

# An Empirical Investigation of Affordances and Conventions

**Jeremiah D. Still and Veronica J. Dark**

*The Human Computer Interaction Program and Department of Psychology at Iowa State University, USA*

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 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 invocable 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 shortcuts).

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.

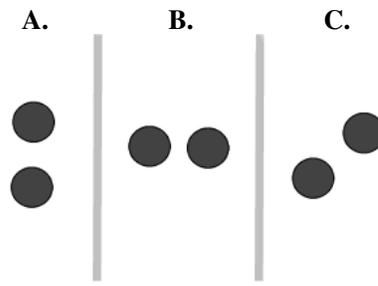
The light switch study assessed participant *knowledge* without providing feedback. The results suggested the existence of a different convention in the US and the UK regarding how to interact with a light switch. Our investigation focuses on how users *act* in a very simple situation that sometimes 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.



**Fig. 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](http://www.pstnet.com)).



**Fig. 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.

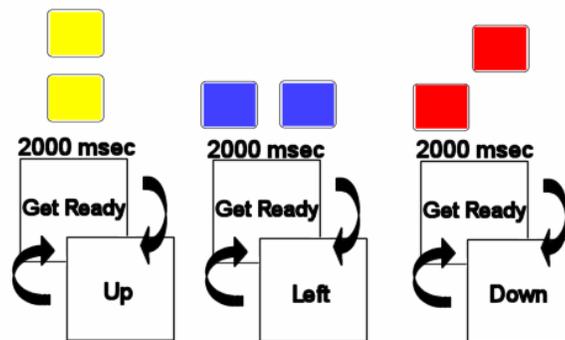
### Procedure

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.



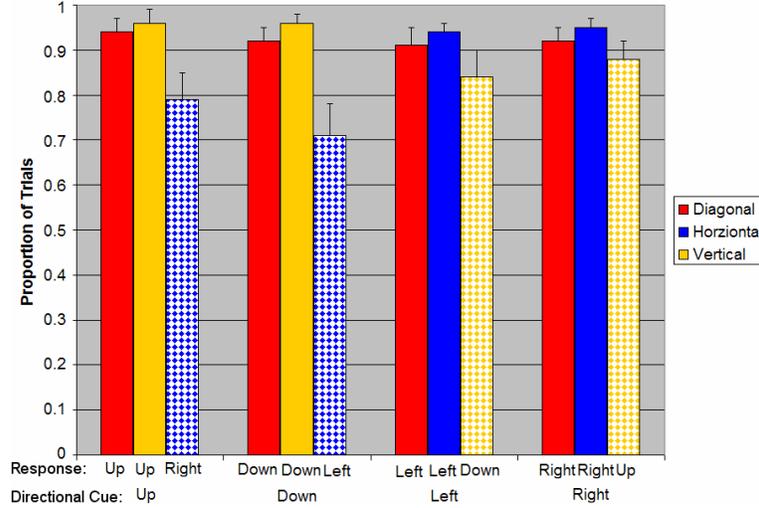
**Fig. 3** Schematic representation of the experimental procedure. The arrows indicate the change of displays. The three colors and button configurations represent the three blocks of trials. For each button configuration, a participant was presented 20 trials, each of which began with 'Get Ready' for 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.



**Fig. 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.

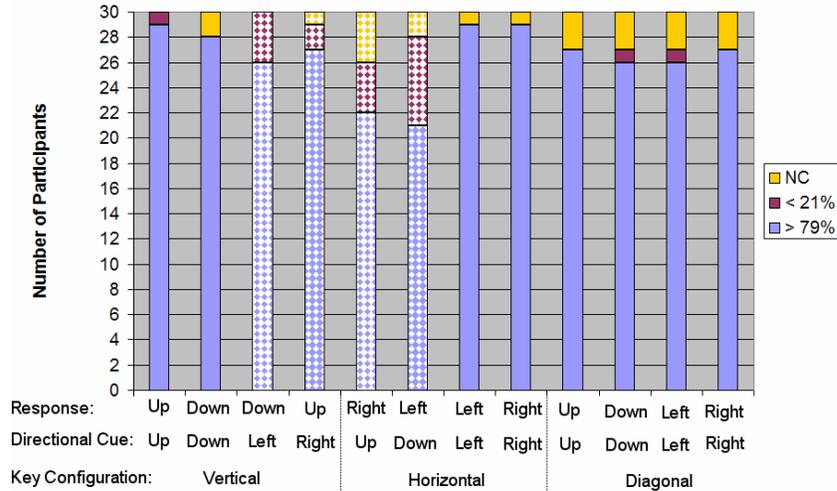
The first thing to notice is that the proportion of what we had assumed were afforded responses (solid bars) is above 0.90 for all button configurations, suggesting that we were correct in our assumption that affordances were present. The second thing to notice is that although the proportions in the 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.



**Fig. 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.

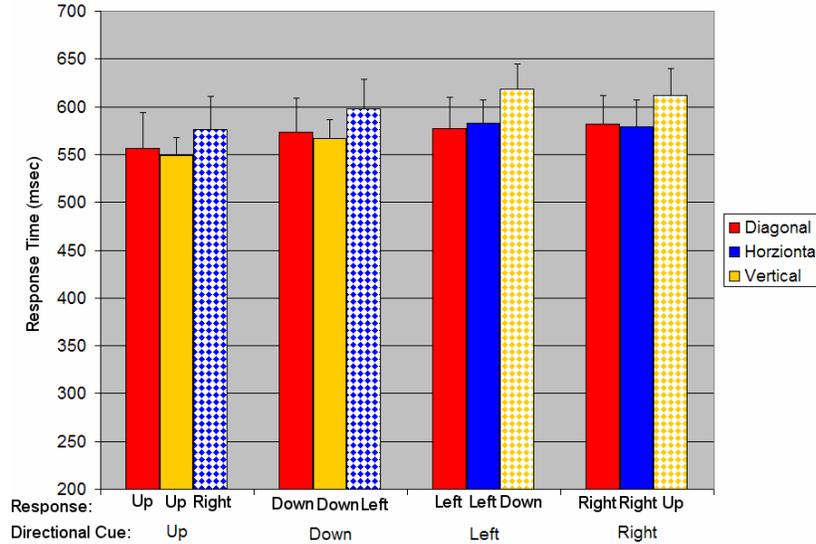
The data presented in Figure 5 confirm the interpretation offered for the response choice data. Affordances exist in the spatial configuration of the buttons. In an ambiguous situation in which individuals are told to make a response indicating a movement, a button that matches the direction of the movement 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 dif-

ference 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.



**Fig. 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.



**Fig. 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 bat-

tle 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**

We would like to thank Matthew Fee for running participants and Mary L. Still for suggestions on earlier drafts.

The original publication is available at [www.springerlink.com](http://www.springerlink.com)

### **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 13<sup>th</sup> 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