Applying the Norman 1986 User-Centered Model to Post-WIMP UIs: Theoretical Predictions and Empirical Outcomes

G. MICHAEL POOR, Baylor University
SAMUEL D. JAFFEE and LAURA MARIE LEVENTHAL, Bowling Green State University
JORDAN RINGENBERG, The University of Findlay
DALE S. KLOPFER and GUY ZIMMERMAN, Bowling Green State University
BRANDI A. KLEIN, Missouri University of Science and Technology

30

In recent decades, "post-WIMP" interactions have revolutionized user interfaces (UIs) and led to improved user experiences. However, accounts of post-WIMP UIs typically do not provide theoretical explanations of why these UIs lead to superior performance. In this article, we use Norman's 1986 model of interaction to describe how post-WIMP UIs enhance users' mental representations of UI and task. In addition, we present an empirical study of three UIs; in the study, participants completed a standard three-dimensional object manipulation task. We found that the post-WIMP UI condition led to enhancements of mental representation of UI and task. We conclude that the Norman model is a good theoretical framework to study post-WIMP UIs. In addition, by studying post-WIMP UIs in the context of the Norman model, we conclude that mental representation of task may be influenced by the interaction itself; this supposition is an extension of the original Norman model.

Categories and Subject Descriptors: H.5.2 [**User Interfaces**]: Interaction Styles

General Terms: Human-Computer Interaction

Additional Key Words and Phrases: Reality-based interaction, direct manipulation interaction, post-WIMP interaction

ACM Reference Format:
G. Michael Poor, Samuel D. Jaffee, Laura Marie Leventhal, Jordan Ringenberg, Dale S. Klopfer, Guy Zimmerman, and Brandi A. Klein. 2016. Applying the Norman 1986 user-centered model to post-WIMP UIs: Theoretical predictions and empirical outcomes. ACM Trans. Comput.-Hum. Interact. 23, 5, Article 30 (October 2016), 33 pages.
DOI: http://dx.doi.org/10.1145/2983531

Some aspects of this work were supported by a Bowling Green State University Faculty Research Committee Research Initiative Grant awarded to Leventhal and Klopfer entitled "Using Interactive Three-Dimensional Models to Study Spatial Ability." Some software was developed with support from Summer 2012 (NSF Science, Technology, Engineering, and Mathematics Talent Expansion Program (STEP) under Award No. 0757001), and S-STEM (under Award No. DUE-0850026).
Authors' addresses: G. M. Poor, Computer Science Department, Baylor University, Waco, TX, 76798; email: Michael_Poor@baylor.edu; S. D. Jaffee, Psychology Department, Bowling Green State University, Bowling Green, OH, 43403; email: jaffees@bgsu.edu; L. M. Leventhal, Computer Science Department, Bowling Green State University, Bowling Green, OH, 43403; email: leventha@bgsu.edu; J. Ringenberg, Computer Science Department, The University of Findlay, Findlay, OH, 45840; email: ringenberg@findlay.edu; D. S. Klopfer, Psychology Department, Bowling Green State University, Bowling Green, OH, 43403; email: klopfer@bgsu.edu; G. Zimmerman, Computer Science Department, Bowling Green State University, Bowling Green, OH, 43403; email: gzimmer@bgsu.edu; B. A. Klein, Psychology Department, Missouri University of Science and Technology, Rolla, MO, 65409; email: kleinb@mst.edu.

DOI: http://dx.doi.org/10.1145/2983531

1. INTRODUCTION

Human-Computer Interaction (HCI) as a discipline stands at a crossroads between computer science and a number of social science disciplines, including cognitive psychology. As such, HCI offers numerous opportunities to apply cognitive theories to explain user behavior with specific types of user interfaces (UI) and conversely for behaviors to inform theory. Those opportunities to use and to inform theory are particularly rich when interaction styles change dramatically. For example, in the 1980s when the direct manipulation and the Window Icon Menu Pointer (WIMP) paradigms were revolutionary [Shneiderman 1982], theoretical models such as Norman [1986] were used to explain enhanced user performance with the new UIs. Simultaneously, observations of behavior with the new UIs informed and enhanced theoretical models, such as Norman [1986].

Over the last 20 years, the introduction of "post-WIMP" UIs have presented new opportunities to both use and build theory. Post-WIMP interactions provide enhanced user experiences and support improved user performance. Cognitively, post-WIMP UIs allow users to engage parallel perceptual systems and sophisticated knowledge structures, as well as to use rich problem-specific information processing systems [Jacob et al. 2008].

In this article, we apply the Norman [1986] model of interaction, a classic theory that has been applied for decades to WIMP interactions, to post-WIMP UIs. As we apply this theory, we also explore ways the theory itself may be informed and updated. Our specific strategy for the article is as follows:

(1) *Overview Norman [1986]'s model of interaction*: For simplicity, we call this model *the Norman model*. The Norman model drew together the work of many and was based on cognitive theories of problem solving that emphasized the roles of users' mental representations in UI outcomes.
(2) *Apply the Norman model to post-WIMP interactions*: We examine the two mental representations that are highlighted in the Norman model and how they might be influenced by post-WIMP interactions. Specifically, we hypothesize that post-WIMP UIs (1) enhance the mental representation of UI and (2) enhance the mental representation of task. We explore these theoretical assertions by reviewing a number of theoretical and empirical studies that investigated mental representations of UI and task.
(3) *Empirically test our theoretical assertions*: Following our review of background literature, we present a study of three UIs. The UIs lie along a continuum, from a traditional WIMP UI to a clearly post-WIMP UI. In the study, participants completed a standard three-dimensional object manipulation task, a type of task that is typical of the targeted designs of many post-WIMP UIs. The study tested the two hypotheses that post-WIMP interactions enhanced mental representations of UI and task. We discuss our empirical results in the context of the Norman model and our theoretical assertions.
(4) *Further empirically explore our theoretical assertions*: We conduct further exploratory analyses to answer unanswered questions *vis-a-vis* our assertions of the relationships between mental representations and post-WIMP interactions.
(5) *Discuss results and make conclusions*: We conclude the article with a discussion of our results, their generalizability and applicability to theory.

