The expectancies in this task are quantitatively rather than qualitatively different. Therefore, in the remainder of this article, we will treat all nonaffordance conditions as conventions. When we want to draw attention to the fact that the nonaffordance expectancies associated with the horizontal configuration are weaker than those associated with the vertical configuration, we will refer to those weaker expectancies as a "bias". Figure 2 summarizes the empirical findings
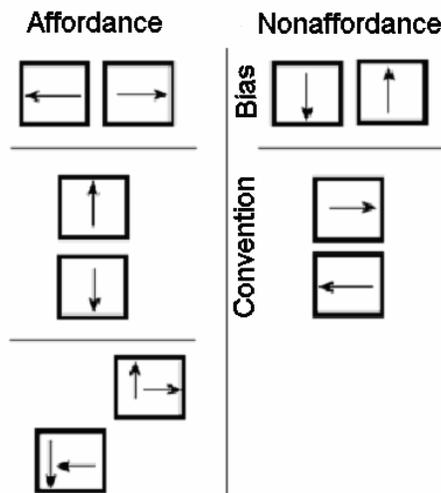

Figure 2. A visual representation of button-to-action empirical findings. The squares represent the button spatial configuration and the arrows show the mapping of the buttons to the directions.

We hypothesize that conventions and biases within a user group are indicative of the differing amounts of interaction regularity experienced by the user. The results suggest that

users may have learned that across numerous devices the top buttons are typically associated with moving right within vertically configured buttons, thus, leading to existence of a strong convention at the group level. A group bias could arise for the horizontally configured buttons if most devices map the right button with moving up but a sizable proportion also map the right button with moving down. We predict that the effect of multiple interactions with devices in this case will differentially influence individual users' button performance with some users adopting the more frequent interaction mapping and others adopting the less frequent interaction mapping. Thus, we suggest that strong conventions are associated with consistent learning episodes across devices and that those episodes lead to strong associations between button and action mappings. In contrast, biases are associated with more varied learning episodes between button and action mappings and the variation leads to a weaker association.

Experiment 2

Introduction

As the results of Still and Dark (2008) and Experiment 1 suggest, it is important for designers to understand the users' previous experiences because that knowledge determines participants' expectancies and which interactions will be most effective. For example, if a new interface requires responses that conflict with an already learned convention, users may be more likely to make repeated errors. The repeated errors would be the result of responses being directly shaped by long-term memory structures built during past interactions (see Chapter 2). Research has shown that such highly practiced consistent mapping situations are difficult to retrain because they are automatic and no longer under conscious control (Besnard, & Cacitti, 2005; Moors, & Houwer, 2006; Norman, & Shallice, 1986; Shiffrin, &

Schneider, 1977). To overcome the previously learned interaction requires monitoring and rebuilding of memory structures through controlled processing. Controlled processing requires working memory resources.

The Role of Working Memory

Users only have access to a limited amount of working memory resources and the availability of these resources is critical for minimizing cognitive effort and task errors (Reason, 1990). Working memory is most often described as a temporary storage usually discussed as a mental workspace for conscious processing (Baddeley, 1992). Although working memory and long-term memory have been discussed as two separate systems, it is commonly assumed that the two systems interact (e.g., Baddeley, 2002). Long-term memory structures mold current representations being operated on within working memory. Another conceptualization of working memory was provided by Cowan (1988), who suggested that short term storage is actually highly activated regions within long-term memory. By this view the representations available to working memory depend on knowledge and structures already present in long-term memory.

An understanding of working memory and its role in interactions has practical implications for design. For example, novice chop stick users must spend a lot of working memory resources towards correctly operating chop sticks when eating noodles. That is, the user needs to focus on how to hold the chop sticks in a way that gathers noodles. However, an expert user can eat quickly while reserving working memory resources to carry on a discussion (Allport, Antonis, & Reynolds, 1972). These extra working memory resources are available to the expert user because the procedure of how to use chop sticks has already been learned and stored in long-term memory, making that procedure directly accessible without

consuming limited working memory resources. Therefore if one were designing a product with an interface that had similar functionality to chop sticks, one might want to consider whether or not the user already has that knowledge in memory and how that knowledge will affect the user's interpretation of the design.

According to Baddeley (1986), working memory includes both verbal and visual systems. There is a phonological loop for rehearsing verbal information and a visuospatial sketchpad where visual images are briefly stored. Both subsystems are controlled by the central executive. Baddeley's model of working memory has been examined under an array of dual task methodologies in which researchers examine the effects of completing two tasks simultaneously (e.g., Logie, Zucco, & Baddeley, 1990). The two tasks may or may not draw on common limited working memory resources. If a primary task is effortful and consumes working memory resources, the secondary task will only be affected if it consumes similar limited resources.

If affordances and conventions are behaviorally different, they also may be affected differently by availability of working memory resources. Due to the perceptual nature of affordances, they are most likely automatic and should require few working memory resources. A convention may or may not still be a controlled process that requires working memory resources. Thus, there may be design implications if working memory load affects affordances and conventions differently. Experiment 2 was designed to examine whether working memory load interferes with convention-based interactions and not with affordance-based interactions. Answering this question would provide critical insight into cognitive differences between the two interaction types.

The Role of Congruency

As mentioned earlier, consistency plays a critical role in forming stable long-term memory structures that support automatic processing of an interface leading to effortless perception of a design's functions. Such automatic processes, or skills, are often used by experts to quickly find solutions to complex problems. Thus, technology experts develop skills that depend on consistent interface interactions. Like expert chop stick users, technology experts have interaction patterns that are stored in long-term memory. These patterns are directly accessible without consuming many limited working memory resources (Ericsson & Kintsch, 1995). Due to working memory's limited capacity, humans must depend on recognition of patterns to perform multiple tasks more efficiently.

We know from previous research that when expectations are violated, interactions become slower and more effortful. The incongruent mapping of previous interaction knowledge to a current and different interaction is known as negative transfer (Besnard & Cacitti, 2005). When an interface uses familiar visual features, but those features have a new function, it creates a potentially dangerous situation (Norman, 1983). In a negative transfer situation, users expect a learned outcome based on their actions and yet an unexpected outcome occurs. This incongruent mapping between a user's actions and the interface response requires additional working memory resources as the interaction is now required to be consciously monitored. Thus, when working memory resources are limited, dealing with incongruency may be especially difficult. One of the questions examined in Experiment 2 was what is the cost of violating the expectancies in affordance-based and convention-based interactions (as operationalized by incongruent mappings) and how does the cost compare?

In Experiment 2, participants performed a target selection task while maintaining information in working memory. Participants were asked to remember two (low load) or four letters (high load) and the *locations* of those letters within a three by three grid and then to engage in four interaction, or movement, trials in which they "moved" towards the target by hitting one of two buttons. After four movement trials, participants were presented a single letter and asked for its location within the grid. This working memory task required participants to combine information from both verbal and spatial working memory and therefore should also have used resources from the central executive (Baddeley, 2007; Fuster, 2002; Wickens, 2002). The working memory task was designed so that under high load fewer working memory resources would be available for any other processing than under low load.

The button pressing task used Experiment 2 was very different from that used in Experiment 1. In Experiment 2, there was a correct response and the response produced a visible action in which a cursor moved in a prescribed button-to-action mapping, providing visual feedback. Only the vertical and horizontal spatial button configurations were employed and the mappings were either congruent or incongruent (i.e., inverted) given the movement type. For example, an incongruently mapped vertical affordance movement would associate the top button with moving down and the bottom button with moving up. We assumed that the congruent mappings would fit user expectancies and incongruent mappings would violate user expectancies.

Participants were encouraged to learn from their interactions so that they could be as fast and as accurate as possible within each mapping block. There were 40 movement trials in each block and we examined performance as a function of practice in order to determine

whether practice differentially affected responses based on conventions and affordances under manipulations of working memory and button-to-action mapping congruency.

Method

Participants

The same 27 participants from Experiment 1 also participated in Experiment 2.

Stimuli and Apparatus

The equipment and software were the same as used in Experiment 1 except that only two pairs of keys (four buttons) were used. The horizontal configuration consisted of keys one and two covered in blue stickers. The vertical configuration consisted of keys three and six covered in yellow stickers. As in Experiment 1, participants were run at individual computer stations in groups of up to three. They wore sound deadening earmuffs and there were dividers between computer stations preventing the participants from viewing or hearing each other.

Experimental Design

The experiment employed a within-subject factorial design with two button-to-action mappings (congruent or incongruent with expectations), two different working memory loads (two or four letters), and two button configurations (vertical or horizontal). Thus there were eight different blocks. Block order was randomly determined for each participant.

Each block contained 10 working memory trials and embedded within each working memory trial were 4 movement trials in which one target and three distractors were presented. The task was to move to the target's spatial location. The order of movement trials within a working memory trial was random without replacement. Thus, there were 40 movement trials in each block, half reflecting affordance-based movements and half

reflecting convention-based movements. Figure 3 illustrates the displays comprising a

working memory trial with the embedded movement trials.
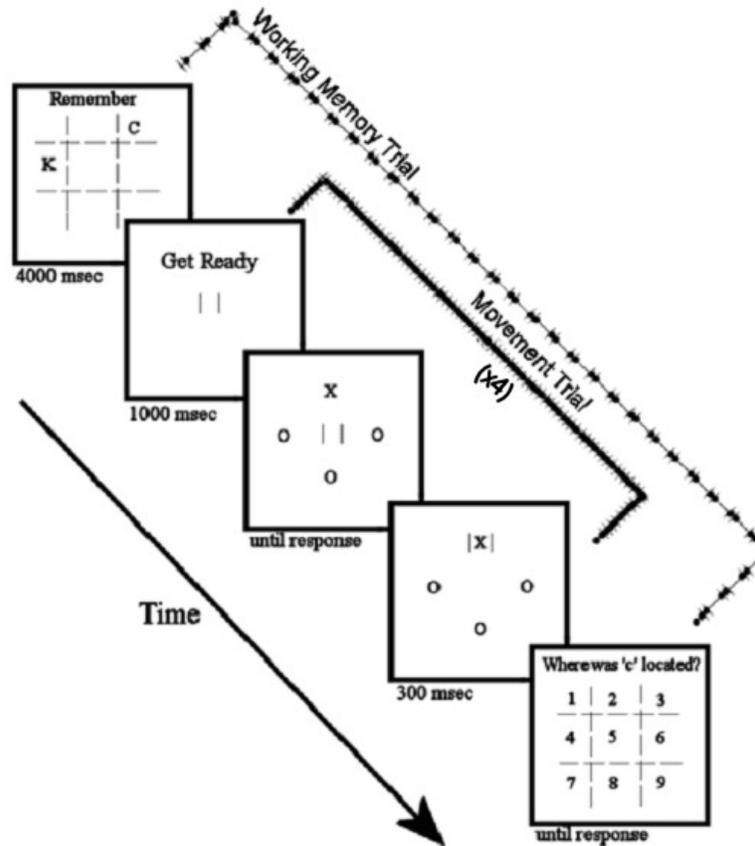
Figure 3. Schematic representation of a working memory trial, which included four embedded movement trials. The arrow represents passage of time. Each trial began with the memory display (4000 msec) in which two or four letters were presented, one at each of 9 possible locations. Next, the display prompted the participant to "Get Ready!!" (1000 msec) then a target and distractor array appeared until the participant responded with a button press that caused the cursor to move to a target or distractor, depending on the mapping in place. The cursor stayed at the associated location for 300 msec, providing visual feedback. Four movement trials occurred, one with the target at each possible location, followed by the letter probe. Participants pressed a letter key indicating the location the letter probe had occupied in the memory display. A break and reminder of the instructions occurred at the end of each working memory trial. Ten consecutive working memory trials under each set of button-to-action mappings comprised a block. After completing a block participants were made aware that a new block was about to begin and they were encouraged to take a break.

The first set of four movement trials in a block served as the "instructional set" because feedback on that block indicated that block's button-to-action mappings. The remaining 9 blocks were divided into thirds to assess practice effects.

Procedure

Participants received general instructions explaining the nature of the working memory task and the movement task. For the working memory task, two or four letters were presented with each letter in a randomly chosen cell in a 3 x 3 grid outlined on the screen. The display lasted 4000 msec. Participants were told that they needed to remember which letter occurred in which locations because their memory for this information would be tested after four movement trials. For the test, one of the letters was randomly selected and was presented at the top of the screen and participants were asked to specify in which position in a three by three grid the letter had appeared at the beginning of the working memory trial. Participants entered the response by using the numeric keys on the upper row of the keyboard. Only accuracy was stressed for the working memory task. Working memory load was manipulated between blocks, so all 10 working memory trials within a block were either load 2 or load 4.

For the movement task, participants were told that they would see a display with a target 'x' and three distractor circles. Their task was to move the cursor in the direction of the 'x' to the best of their ability given the two available buttons. The cursor the participant moved was two black lines vertically arranged. The cursor appeared in the middle of the screen along with the words 'Get Ready' for 1000 msec prior to onset of the target display. Upon seeing the target and distractors array, the participants pressed one of the two buttons and that choice caused the cursor to move in a prescribed way (i.e., the cursor moved to the

location associated with that response for the current block). As soon as the participant made a response, the cursor moved to the location associated with that response for the current block. The cursor remained at the chosen location for 300 msec. As already noted, participants were instructed to be as quick and as accurate as possible on this task.

Participants completed four consecutive movement trials before the working memory test. During these four trials, the location of the target was random with the constraint that it occurred once at each location (Top, Bottom, Left, or Right of the cursor). The button-to-action mappings were either congruent or incongruent with previously identified affordances and conventions/biases. The button-to-action mapping varied across blocks (i.e., they changed after each 10 working memory trials). Participants were not informed of what the button-to-action mapping was, but they were told at the beginning of each block that a new mapping was in place. Participants needed to use the visual feedback relating their button press to cursor movement on the first set of movement trials to infer the current button-to-action mapping. Thus, the first set of four movement trials served as instructions for the block. Table 2 summarizes the four button-to-action mappings that were used.

Table 2. Description of all the button-to-action mappings examined in Experiment 2. The arrows represent the cursor's directional movement when the associated spatial button was pressed.