Following the strategy outlined above, we hope to provide a better understanding of post-WIMP UIs and contribute to relevant theory in HCI.

Table I. Norman's 1986 Model of User Interaction

Stage 1.	The user forms a task goal. In other models, this activity is characterized as constructing a mental representation of the problem space (see Newell and Simon [1972]).
Stage 2.	The user forms a task intention within the task goal. Mentally, in forming the task intention, the user determines how to achieve the task goal. The particular task intention is influenced by the richness of the user's mental representation of the task.
Stage 3.	Mentally, the user translates the desired task intention to a sequence of UI actions. The quality of the user's mental representation of the UI influences the user's success on this step. In turn, the quality of the user's mental representation of the UI is influenced by features of the UI [Norman 1990].
Stage 4.	Physically and mentally, the user executes the UI action sequence.
Stage 5.	The user perceives the state of the physical world (UI).
Stage 6.	The user interprets the state of the physical world.
Stage 7.	The user evaluates the state of the UI in the context of the task. Specifically, the user evaluates how well have their task intentions been realized.

2. BACKGROUND

2.1. The Norman User-Centered Model of Interaction

In 1986, Norman proposed a seven stage model of user-centered interaction [Norman 1986]. The model described mental and physical activities that users performed while completing task goals (see Table I). The activities in the first two stages were task specific and mental (independent of UI), activities in Stage 3 were mental but focused on the specifics of actions within the UI, activities in Stage 4 were UI specific, and activities in Stages 5–7 evaluated UI actions, compared to the task intention. Within the Norman model, the distance between the user's task intention and operation of the UI was called the *gulf of execution*. The distance between the operation of the UI and the evaluation of the final state of the UI was called the *gulf of evaluation*. Norman [1986] and colleagues [Hutchins et al. 1986] suggested that the shorter these gulfs, the less cognitive resources required and the more effective the UI. In the last 30 years, the Norman model has been used to characterize a wide variety of UIs (e.g., Lam [2008]).

2.1.1. Mental Representations in the Norman Model. Historically, *mental representation* has referred to a person's internal portrayal of information (knowledge and skills) to make inferences about the world, solve a problem, operate a device, and so on [Chase and Simon 1973]. The stages of the Norman model linked user activities with mental representations of task and UI. A mental representation of UI was believed to contain operational and functional information about the UI. An enhanced representation of UI enabled the user to efficiently develop a sequence of UI actions in Stage 3 of the Norman model [diSessa 1986]. A mental representation of task was thought to contain information about the current task state, goal state, and possible transformations. An enhanced task representation made the goal state and the path through the problem space to the goal state more clear and explicit [Chi et al. 1981; Moreno and Mayer 2007]. Together, enhanced mental representations of task and UI were held to reduce cognitive load and gulfs of execution and evaluation.

2.1.2. Mental Representations of Spatial Information. Post-WIMP UIs are typically UIs that use controllers, matched to the task, to manipulate virtual objects in virtual space. Cognitively, object manipulation tasks supported by these post-WIMP UIs involve processing of spatial information. Prior research has suggested that spatial information is mentally represented as analogous imagery, like a picture of a real object in space. Therefore, throughout the remainder of this article, we assume that an *enriched* mental representation is one in which key information, in the form of imagery, is isomorphic to reality (see also Shepard [1984], Shepard and Metzler [1971], and Finke [1980]).

2.1.3. The Norman Model and Problem Solving. The Norman model was ultimately a model of problem solving. However, it was a model of a special type of problem solving, where problem solving was significantly influenced by the user interaction. Within the Norman model, there was an implied sequence of events and dependencies, specifically that the representation of task influenced the representation and operation of the UI (Stages 3 and 4). The reverse theoretical relationship, that is the representation and operation of UI influenced the representation of task, typically was not made explicit. In other words, the idea that interaction itself might influence the representation of task, was not a usual focus of work involving the Norman model. For example, the model indicated that the user would reset their UI actions if the actions had not achieved their task intention (Stage 7). The model, while not explicitly prohibiting, did not predict that the user would reset their task intention as well. One of our assertions about post-WIMP UIs, that the experience of interaction influenced the mental representation of task, hypothesizes that this reverse relationship could also occur.

2.2. Mental Representations of UI

Users of post-WIMP UIs are often highly effective at their target tasks, yet post-WIMP UIs are quite varied in their tasks and interaction specifics [Jacob et al. 2008; Ullmer et al. 2005; Hippler et al. 2011; Poor 2008; Mott et al. 2012]. It is reasonable to ask, "How this is the case?" Jacob et al. [2008] has suggested that "new interaction styles share salient and important commonalities, which can help us understand, connect, and analyze them" (p. 201). In Jacob et al.'s view, post-WIMP UIs enabled users to draw upon *preexisting* knowledge of the real world as they formed mental representations of UI. In other words, even if a user had not experienced a particular post-WIMP UI, their prior knowledge of the real world informed their mental representations of the UI. Jacob et al.'s [2008] Reality-Based Interaction (RBI) model described four types of implicit knowledge that users brought to their mental representations of post-WIMP UIs. RBI posited three types of implicit knowledge related to object manipulation (1) naïve physics awareness, (2) body awareness and skills, and (3) environmental awareness.[1]

(1) Jacob et al. defined *naïve physics awareness* as a person's knowledge about the physics of objects. Jacob et al. suggested that when a person swiped a smartphone object and the object moved across the screen, the object appeared to have mass and constraints in space. Therefore, users had naïve knowledge of the physics of the virtual object and were not surprised by responses of the UI to user actions.
(2) Jacob et al. indicated that people had knowledge about, and skills for, controlling their own bodies, called *body awareness and skills*. Using gestural interaction as an example, Jacob et al. indicated that when users waved an arm downward as a part of the interaction, they were not surprised when an element of the interaction acted as if it had been pushed in response to their arm movement.
(3) Jacob et al. characterized *environmental awareness knowledge* as users' knowledge about their own orientation in space and, therefore, the orientation of objects around them in space. Poor et al. [2013b] noted that some virtual reality interactions allowed the user to perform interactions from a true egocentric perspective, rather than from display-defined frame of reference.

[1]The fourth theme of social awareness is not relevant to the present work.

To reiterate, RBI identified types of real-world knowledge that users incorporated into their mental representations of UI. Users applied this knowledge to form sequences of UI actions (Stage 3 of the Norman model). "By drawing upon these themes of reality, emerging interaction styles often reduce the gulf of execution [Poor 2008, p. 201] ."

Zhang [2012] noted that mental representations of UI contained both operational and functional elements. Older studies of the mental representations used during programming comprehension made a similar point [Pennington 1987]. In post-WIMP UIs, operation and function are the same or closely related. RBI knowledge, especially as it pertains to object manipulation, informs the functional aspects of the mental representation of the UI, and presumably the operational aspects as well. By contrast, in a WIMP UI, a user's mental representation can be dominated by operational details, which are not *a priori* informed by prior real-world knowledge (see also Beaudouin-Lafon [2000]).

In summary, we hypothesize that at least some of the performance improvements reported with post-WIMP UIs are due to users' enhanced mental representations of UI. In post-WIMP UIs, users incorporate their own real-world knowledge into their mental representations. Enhanced mental representations of UI allow users to make predictions about the outcomes of their actions and therefore form more effective and efficient action sequences (Stage 3 of the Norman model).[2]

2.3. Mental Representations of Task

The Norman model (see Table I) indicated that users' mental representations of task influence performance, strategy, and evaluation in Stages 2 and 7. Is it possible that the UI itself may influence the user's mental representations of the task as well? While the RBI notions say little about the nature of mental representation of task, there are additional lines of research that suggest that any action (even UI action) can affect the mental representation of a task. In many post-WIMP UIs, the UI action involves multiple modalities (e.g., vision, touch, proprioception). A UI action potentially enriches the quality of mental representation of the task by mapping multi-modal information onto the representation. The resulting representation includes a more well-structured problem space. Research in non-UI-situations has indicated that multi-sensory inputs enhance mental representations (see Wraga et al. [2000], Vidal et al. [2009], and Woods et al. [2008]). Additionally, in the mental representation literature, a handful of studies have suggested that tactile and visual information map onto a common representational system [Easton et al. 1997a, 1997b].

2.4. Other Factors Influencing Mental Representations of UI or Task

In addition to real-world knowledge, Jacob et al. [2008] posited that individual differences might lead to differences in mental representations of UI. Within the psychological literature, a number of studies have shown that individual characteristics, particularly spatial ability (SA), influenced mental representations, and user performance in object manipulation tasks [Just and Carpenter 1985; Chase and Simon 1973]. Additionally, Jacob et al. [2008] noted that ergonomic factors might play a role in users' experiences with post-WIMP UIs as well, masking or compensating for deficiencies in mental representations.

2.5. Alternative Theoretical Models of Post-WIMP UIs

While we focus on the impact of mental representations of task and UI through the Norman model, Lam [2008] presented an alternative cognitive framework, based on

[2]Some authors have explained the efficacy of post-WIMP UIs in terms of distributed cognition [Hollan et al. 2000], a concept which may have explanatory power but is outside the realm of the current work.

Norman, to compare post-WIMP UIs. In her work, Lam [2008] focused on cognitive costs. For example, in Norman Stage 2, her focus was on the decision-making and information gathering costs involved in building task goals and intentions. Costs of building and using mental representations and richness of representations are closely related. Unlike Lam [2008], we opt to focus on richness of representation.

Beaudouin-Lafon's [2000] Instrumental Interaction Model offered a second alternative model to characterize contemporary UIs. Their framework focused on UIs as mediators between users and tasks, described by the following three measures:

(1) *The degree of indirection* was defined as the two-dimensional spatial or temporal offset between the interaction element and the manipulated object. For instance, a scrollbar has minimal spatial indirection but a dialog box, offset from the interaction element that it is controlling, has higher indirection because it is more spatially offset.
(2) *The degree of integration* was defined as the ratio between the degrees of freedom of the task elements and the input device. For example, a scroll bar has one degree of freedom but a mouse controller moves in two directions; the degree of integration between the scrollbar and mouse is 1/2.
(3) *The degree of compatibility* was defined as how closely the physical actions of the interaction mimic those of the manipulated object. For instance, dragging a UI object with a mouse or a touchscreen has higher compatibility than moving an object by a typed command.