	Congruent		Incongruent	
Spatial Button Configurations	Affordance	Convention	Affordance	Convention
Vertical	↑ ↓	→ ←	↓ ↑	← →
Horizontal	← →	(Bias) ↓ ↑	→ ←	(Bias) ↑ ↓
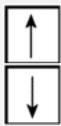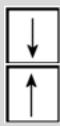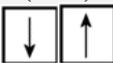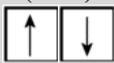
Before each block and after each working memory trial, participants were given an opportunity to rest. They were presented a display in which they were reminded to respond as accurately and as quickly as possible for the next movement trials. The display also instructed them to place their fingers on the appropriate buttons (i.e., vertical or horizontal), ensuring that participants' fingers were on the correct buttons. Importantly, at the beginning of each block, participants were instructed that a new block was about to begin. This warning indicated that the underlying button-to-action mappings or working memory load or both would change.

Across blocks, the button-to-action mappings were manipulated to investigate the impact on user performance of a design that either is congruent with users' current knowledge or that violates their current knowledge (incongruent). Working memory load also was varied across blocks. The response time and accuracy of all responses were collected.

Within a block of trials participants interacted with an affordance and a convention depending on the target's location and the buttons being used. For example, given the vertical button configuration and the four possible locations in which a target may appear, using the buttons to move up/down would be considered an affordance, but moving left/right would be considered a convention. Thus, half of the trials within a block had participants act on an affordance and half on a convention. Moving the cursor up/down and left/right represented different movement types (i.e., affordance or convention) within each block. For each type, the mapping could be congruent (e.g., left button for left) or incongruent (e.g., left button for right). The entire experiment took approximately 30 minutes including instructions and debriefing.

Results & Discussion

Any response with a latency of less than 150 msec was omitted from all analyses (both accuracy and response time); this removed 147 data points out of 8,460 (i.e., 2% of the total data). Further, response time outliers were identified at the individual subject level by computing a ceiling at 2.5 standard deviations above the mean. Separate outlier cutoffs were calculated for the congruent and incongruent distributions because response times were much faster with congruent mappings. Any response time that exceeded the ceiling for a subject was replaced with the ceiling value. These replacements occurred for 202 data points out of 8,460 (i.e., 2% of the total data).

There were three dependent measures across two tasks: working memory accuracy, target selection accuracy, and target selection response time. All statistical tests used an alpha level of .05 and all error bars in the figures represent the standard error of the mean.

Working Memory Task

A 2 x 2 x 2 repeated-measures Analysis of Variance (ANOVA) with the variables button spatial configuration (vertical or horizontal), working memory load (two or four letters), and mapping congruency (congruent or incongruent with expectations) was conducted. As expected, the accuracy on the working memory task for remembering two letters (M = .75, SE = .03) was significantly higher than for remembering four letters (M = .69, SE = .04), $F(1, 26)$ = 5.35, p = .029. Further, performance on the working memory task was well above chance, which was .11, or 1 out 9 possible letter locations. The main effect of mapping congruency was significant, $F(1, 26)$ = 10.53, p = .003, showing that working memory task performance was higher under congruent (M = .77, SE = .03) as compared to incongruent button-to-action mappings (M = .68, SE = .04). No other main effects or

interactions were significant. These results suggest that participants were doing the working memory task and that they should have had more resources available in the low working memory blocks. They also suggest that the incongruent mapping conditions were more difficult than the congruent mapping conditions.

Movement Trials

Four movement trials were embedded within each of 10 working memory trials in a block. The movement trials embedded within the first working memory trial in each block were treated as instruction trials and were not analyzed. Interactions on these trials are qualitatively different from interactions in the other trials because during these trials the participant had to determine by trial and error the nature of the underlying button-to-action mapping. Thus, the feedback provided during the first four movement trials served as button-to-action mapping instructions for a given block. The movement trials within the remaining nine working memory trials were grouped into thirds in order to examine practice effects. The first third was comprised of the 12 movement trials embedded in working memory trials 2-4, the middle third was comprised of the 12 movement trials embedded in working memory trials 5-7, and the final third was comprised of the 12 movement trials embedded in working memory trials 8-10.

Target Selection Accuracy

A 3 x 2 x 2 x 2 x 2 repeated-measures ANOVA with the variables practice (thirds), movement type (affordance or convention), button spatial configuration (vertical or horizontal), working memory load (two or four letters), and mapping congruency (congruent or incongruent with expectations) was conducted. The ANOVA revealed a significant main effect of practice, $F(2,52) = 9.92$, $p < .001$. Target selection accuracy increased as practice

increased (first third: $M = .83$, $SE = .02$; middle third: $M = .85$, $SE = .02$; and last third 4: $M = .87$, $SE = .02$). A linear contrast accounted for 99% of the variance among these means. Practice did not interact with any other experimental variable. So, although performance improved with practice, the small amount of practice within each block was insufficient to produce the level of change within a button-to-action mapping long-term memory structure that would support qualitative differences in performance.

Figure 4 shows mean accuracy, collapsed over practice, as a function of movement type, button spatial configuration, working memory load, and mapping congruency. There was a significant main effect of movement type, $F(1,26) = 42.88$, $p < .001$, showing that accuracy in affordance movements ($M = .89$, $SE = .02$) was higher than in convention movements ($M = .81$, $SE = .02$). There was a main effect of mapping congruency, $F(1,26) = 46.66$, $p < .001$, showing that congruent button-to-action mappings ($M = .92$, $SE = .01$) allowed more accurate responses as compared to incongruent mappings ($M = .78$, $SE = .02$). The Movement type x Button configuration interaction effect was significant, $F(1,26) = 8.64$, $p = .007$, as was the Movement type x Button configuration x Working memory load x Mapping congruency interaction effect, $F(1, 26) = 4.81$, $p = .04$. In order to examine the nature of these interactions, separate analyses were computed for each interaction type.

Affordances

A 2 x 2 x 2 repeated-measures ANOVA on target selection accuracy in the affordance movements with the variables button spatial configuration (vertical or horizontal), working memory load (two or four letters) and mapping congruency (congruent or incongruent with expectations) revealed only a main effect of button-to-action mapping congruency. Accuracy was significantly higher in the affordance conditions with a congruent button-to-action
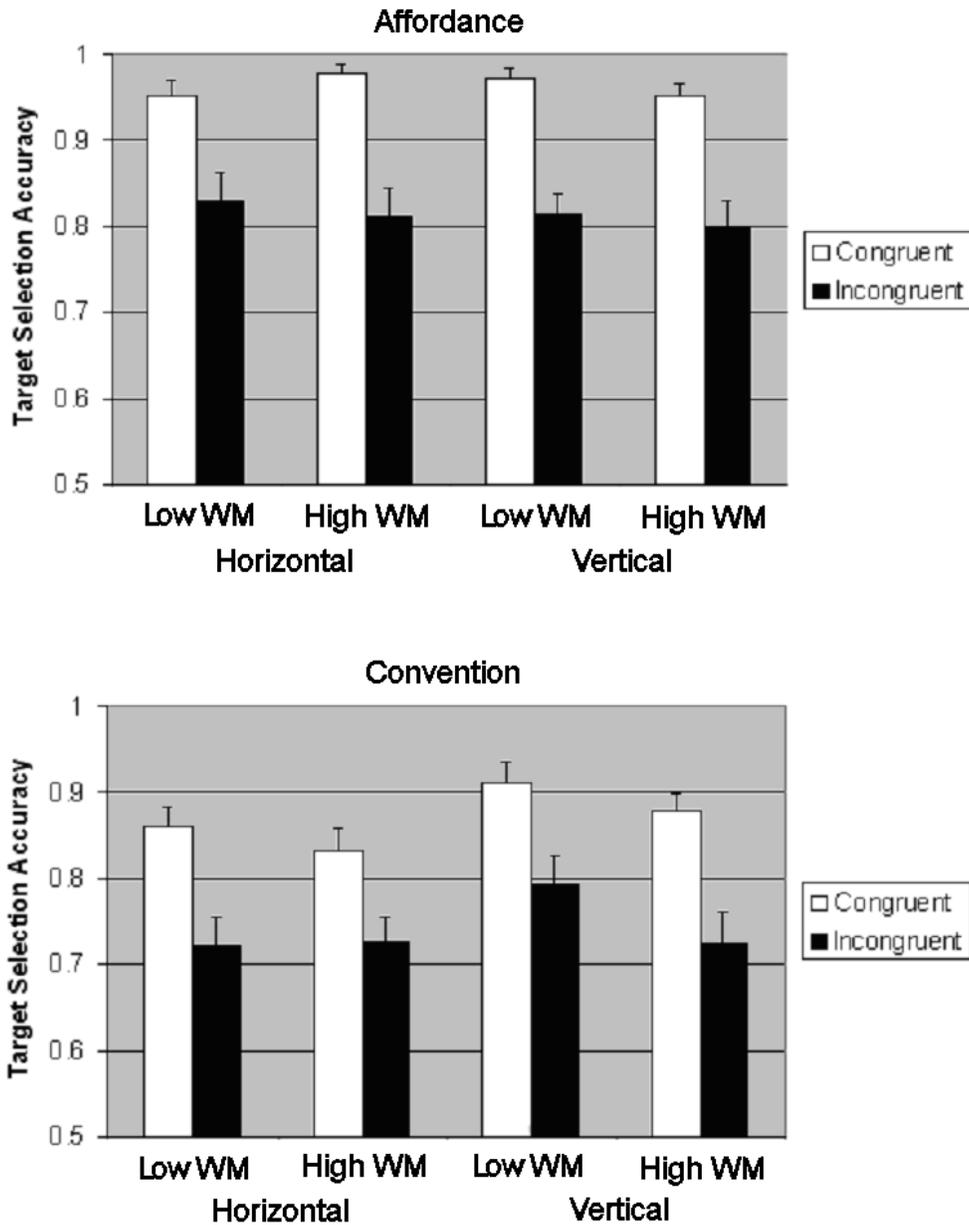
Figure 4. Mean target selection accuracy as a function of movement type (affordance vs. convention), button configuration (horizontal vs. vertical), working memory load (low vs. high), and mapping congruency (congruent vs. incongruent).

mapping ($M = .96$, $SE = .01$) than with an incongruent button-to-action mapping ($M = .81$, $SE = .03$), $F(1,26) = 39.65$, $p < .001$. Target selection accuracy in the congruent affordance movements was very near the ceiling.

Conventions

A similar 2 x 2 x 2 repeated-measures ANOVA on accuracy in the convention movements revealed a more complex pattern. There was a main effect of button configuration in which accuracy was slightly higher with the vertical buttons ($M = .83$; $SE = .02$) than the horizontal buttons ($M = .79$, $SE = .02$), $F(1,26) = 5.33$, $p = .029$. This difference is in line with the results of Experiment 1 suggesting that the vertical convention is stronger than the weaker horizontal bias. Working memory load affected accuracy, $F(1,26) = 5.11$, $p = .03$, with slightly higher accuracy under a low working memory load ($M = .82$, $SE = .02$) compared to a high load ($M = .79$, $SE = .02$). And, as with the affordance movements, mapping congruency affected accuracy, $F(1,26) = 30.32$, $p < .001$. Accuracy was higher with congruent button-to-action mappings ($M = .87$, $SE = .02$) than incongruent button-to-action mappings ($M = .74$, $SE = .03$).

Comparing Affordances with Conventions

Button configuration had a main effect on conventions with higher accuracy on vertical buttons than horizontal buttons, however affordances were unaffected by button configuration. Affordances are simply apparent, so there is no reason to expect that a vertical affordance should differ in strength from a horizontal affordance. As noted however, higher accuracy on conventions with a vertical configuration than a horizontal configuration makes sense if the vertical conventions are stronger than the horizontal biases. Working memory load affected target selection accuracy when participants were acting on conventions,

however the affordance actions were unaffected. These two differences between affordances and conventions show that behavior based on affordances is different from behavior based on conventions. Further, the effect of working memory load on conventions but not biases suggests that conventions are drawing from a limited resource pool, but that affordances are not.

Both affordances and conventions were affected by button-to-action mapping congruency. Inverting the affordances' button-to-action mappings was expected to decrease performance, and it did. The finding that inverting the mappings for conventions also decreases accuracy confirms our interpretation of conventions as being the result of learning based on past interactions. Participants appear to have expectations concerning which button goes with which action, and these expectations affect current actions. Interestingly, the effect of congruency did not appear to be different for affordances and conventions (i.e., the Movement type x Congruency interaction effect was not significant), suggesting that the cost to target selection accuracy of violating an affordance and a convention are roughly comparable. Although this evidence stems from a null result, it is theoretically unexpected. However, it is important to note that accuracy in the congruent affordance movements approached ceiling, which may have prevented detection of an interaction. Future research designed to examine whether congruency effects are similar between affordances and conventions will need to use a more difficult task that is not limited by a ceiling.

Target Selection Response Time

Only response time data associated with a correct target selection were included within these analyses. The data from one participant were omitted because there were too many missing values. A 3 x 2 x 2 x 2 x 2 repeated-measures ANOVA with the variables

practice (thirds), movement type (affordance or convention), button spatial configuration (vertical or horizontal), working memory load (two or four letters) and mapping congruency (congruent or incongruent with expectations) was conducted. All five main effects were significant. There was a significant main effect of practice, $F(2,50) = 10.73$, $p < .001$ in which response time linearly decreased with practice (first third: $M = 755$ msec, $SE = 42.7$; middle third: $M = 705$ msec, $SE = 38.0$; and last third 4: $M = 677$ msec, $SE = 35.0$). A linear contrast accounted for 97% of the variance among the means. Thus, the response times confirm what was found in accuracy: Practice improves performance. Practice did not interact with any other experimental variable.

A significant main effect of movement type, $F(1, 25) = 20.06$, $p < .001$, showed faster responses for affordance-based movements ($M = 673$ msec, $SE = 31.9$) than for convention-based movements ($M = 751$ msec, $SE = 43.9$). A significant main effect of button configuration, $F(1,25) = 4.74$, $p = .039$, showed that movements with the horizontal button configuration ($M = 694$ msec, $SE = 34.7$) were faster than with the vertical button configuration ($M = 730$ msec, $SE = 41.5$). A significant main effect of mapping congruency, $F(1,25) = 18.79$, $p < .001$, showed that movements involving congruent button-to-action mappings ($M = 625$ msec, $SE = 26.03$) were faster than those involving incongruent mappings ($M = 799$ msec, $SE = 54.09$). The congruency effect was as predicted and was also present in the accuracy data.