In their comparison of various UIs, Beaudouin-Lafon [2000] concluded that graspable, post-WIMP UIs were high on these three characteristics. Traditional WIMP UIs were low, especially on parameters of indirection and compatibility. For our purposes, we acknowledge that the WIMP and post-WIMP UIs might engender identical mental representations, but have key mediation differences under the Beaudouin-Lafon [2000] model. However, we believe that the dimension of *compatibility* was in line with the ideas of Jacob et al. [2008]. In other words, connecting the Instrumental Interaction Model and RBI suggests that the more *compatible* a UI with the user's task, the more that the user should use their own *real-world knowledge to generate sequences of UI actions*. Jacob et al. admitted that WIMP and post-WIMP also might differ in ergonomic properties, properties well described by the Beaudouin-Lafon [2000]'s notions of integration and indirection [Jacob et al. 2008]. We conclude that Beaudouin-Lafon [2000] does not *a priori* discount our notion that richer user mental representations of post-WIMP UIs explains much of their power. Rather the Instrumental Interaction is a tool that we will use to evaluate the degree to which the UIs in our study are post-WIMP.

Finally, while many authors have utilized the Norman model to explain performance differences across interaction styles, it is possible that other factors, outside of those inherent in the Norman model, are in play in post-WIMP UIs. For example, Svendsen [1991] posited that time spent on problem solving, including time to manipulate the UI, influenced the user's overall performance. A recent study by Donahue et al. [2013] put Svedsen's [1991] explanation to the test for post-WIMP UIs. Using Svedsen's [1991] explanation, Donahue et al. [2013] predicted that while the post-WIMP UIs would be easier to interact with (i.e., allowing faster manipulation), they would lead to less efficient task performance due to reduced planning/reflection time. Donahue et al.'s [2013] findings were quite the opposite. They found that post-WIMP UIs were easier to interact with (as evidenced by faster manipulation), and produced more efficient problem solving. These results suggested that performance improvements with post-WIMP UIs cannot be explained simply by the time taken to complete the task.

2.6. Summary of Background

To summarize, we propose that post-WIMP UI performance improvements are well explained by the Norman model. Specifically, the two mental representations, central to the Norman model, are enhanced within the post-WIMP UI experience:

(1) *Enhanced mental representation of UI*: Users' abilities to draw on pre-existing knowledge about the world and their bodies in the world allow them to build enriched mental representations of the UI. Enhanced representations allow users to more efficiently build UI action sequences needed to achieve their intentions (Stage 3 of the Norman model). Post-WIMP UIs with higher levels of task/UI compatibility allow users to more easily apply their real-world knowledge. Because of low compatibility [Beaudouin-Lafon 2000], mental representations for WIMP UIs must support both functional and operational knowledge, with a heavy draw on cognitive resources to align the two.
(2) *Enhanced mental representation of task*: The meaningful multi-sensory properties inherent in the UI controllers of some post-WIMP interactions enrich the user's mental representation of the task, influencing both task intention and evaluation of how well intentions were met (Stages 2 and 7 of the Norman model).

In addition, we note that individual differences may influence a person's mental representations of task and UI, and by extension, user outcomes with post-WIMP UIs as well. Ergonomic factors may play a role in users' experiences with post-WIMP UIs.

Relative to the Norman model itself, the background literature indicated that the representation and operation of UI informed the representation of task. This idea typically has not been an explicit element of the Norman model. Post-WIMP UIs with their rich multi-sensory controllers may inform the Norman model and representation of task in ways that prior generations of UIs did not.

3. THE CUBE COMPARISON TASK

RBI connected post-WIMP UIs to three categories of real-world knowledge. The categories were related to object manipulation and the associated cognition of objects in space. To test our theory that RBI knowledge was connected to enhanced mental representations of UIs in post-WIMP UI situations, it was critical that we identified a task that involved cognition of objects. Of course our task had to be one for which we could build a matched post-WIMP UI. Finally, any matched post-WIMP UI needed to have multi-sensory properties so that we could test our theoretical assertion that the multi-sensory interaction enhanced mental representations of task. We reasoned that an object rotation task would be ideal. Mental rotation tasks have long been used to study the nature and quality of mental representations of space (and objects). In these tasks, participants were presented with two objects, one of which could be rotated in either two-dimensional [Cooper and Shepard 1975] or three-dimensional space [Shepard and Metzler 1971]. The task of the participant was to, as quickly as possible, mentally rotate one or both objects to determine if the objects were the same or different.

A number of different factors are known to affect a person's responses in mental rotation tasks, including individual differences in SA [Just and Carpenter 1985], stimulus complexity [Bethell-Fox and Shepard 1988], and angular disparity between the orientations of the two objects [Cooper and Shepard 1975; Shepard and Metzler 1971]. In particular, as angular disparity increased, the time to solve the problem also increased, often in a linear fashion. Such an effect is believed to reflect transformation processes on the mental representation of objects; longer rotations take more time, much like rotations of real-world objects. Thus, mental rotation tasks can be used to study the nature of mental representations. Additional neuroscience research has suggested that rotational processes in play in mental rotation tasks are tied to perceptual systems involved in mental representations [Shepard 1984; Gauthier et al. 2002]. Finally, prior research has indicated that motor and mental rotation of objects draw on the same cognitive processes [Wexler et al. 1998; Wiedenbauer et al. 2007]. Thus, UIs that manipulate virtual objects engage the same cognitive mechanisms at work in mental rotation tasks and *vice versa*.

In our prior work and in the study presented below, we used an object rotation task called the Cube Comparison Task (CCT). The CCT was based on the original task developed by Thurstone [1938] and researched extensively by Just and Carpenter [1985]. In the CCT, the participant was presented with two opaque cubes placed side by side, with patterns (usually symbols or letters) on the visible faces. The patterns could be oriented in any way, but they were not repeated on more than one cube face. The task of the participant, like in other mental rotation tasks, was to determine as quickly and accurately if the two cubes could be the same (i.e., possessed the same patterns in the same organization) or if they were definitely different.

Unlike other studies of mental rotation, in the CCT, it was not simple to define angular disparity. Fortunately, in their work with the CCT, Just and Carpenter [1985] identified six main problem types of varying difficulty (defined by the number of rotations needed to align the cubes and number of visible matching faces). They demonstrated that, within the CCT, these problem types manifested the same results as angular disparity in other mental rotation tasks.

In addition to the mental rotation aspects of the CCT, Just and Carpenter [1985] treated the CCT as a problem-solving task. Using this secondary focus, Just and Carpenter [1985] explored both task completion variables (accuracy and reaction time), and participants' strategic variables. As the CCT can be solved in a number of different ways (due to the mathematical isotropy of cubes), choice of strategy was thought to be reflective of the person's mental representation of the task. In particular, Just and Carpenter [1985] found that persons with high levels of SA, as measured by standardized tests and who should have enriched mental representations, showed superior strategies in solving the CCT. For example, in problems that required several rotations to solve, high SA individuals were able to visualize the entire needed chain of rotations at once. Low SA persons often could visualize only one rotation at a time. Thus, Just and Carpenter [1985] concluded that strategies shown in the CCT provided a window into the people's mental representations of task and their approaches to problem solving.

Though much of the work with the CCT (and mental rotation tasks in general) has involved the use of static, non-interactive objects, such an approach has its limitations. For example, the only way to determine a participant's strategy in a static task is to use verbal protocols. Verbal protocols do not necessarily capture all underlying problem-solving processes, as these processes may not be accessible to conscious awareness. In order to better measure strategy, much of our work with the CCT [Klopfer et al. 2007; Jaffee et al. 2010; Hippler et al. 2011; Poor et al. 2011], including the present study, used an interactive version of the task. In the interactive version of the CCT, participants rotated (in either direction) one of the cubes along the three Cartesian axes that pass through the faces of the cube. Using the interactive CCT, we have analyzed participants' patterns of rotations to understand their mental representations and problem-solving strategies. We have found that individual differences in SA [Jaffee et al. 2010], angle of projection of the cubes [Jaffee et al. 2010], and type of UI/interactive experience [Klopfer et al. 2007; Poor et al. 2011] impacted CCT performance and strategy. In addition, prior research has indicated that in mental rotation tasks which involved real objects or related aspects, users preferred rotations around the vertical axis. This finding reflected the evolutionary import of gravity (see also Gibson [1979], Friedman and Hall [1996], Waszak et al. [2005], and Battista and Peters [2010]). Jaffee et al. [2010] and

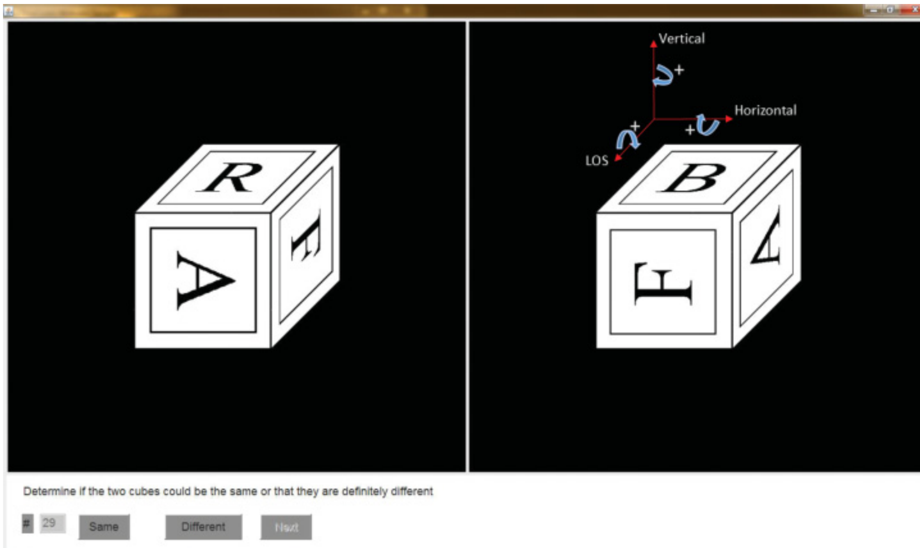

Fig. 1. Sample 270 degree rotation problem.

Hippler et al. [2011] found preferences for vertical axis rotations in some interactive CCT situations.

3.1. Summary of CCT

We chose a well-known object rotation task, the CCT, as the basis of our study of the theory presented in Section 2. Prior research has indicated that behavior on the static CCT reflected users' mental representation of task. Behavior on the interactive CCT has revealed features of mental representations of task and UI. Further, manipulating a UI for the interactive CCT potentially draws on the real-world knowledge identified by Jacob et al. [2008]. Thus, the interactive CCT is a useful vehicle to study mental representations of task and UI. However, prior studies of the interactive CCT have not explicitly separated measures of mental representation between task and UI. These are critical measures if we are to understand the influence of specific UIs on the activities and representations from the Norman model.

We now present an example problem from the CCT based on Just and Carpenter [1985] and shown in Figure 1.[3] To solve the problem, the person needs to rotate, mentally and/or with the UI, the right cube 270 degrees; small arrows superimposed onto Figure 1 show the required rotations. Table II and its associated figures illustrate activities from the interactive CCT and mental representations for WIMP and post-WIMP UIs involved in the example. As Table II shows, gulfs of execution and evaluation may be longer for interactive CCT WIMP UI users, due to limitations in mental representations of task and UI.