The Movement type x Button configuration interaction was significant, $F(1,25) = 8.49$, $p = .007$. Mean response times for each of the four cells in the interaction are shown in Figure 5. This two-way interaction was examined by considering the effect of button configuration for each movement type. There was a significant main effect of button

configuration on affordance movements, $F(1,25) = 14.79$, $p = .001$, but not on convention movements, $p > .97$. Affordance-based movements with the horizontal buttons were faster than with the vertical buttons. This effect was not predicted and was not present in the accuracy data.
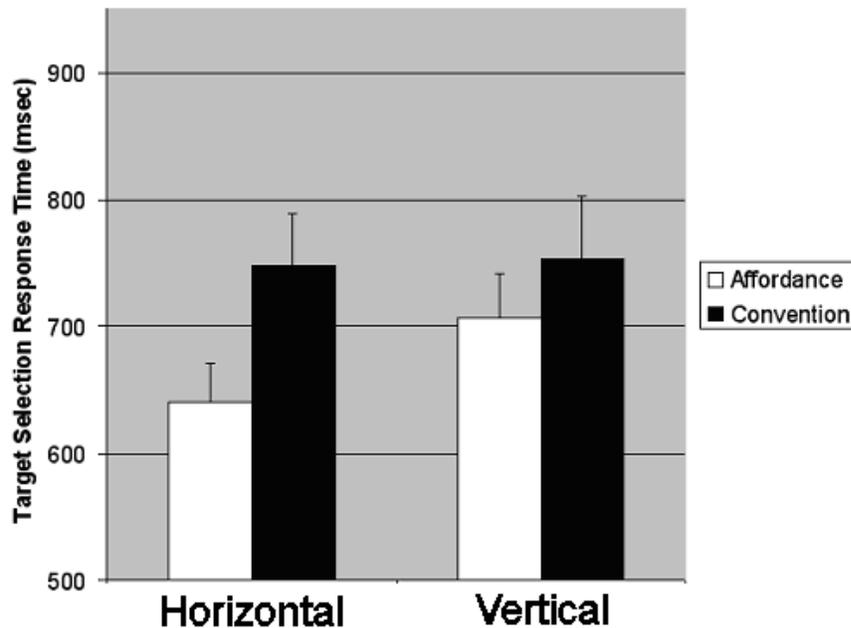

Figure 5. Mean response time as a function of movement type (affordance vs. convention) and button configuration (horizontal vs. vertical).

We offer two possible explanations. First, it is possible that this difference may be reflecting a goodness of fit to a perceptual mapping for moving left/right with the horizontal buttons compared to the vertical buttons. The keyboard was slightly tilted, so the top button was 'up' in relation to the bottom button, however a far/near distinction between the two buttons may have been more apparent. Therefore, a far/near mapping to moving up/down with a vertical button configuration may not be as effective as a left/right mapping to moving left/right with a horizontal spatial button configuration. Second, because the difference was

only found in response time, it might reflect differences in how easily participants were able to maintain their fingers on the appropriate keys. The orientation of the hand for the horizontal buttons was more like the typical orientation for keyboard interaction than the orientation for the vertical buttons.
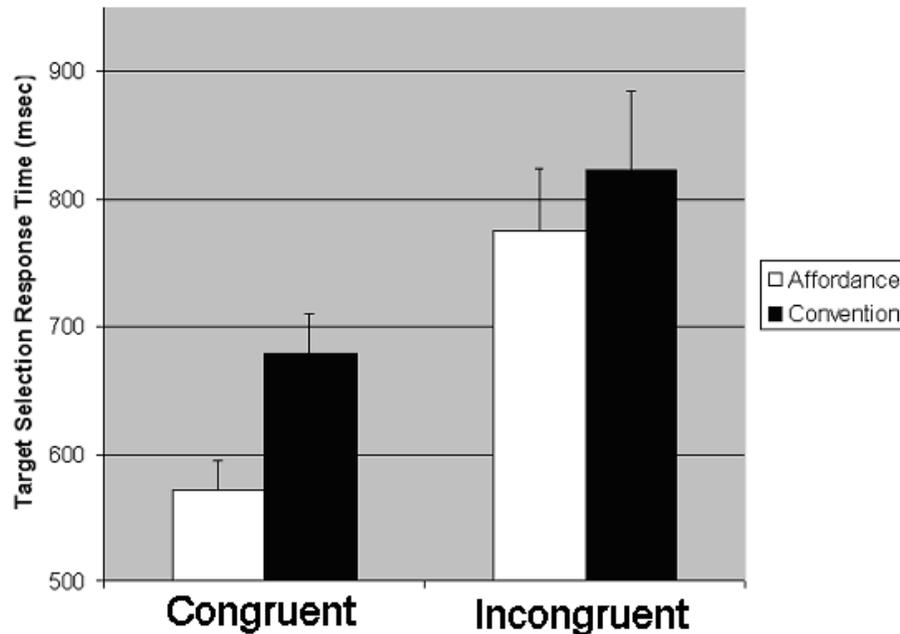

Figure 6. Mean response time as a function of movement type (affordance vs. convention) and mapping congruency (congruent vs. incongruent).

The Movement type x Mapping congruency interaction was significant, $F(1,25) = 5.12$, $p = .03$. Mean response times for each of the four cells in the interaction are shown in Figure 6. This two-way interaction was examined by looking at the effect movement type for each level of mapping congruency. There was a significant main effect of movement type on the congruent mappings, $F(1,26) = 48.75$, $p < .001$, but not on the incongruent mappings. The congruent affordance responses were faster than congruent convention responses, indicating another behavioral difference between affordances and conventions. There was no effect of movement type for the incongruent button-to-action mappings, $p = .078$. Although,

the evidence is provided through a null result, it appears that responding to an incorrectly mapped affordance was just as slow as responding to an incorrectly mapped convention. This provides some empirical evidence that violating learned conventions button-to-action mapping is as detrimental as violating an affordance's perceptual mapping.
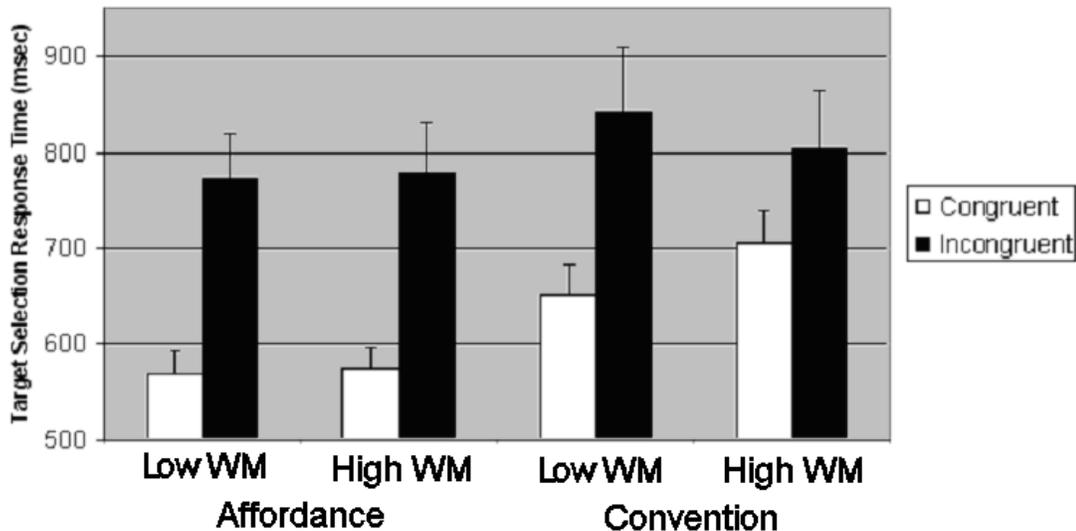

Figure 7. Mean target selection accuracy as a function of movement type (affordance vs. convention), working memory load (low vs. high), and mapping congruency (congruent vs. incongruent).