In the example shown in Table II, the post-WIMP users had a richer representations of the UI and task. Richer representation of the UI allowed users to execute their rotations as they intended. When WIMP users' representation of UI failed them, they improperly manipulated the UI. In addition, with richer representations of task, the post-WIMP participants of Table II had clearer and better-structured understandings of the problem space. Like the high SA CCT participants, observed

[3]Figure 1 and its illustrations of axes is an example rather than a specific UI from this or any other study.

Table II. Norman's 1986 Model of User Interaction, Post-WIMP UI, WIMP UI Example, and Gulfs of Execution/Evaluation

Norman stage	Example post-WIMP UI CCT	Example WIMP UI CCT	Gulfs of execution and evaluation
Stage 1. User forms a task goal.	The user sets a task goal of 270 degree rotation (see Figure 1).	The user sets a task goal of 270 degree rotation (see Figure 1))	
Stage 2. User forms a task intention within the task goal. Mentally, in forming the task intention, the user visualizes steps to achieve the task goal. The particular task intention is influenced by the richness of the user's mental representation of the task.	User has a richer representation of task, due to multi-sensory UI properties, experiences the cube as real and can visualize the necessary chain of rotations. He/she sets a task intention of clockwise rotation of the cube around the vertical axis as the first of the steps needed to achieve the match between the cubes. The intended rotation is to move the side cube face to the front face, as shown in Figure 4 (move "A" to the front face).	With a less rich representation of task, user experiences the interactive cube is less real and can only visualize one rotation at a time. To get started, user sets task intention of moving the side face (the "A" to the top face, as shown in Figure 2). This is a counterclockwise line-of-sight rotation.	.
Stage 3. Mentally, the user translates the desired task intention to a sequence of UI actions. The quality of the user's mental representation of the UI influences the user's success. In turn, the quality of the user's mental representation of the UI is influenced by RBI knowledge.	The user identifies the UI action needed to rotate the cube clockwise around the vertical axis. In a post-WIMP UI, the user real-world knowledge tells them to *push* the side face of the cube to the front to achieve the task intention. The UI actions are illustrated in Figure 4.	The user cannot apply RBI knowledge. Therefore, the user will translate their intention of rotation from side face to top into, for example, a series of button clicks. Unfortunately, in this example, the user selects button clicks that will place the top face onto the side, rather than the intention to rotate in the other direction. The WIMP user's planned UI action sequence is shown in Figure 3.	Translating intention to button clicks instead of simply planning to push, increases gulf of execution for WIMP UI user.
Stage 4. Physically and mentally, user executes action sequence in the UI.	The user pushes the side face to the front.	The user executes the planned button clicks.	Poor execution of UI actions can increase gulf of execution.
Stage 5. User perceives the state of the physical world (UI).	User perceives the new cube configuration.	User perceives the new cube configuration.	
Stage 6. User interprets the state of the physical world.	User interprets the new state of the rotated cube.	User interprets the new state of the rotated cube.	
Stage 7. User evaluates the state of the world in the context of the task. Specifically, the user evaluates how well their task intentions been realized.	Post-WIMP UI user determines that the side face is now on the front of the cube. User returns to Stage 2 to execute the next step toward 270 degree rotation. The sequence of rotations that were a part of the original mental representation of task are strengthened and clarified as the mental representation of task is enriched. Even if the post-WIMP user *had* chosen or made an unintended rotation at Stage 4, their richer task representation should better inform their assessment at Stage 7.	WIMP UI user has lost sight of the original side face (the "A"). The outcome does not match task intention. The process returns to Stage 2 to set a new intention in context of original goal and state of UI.	Gulf of evaluation is lengthened for WIMP UI user as execution of the task intention via the UI has gone awry. As the original side face is non-visible, the WIMP user uses cognitive resources determining whether the problem is with their intention, UI actions, or both.

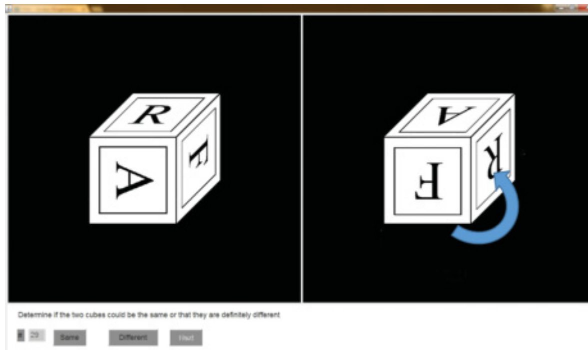

Fig. 2. Task intention WIMP user. This is a counterclockwise line-of-sight rotation.

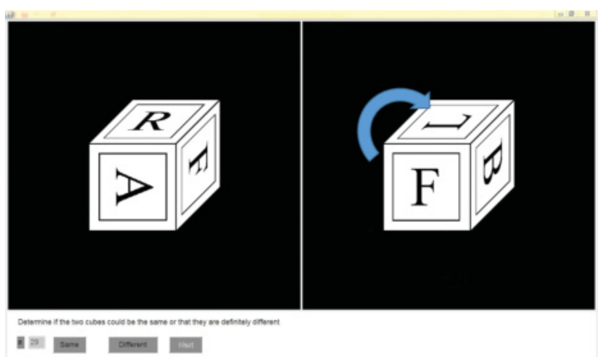

Fig. 3. WIMP user's UI action. This is a clockwise line-of-sight rotation.

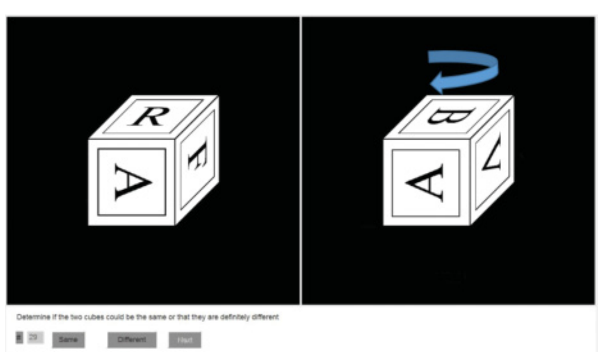

Fig. 4. Post-WIMP user's task intention and UI action. This is a clockwise vertical rotation.

by Just and Carpenter [1985], the post-WIMP users could "see" the steps to solve the problem of aligning the two cubes. But similar to the low SA participants of Just and Carpenter [1985], the WIMP users of Table II could only see one rotation at a time.

4. THE CURRENT STUDY

Theoretically, as shown in the example of Table II, post-WIMP UIs should enrich the mental representations of task and UI compared to WIMP UIs. The quality of these mental representations informs activities in the Norman model and influence gulfs of execution and evaluation. Our specific research question is whether or not post-WIMP UIs *actually* enrich these mental representations in the CCT. In the current study, we compared three UI conditions and two levels of the individual difference characteristic of SA, using the interactive CCT. We measured variables of richness of mental representations of UI and task as well as general performance measures. Our study had three hypotheses:

(1) *H1*: UIs which exploit RBI knowledge themes of naïve physics awareness, body awareness, and/or environmental awareness should lead participants to build enriched mental representations of the UI. Enriched representations allow users to more efficiently translate their task intentions into UI actions. In particular, persons who can exploit RBI knowledge should better predict impact of physics, body motion, and body position on on-screen objects and minimize UI actions that are outside their intentions (Stage 3 of the Norman model).
(2) *H2*: UIs that engage multi-sensory inputs in ways that are relevant to the CCT should yield enriched mental representations of task. Enhanced mental representations lead users to set task goals and intentions that differ from those who use a WIMP UI. In post-WIMP UIs, cube rotations, as indications of task intentions, should resemble the rotations of real objects. Resultant task intentions should be clearer and facilitate evaluation. (Stages 2 and 7 of the Norman model)
(3) *H3*: Persons of low SA by definition have less rich mental representations of task in the CCT, manifesting in patterns of cube rotations unlike those expected of real objects.

Based on prior studies of the interactive CCT, participants in the current study should successfully solve the problems of CCT with high degrees of accuracy (e.g., Klopfer et al. [2007], Jaffee et al. [2010], Hippler et al. [2011], and Poor et al. [2011]). Differing ergonomic demands on users may influence performance, mental representations, and/or observed behaviors as well [Jacob et al. 2008].

4.1. General Description: Experimental Design

Our overarching goal in the design of our study was to test the three hypotheses stated in the prior section. First, we controlled for visual issues that might impact the CCT as follows. All presentations were shown in the same graphical projection of the cube onto the two-dimensional plane of the screen on identically sized screens (19inch). Additionally, the look of the presentations was the same; for example, the spacing of the cubes was the same. Also, numerous studies have shown that performance on the CCT is influenced by problem complexity, analogous to angular disparity in other object rotation tasks. Problem complexity can dominate the variance of measured variables in CCT studies. As the focus of the current work was not on the role of particular problem types, we opted for all participants to see the same problems, with the six "standard" problem types equally and randomly represented across the stimuli and collapsed for our dependent measures. The specific problem types, drawn from Just and Carpenter [1985] are shown in Figure 5.

4.1.1. Independent Variable – UI Conditions. Arguably for a given task, there is a continuum of UIs varying from WIMP to post-WIMP, with many possible UIs in-between. The in-between UIs might draw on some types of RBI knowledge, have some level of multi-sensory interaction and/or be evaluated at moderate levels of the metrics from the Instrumental Interaction Model. For our UI conditions, we chose three UIs that were at distinguishable points along this continuum for the CCT (1) Mouse (M) a WIMP UI, based on a mouse controller, (2) Touchscreen (TS) based on a touchscreen, and (3) Tangible (TG) a post-WIMP UI, based on a tangible controller. While these three

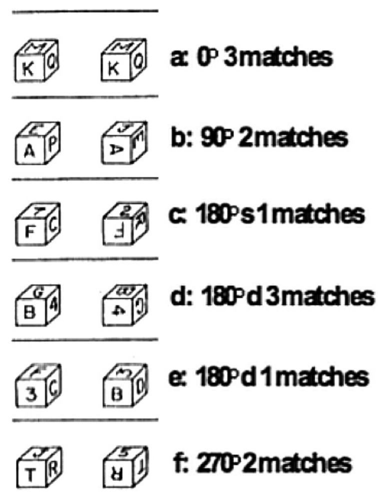

Fig. 5. Six problem types; reprinted with permission under the rules of the *Psychological Review* and the American Psychological Association.

interactions were arguably three of thousands that we could have selected, we chose them specifically because they differed on the features that we were interested in. In particular, the operation of the three UIs was differently informed by users' real-world knowledge [Jacob et al. 2008] while also differing in their multi-sensory properties. In addition, as indicated by the Instrumental Interaction Model [Beaudouin-Lafon 2000], the interactions had different moderator properties between the CCT and UI. We recognized that it would be not only naive but very difficult to select three interactions that *only* differed on the two issues (ties to real-world knowledge and multi-sensory properties) that we associated with the Norman model. For example, our three UI conditions had unavoidably different ergonomic characteristics. Therefore, to better characterize our three interactions, we provide detailed analyses of each of the three interactions. Additionally, we hope to explain why we chose the three interactions that we did, in light of the goals of this article and to provide a framework from which to understand our results. For each interaction, we describe (1) real-world knowledge engaged (RBI), (2) multi-sensory properties, (3) potential ergonomic challenges, and (4) characteristics described by the three dimensions of the Instrumental Interaction Model (indirection, integration, and compatibility) [Beaudouin-Lafon 2000].