The Movement type x Working memory load x Mapping congruency interaction was significant, $F(1,25) = 5.80$, $p = .03$. Mean response times for each of the eight cells in the three-way interaction are shown in Figure 7. We examined this three-way interaction by looking at the two-way interactions for each movement type. The Working memory load x Mapping congruency interaction significantly affected convention movements, $F(1,26) = 5.01$, $p = .03$, but not affordance movements, $p > .94$. Examination of these interactions suggests a significant effect of working memory load only for congruent convention

92

responses in which congruent button-to-action mappings were faster under a low working memory load than a high working memory load.

General Discussion

There is debate in the literature about whether a distinction ought to be drawn between affordances and cultural conventions from the designers' viewpoint (McGrenere & Ho, 2000; Norman, 1999; Oshlyansky, Thimbleby & Cairns, 2004). But, little empirical data exist to experimentally contrast these interaction design concepts (Still & Dark, 2008). The purpose of the two experiments reported in this article was to explore the cognitive attributes of design conventions and affordances. Experiment 1 showed that participants had prior expectations about convention-based button-to-action mappings. The button-to-action mappings were the same as we found previously (Still & Dark, 2008). Experiment 2 examined the effect of practice, working memory load and button-to-action mapping congruency on affordances and conventions within the context of a spatial target selection task. Experiment 2 demonstrated some behavioral differences between acting on affordances and conventions. Working memory load did not affect button interactions involving affordances, but did affect interactions involving conventions. As expected, both affordances and conventions were heavily affected by violations of button-to-action mapping expectancies.

Manipulating the availability of working memory resources differentially affected responses based on affordances versus conventions. Working memory load had no affect on affordance-based interactions, however it affected convention-based interactions in that accuracy was lower when working memory load was higher. These results suggest that interactions involving perceptual affordances use very little resources, while those involving

conventions require more. That is, acting on a convention is more accurate and faster under low working memory load compared to a high load. An inference that can be drawn from these data is that interfaces used by operators under a heavy working memory load should be designed to take advantage of perceptual affordances.

Manipulating button-to-action mapping congruency had a large effect on participants' responses. Congruent mappings were always faster and more accurate compared to incongruent mappings. Congruent mappings for affordances were associated with faster and more accurate responses as compared to conventions. This finding provides more evidence that affordances and conventions have differential impact on behavior. Additionally, when considering only the nonaffordance conditions, congruent convention (i.e., horizontal) mappings allowed faster responses than congruent bias (i.e., vertical) mappings.

In contrast to congruent mappings, response times based on incongruent affordance or convention button-to-action mappings were statistically indistinguishable. Acting on a reversed button-to-action mapping produced a similar cost to the user regardless of whether the mapping reflected an affordance or convention. This finding suggests that the convention-based button-to-action mapping is a fundamental part of how the user views the world. The conventions produce expectancies. Violating an expectancy is detrimental regardless of whether the expectancy reflects the natural world or learned constraints.

The findings from Experiment 1 captured the participants' button-to-action mapping expectations. We replicated the basic pattern of our previous findings (Still & Dark, 2008), but with many fewer trials. The results demonstrate that participants were not randomly pressing buttons when perceptual affordances were not available. Rather, they acted in a consistent manner that clearly demonstrated the existence of learned button-to-action

associations. The somewhat lower agreement among participants about the horizontal nonaffordance mappings (biases) than the vertical nonaffordance mappings (conventions) indicates that these learned associations differ in strength.

We examined short term practice effects in Experiment 2. We expected that differences in participants' ability to overcome incongruent button-to-action mapping associated with affordances and conventions would be informative. We fully expected that incongruent affordance mappings would be more disruptive than incongruent convention mappings. We found that target selection accuracy and response times improved in a linear fashion with practice (thirds) within each block. However, practice did not interact with any experimental variable. There simply were too few practice trials to observe differences in the rate at which participants were able to overcome their button-to-action expectations.

Implications for Designers

Currently, designers adhere to interaction affordances and avoid violating what have been considered naturally occurring mappings. Part of the reason affordances are important is because they *naturally* communicate the design's available actions (Norman, 1988). However, one can ask whether all natural interactions stem from innate perceptual information (Gibson, 1979) or whether some natural interactions might stem from learned cultural conventions (Raskin, 1994). We believe that both perceptual information and learned associations affect a user's expectations of what a design ought to offer in response to an action. On the one hand, the results of this experiment show that perceptual affordances, when mapped correctly, outperform conventions. On the other hand, violating the user's expectations of a button-to-action mapping produced a similar cost whether the violation was of an affordance or a convention.

There is one distinction between conventions and affordances that is not controversial; that distinction is that conventions must be learned and physical affordances require minimal learning. Users are continually learning about and reacting within their real and virtual operating environments. Each interaction produces a change in long-term memory and long-term memory influences the current representation of the situation. Thus, continuous learning is cognitively reflected as a constant interaction between long-term and working memory. These memory interactions allow users to effectively operate within an overwhelmingly rich and complex world. Working memory operations affect structures in long-term memory (i.e., learning occurs) and those long-term memory structures directly affect expectations and facilitate interactions. As we have previously suggested (Still & Dark, 2008), participants may have learned button-to-action mappings from interaction with remote controls, vehicle turn signals and computer monitor buttons just to name a few common input devices that support binary responses. The learned mappings lead to changes in long-term memory structures that facilitate real time processing. So, when a similar instance of interaction arises again, relevant structures in long-term memory become active, which leads to effortless and efficient processing of the situation and the response (Baddeley, 2002; Cowan, 1988).

We associate this effortless interaction experience with the directness of an object's function. Further, the degree of effort required for processing relates to the strength of the representation that supports the interaction. A perceptual affordance's representation is very strong; it reflects human evolutionary history. A learned convention has an initially weaker representation, but over many interaction experiences the representation grows stronger. We predict that through enough practice, a convention's representation would grow to be as

strong as a perceptual affordance's representation. Given sufficient practice (not provided in these two experiments) with a convention-based interaction, it should no longer consume limited working memory resources. By that logic, a strong distinction between affordances and heavily learned conventions should not exist within the design literature as both provide direct access to an expected response.

Designers should attempt to reuse existing interaction conventions. The consistency guideline facilitates the transfer of arbitrary design associations into pattern knowledge making conventional actions automatic. As already noted, designs with interaction regularity facilitate the transformation of conscious interactions that consume working memory processing into effortless interactions that are automatic. Designers should not *just* focus on maintaining consistency to facilitate learning (e.g., Apple Computer, Inc., 1992; Nielsen, 2002; Shneiderman & Plaisant, 2005). They also ought to try to maintain already learned pattern knowledge by maintaining their software's design conventions across product families and updates.

When an interaction is consistent, it allows users to draw upon their past experiences with technology to make predictions about how a system is going to behave. According to Hannon (2008), "…our ability to learn and use new technologies is contingent upon our experience with prior technologies. On a computer, for instance, each time we learn a new interaction idiom such as drag-and-drop, or double clicking, or scrolling, we adopt new ways of understanding how software applications and hardware devices work" (p.60). It is apparent that following previous conventions and including affordances within an interface ought to facilitate the usability of product.

Reusing a design convention requires being able to identify its existence. One difference between an affordance and a convention is the potential visibility of its existence to designers. Affordances should be common across people. We have suggested that many users hold the same conventions, so some might be tempted to simply test new interactions on themselves in order to determine whether a convention exists in their new design. The problem is that designers may not be able to perceive their users' conventions. The perception of a convention depends on an individual's previous interaction experience with an interface. Thus, when designers do not have a similar interaction history as their users, the designers might easily overlook a conventional use. It is often the case that a designer's perception of an interface and motivation towards its use is much different from the users. The difference is explained by the designer's unique experience with the design, which includes the weeks or months spent thinking about and interacting with the new interface. This additional experience with a design may make it very difficult to identify previous conventional usages and in fact may lead the designer to see the new interface as being natural. As a result, a designer may accidentally violate a convention and thus require the users to overcome conventional usage, which requires them to engage in controlled processing and expend additional cognitive effort.

Previously, we suggested that affordances and conventions are not behaviorally distinguishable (Still & Dark, 2008). That conclusion was based on a task in which there was no correct answer and no feedback or working memory load. In the current research, when we added those dimensions in Experiment 2, we demonstrated that behavior based on naturally occurring affordances was differentially affected compared to behavior based on learned conventions. Thus, affordances and conventions can differentially affect behavior.

We also demonstrated that violating a convention produces a performance decrement, just as it does for affordances. Thus, affordances and conventions can similarly affect behavior.

An important takeaway message from our research is that conventions once learned are not arbitrary. A convention may appear unconstrained to a designer making an interface update, but that is not the case. Although conventions initially may have been arbitrary, learning does occur. With sufficient practice, new structures are formed in long-term memory, which leads to appropriate actions being apparent to the user.

In conclusion, our data provide evidence for both sides of the convention and affordance debate. It seems that Norman (1999) was correct in suggesting that there are actual behavioral differences between affordances and conventions and McGrenere and Ho (2000) were right to suggest that learned conventions do structure user responses toward expected actions. We propose that a synthesis of these two viewpoints will best serve designers. Designers should employ perceptual affordances when possible. When perceptual affordances are not available, designers should identify and reuse established conventions. Designers should avoid violating conventions, unless the benefit of the new design outweighs performance costs.

References

Allport, D. A., Antonis, B., & Reynolds, P. (1972). On the division of attention: A disproof of the single-channel hypothesis. *Quarterly Journal of Experimental Psychology*, *24*, 225–235.

Apple Computer, Inc. (1992). *Macintosh human interface guidelines*. Reading, MA: Addison–Wesley.

Baddeley, A. D. (2007). Visuospatial short-term memory. In M. D'Esposito, J. Driver, T.

    Robbins, D. Schacter, A. Treisman & L. Weiskrantz (Eds.), Working Memory,

    thought, and action (pp. 63 – 83). New York: Oxford Press.

Baddeley, A. D. (2002). Is working memory still working. *European Psychologist*, *7(2)*, 85-

    97.

Baddeley, A. D. (1992). Working memory. *Science*, *255*, 556-559.

Baddeley, A. D. (1986). *Working memory*. Oxford, England: Oxford University Press.

Besnard, D., & Cacitti L. (2005). Interface changes causing accidents. An empirical study of

    negative transfer. *International Journal of Human-Computer Studies, 62(1),* 105-125.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their

    mutual constraints within the human information-processing system. *Psychological*

    Bulletin, *104*, 163-191.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review,*

    102, 211-245.

Fuster, J. M. (2002). Physiology of executive functions: The perception-action cycle. In D.

    Stuss & R. Knight (Eds.), Principles of frontal lobe function (pp. 96 – 108). New

    York: Oxford Press.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Houghton Mifflin:

    Boston.

Hannon, C. (2008, September). Mental and conceptual models, and the problem of

    contingency. *Interactions*, 58-64.

Logie, R. H., Zucco, G., & Baddeley, A. D. (1990). Interference with visual short-term

    memory. *Acta Psychologica*, *75*, 55-74.

McGrenere, J., Ho, W. (2000). Affordances: Clarifying and evolving a concept. *Proceedings of Graphics Interface*, 1-8.

Moors, A., & Houwer, J. D. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin, 132(2),* 297-326.

Nielsen, J. (Ed.), (2002). *Coordinating user interface for consistency*. Morgan Kaufmann: New York.

Norman, D. A. (1983). Design rules based on analyses of human error. *Communications of the ACM*, *26*, 254-258.

Norman, D. A. (1988). *The psychology of everyday things*. Basic Books: New York.

Norman, D. A. (1999, May). Affordance, conventions, and design. *Interactions*, 38-42.

Norman, D. A. and Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, and D. Shapiro (Eds.), Consciousness and self-regulation: Advances in research and theory, vol. 4 (pp.1-18). New York: Plenum Press.

Oshlyansky, L., Thimbleby, H., & Cairns, P. (2004). Breaking affordance: Culture as context. *Proceedings of the third Nordic conference on Human-computer interaction*, 81-84.

Raskin, J. (1994, September). Intuitive equals familiar. *Communications of the ACM, 37*(9), 17.

Reason, J. (1990). *Human error*. New York: Cambridge University Press.

Shiffrin, R. N., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic, attending, and a general theory. *Psychological Review*, *84*, 127-190.

Shneiderman, B., & Plaisant, C. (2005). *Designing the User Interface (Fourth Edition): Strategies for effective human-computer interaction*. Pearson Addison Wesley.

Still, J. D. & Dark, V. J. (2008). An empirical investigation of affordances and conventions. In J. S. Gero and A. K. Goel (Eds.), Design Computing and Cognition '08, (pp.457-472). Netherlands: Springer.

Visser, W. (2006). *The cognitive artifacts of designing*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, *3*(2), 159-177.

CHAPTER 5. GENERAL CONCLUSIONS

My dissertation investigated the cognitive basis of affordances and conventions with the intent to provide a means by which designers could better predict the success of user interactions. Examination of similarities between affordances and conventions provides human-centered insight into how a user will perceive, process, and respond to design elements. Affordances are often highlighted for their properties of presenting possible actions to users that are easy to understand, thus making affordance an important usability concept. Determining the extent to which this is true of conventions also is important. My hope is that this research will allow future designers to create devices that make our lives easier.

Traditionally, affordances were discussed within a Gibsonian framework where affordances arise from direct perception (Gibson, 1979). However, some authors describe affordances as being mostly perceptual while others describe them as being culturally bound (McGrenere & Ho, 2000; Norman, 1999; Oshlyansky, Thimbleby & Cairns, 2004). We suggest that both of these descriptions are correct, and that they can be explained from a cognitive conceptualization of *perceived* affordances. The literature has yet to describe the cognitive characteristics of affordances or how a perceived affordance can be created. This dissertation introduces a framework that explicitly describes the process from which perceived affordances arise and the cognitive attributes associated with such an outcome. Within the framework described, perceived affordances result from automatic processing of information, that is, they result from an interaction between long-term memory structures and current information in the environment. As such, perceived affordances can reflect many different types of constraint; they are not limited to just physical constraints. Automatic processing is defined as occurring unconsciously and quickly without any effort or need for

working memory resources. These attributes are also used to describe affordances. Though a perceived affordance is not exactly what Gibson meant by affordance, it does have relevance and applicability for designers. A number of authors have mapped affordances onto a number of different frameworks trying to achieve more applicability (see Zhang & Patel, 2006; Baerentsen & Trettvik, 2002; Vyas, Chisalita & Veer, 2006); the view laid out in my dissertation unifies these different frameworks into one appropriate for design applications.

The cognitive view of perceived affordance clarifies the aspects of an interaction that are important for design. It does not focus on a particular type of constraint, but rather on the possibility that a constraint may be in play. A constraint may be the result of a history of consistent interactions, which leads to the formation of long-term memory structures that directly affect working memory representations without any intention or effort on the part of the user.

Affordances emerge from constraints formed through interactions between an external object and the user's personal characteristics (Gaver, 1991). In Gaver's view, affordances are the result of an interaction between the external object provided physical constraints and the user's perception. The current view extends this to include learned constraints and thus highlights the need for consistency in design. Design consistency is critical for producing effortless usage, because interaction regularity facilitates the formation of long-term memory structures that underlie perceived affordances. Current representations are molded by long-term memory structures and are activated by ongoing task needs, which drive the meaning of the specific affordance in play.

There is a debate in the literature concerning whether a distinction between affordances and cultural conventions ought to be drawn (McGrenere & Ho, 2000; Norman,

1999; Oshlyansky et al., 2004). It is possible that in the absence of affordances, users develop conventions to resolve interaction ambiguity. In this dissertation, differences between affordances and conventions were explored through a button pressing task. The results show that affordances exist when the spatial button configuration is congruent with directional cues. When affordances were not available, most participants demonstrated consistent button-to-action mappings that define conventions.

The results reported in Chapter 3 showed that affordances and conventions are not behaviorally distinguishable when participants perform button presses without feedback or working memory load (Still & Dark, 2008). However, the results reported in Chapter 4 demonstrated that in comparison to interactions based on naturally occurring affordances, those based on conventions are more affected by manipulations of working memory and button configuration. Interestingly, violating the user's expectations of a button-to-action mapping produced a similar cost whether the violation was of an affordance or a convention. It is my hope that the findings will impress upon designers that violating conventions should be avoided, just as is violating affordances, unless the benefit of the new design outweighs performance costs.

The results reported provide evidence for both sides of the convention and affordance debate. It seems that Norman (1999) was correct in suggesting that there are actual behavioral differences between affordances and conventions and McGrenere and Ho (2000) were right to suggest that learned conventions do structure user responses toward expected actions. However, as implied by the statement that they both are supported, neither by itself is fully supported. Norman suggested that conventions are qualitatively different than affordances, but the data suggest that the difference may be more qualitative. The cost of

violating interaction expectations were similar for both affordances and conventions.
McGrenere and Ho suggested that a distinction between conventions and affordances should
not be made, but working memory load did differentially influence responses for the two
interaction types. The most likely explanation, in my opinion, is that after the initial learning
period, conventions play a critical role in the perception of a design's available actions just as
do perceptual affordances. Based on the information, both theoretical and empirical,
presented in this dissertation, I suggest that a synthesis of these two viewpoints will best
serve designers.

My future research will further investigate whether practical differences exist
between acting on a convention (e.g., clicking a hyperlink) and acting on a physical
affordances (e.g., pushing a remote control button). Additionally, I will explore different
methods of capturing the conventional interactions that may be driving user expectations.
Designers need appropriate methods for capturing those aspects of a user's long-term
memory structures that support the perception of conventions.

In conclusion, knowing how the user might cognitively view an interface is necessary
for the future development of devices that optimize successful system interactions. The
ultimate goal of my research is to help facilitate the creation of technologies that effortlessly
support the user's task goals.

The Designer's Summary

If I were to do an executive summary of the three most important points made in this
dissertation for designers, they would be as follows:

First, it is important to remember that interaction experiences are directly molded by
long-term memory structures. That is, users are continuously learning from everyday

interactions and these learning instances directly affect future interactions expectations. Conventions contain arbitrary mappings, but with sufficient practice even arbitrary action to function mappings may grow to be strongly associated and become automatically processed. Such strong associations between actions and functions produce perceived affordances within interface designs.

Second, when making design decisions about whether to employ a convention or perceived affordance within a design, it is important to consider the cognitive attributes of the user. Designers already take into account the limitations of working memory. They also need to take into account the impact of long-term memory structures that support conventions and perceived affordances.

Third, the actions of conventions and other perceived affordances are quickly apparent without any effort on the part of the user. For this reason, they are difficult to overcome with only limited amounts of practice, just as are physical affordances. Violating an affordance or a convention produces a similar cost on performance. Therefore, violating a convention should be avoided in design just as much as current designs avoid violating physical affordances. Designers ought to maintain consistency across interactions whether the affordance is learned or innately available.

In sum, the empirical results from the experiments described in this dissertation provide designers with a couple of interface design guidelines. First, designers need to employ perceptual affordances when possible, especially when the user may be under a high working memory load. Second, when perceptual affordances cannot be implemented, established interface conventions should be identified and reused. This is because incorrectly

mapping an affordance's or a convention's action to function will lead to products that are more difficult to use.

References

Baerentsen, K. B., & Trettvik, J. (2002). An activity theory approach to affordance. *NordiCHI*. New York: ACM, 51-60.

Gaver, W. W. (1991). Technology affordances. *Computer-Human Interaction Conference Proceedings*, 79-83.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Houghton Mifflin: Boston.

McGrenere, J., & Ho, W. (2000). Affordances: Clarifying and evolving a concept. *Proceedings of Graphics Interface*, 1-8.

Norman, D. A. (1999, May). Affordance, conventions, and design. *Interactions*, 38-42.

Oshlyansky, L., Thimbleby, H. & Cairns, P. (2004). Breaking affordance: Culture as context. *Proceedings of the third Nordic conference on Human-computer interaction*, 81-84.

Still, J. D. & Dark, V. J. (2008). An empirical investigation of affordances and conventions. In J. S. Gero and A. K. Goel (Eds.), Design Computing and Cognition '08, (pp.457-472). Netherlands: Springer.

Vyas, D., Chisalita, C. M. & Veer, G. C. (2006). Affordance in interaction. *ACM Proceeding of the 13th European conference on Cognitive ergonomics: trust and control in complex socio-technical systems*, *250*, 92-99.

Zhang, J., Patel, V. L. (2006). Distributed cognition, representation, and affordance. *Cognition & Pragmatics, 14,* 333-341.

ACKNOWLEDGMENTS

I must thank my wife Mary L. Still, for her loving support and encouragement throughout the completion of my dissertation. Further, she has continually provided intellectual insight and support throughout my research career.

I would like to thank my advisor Veronica J. Dark, for her enduring patience through my continued academic development. Veronica has an uncanny ability to remember and pay attention to details. I believe this ability makes her a brilliant professor. Over the years, she has shared an immeasurable amount of knowledge about cognitive psychology, statistics, technical writing, experimental design and professional issues, for which I will always be grateful. I have very much valued meeting weekly with Veronica across the last five years. I will miss the intellectual exchange.

Finally, my committee members are acknowledged for their valuable contributions to this work: Ana Correia, Stephen Gilbert, Alison L. Morris and William S. Robinson.