Mouse interaction: The Mouse interaction (M) was implemented with a mouse controller and is shown in Figure 6. In the M condition, participants rotated the cubes by using the mouse to activate one of the six arrow-shaped buttons. The arrows indicated the direction of rotation along the three axes. The arrows were equidistantly located just beyond the edges of the cubes. In terms of the activities to rotate the cube, the participant selected one of six arrows, aligned the mouse pointer with the chosen arrow and clicked (see example in Figure 6).

The arrow tools resembled those used by Gallimore and Brown [1993] in their studies of interactive mental rotation tasks. Additionally, the UI was *purposefully* not post-WIMP, so as to provide a strong contrast, both in terms of interaction specifics and in terms of purported mental representations of UI and task, to the other two UI conditions. For example, while we could have implemented the cube rotation through a dragging motion of the mouse across the cube, we purposefully did not, so as to limit the multi-sensory engagement of the input controller in this condition. Had this interaction

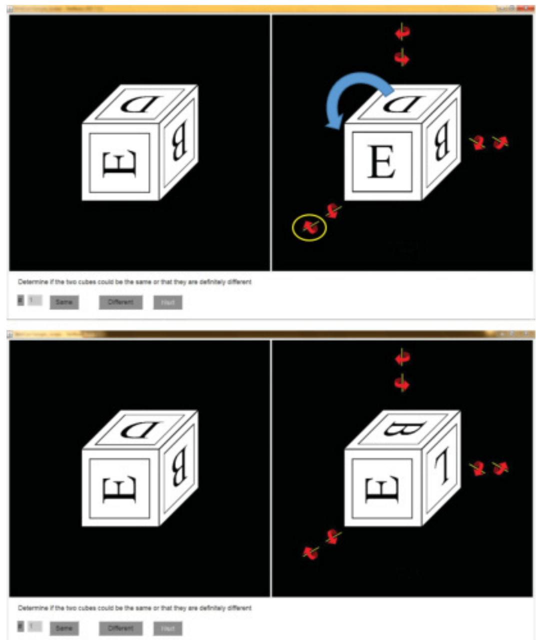

Fig. 6. Mouse interaction.

included dragging to achieve rotation, one could have argued that the sensation of proprioception and/or that RBI knowledge of body awareness were engaged. Using the arrow interaction insured that neither of these arguments about this interaction were true. M was characterized as follows:

(1) *Engagement of real-world knowledge*: Prior real-world knowledge of physics about object manipulation and/or body awareness told the M user little about relationships between arrow motions, body motions, and the behavior of the cube. M users had two frames of reference of the cube to align with their own: the display and the tabletop containing the mouse (see also Poor et al. [2013a]).
(2) *Multi-sensory properties*: M participants received information from the visual presentation exclusively and were not informed by multi-sensory inputs.
(3) *Ergonomic properties*: In M, the mouse controls were clearly visible. Localizing the mouse pointer to the arrows and pushing the mouse button should have offered minimal ergonomic challenges
(4) *Instrumental Interaction Model*[4]:
  (a) Degree of indirection: The interactive elements of M (i.e., the arrow buttons) were offset from the manipulated object (as depicted in Figure 6).
  (b) Degree of integration: For M, the arrow buttons were on/off binary controls, and therefore had zero degrees of freedom. The two-dimensional mouse controller operated in two dimensions and offered two degrees of freedom. Therefore, the degree of integration ratio of M was 0/2.
  (c) Degree of compatibility: M had low compatibility between the actual rotations of the three-dimensional cube and the action of pressing an arrow.

In summary, M was a reasonable representative of a WIMP UI.

[4]Unlike Beaudouin-Lafon [2000], we chose not to include a discussion of the temporal offset as it was more or less constant across UI conditions.

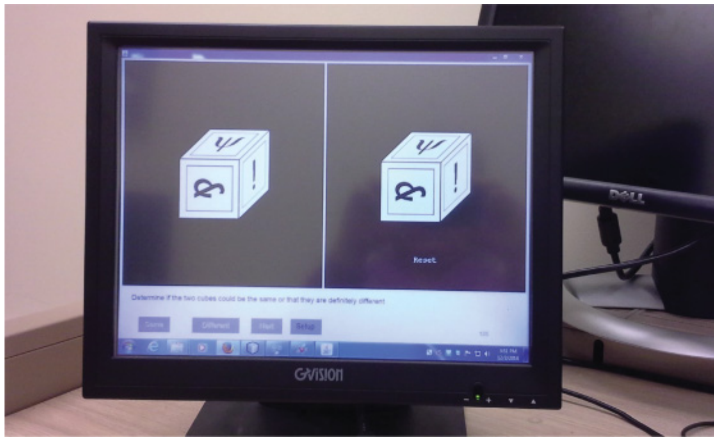

Fig. 7. Touchscreen interaction, cube rotation control via finger drag on cube hot spots.

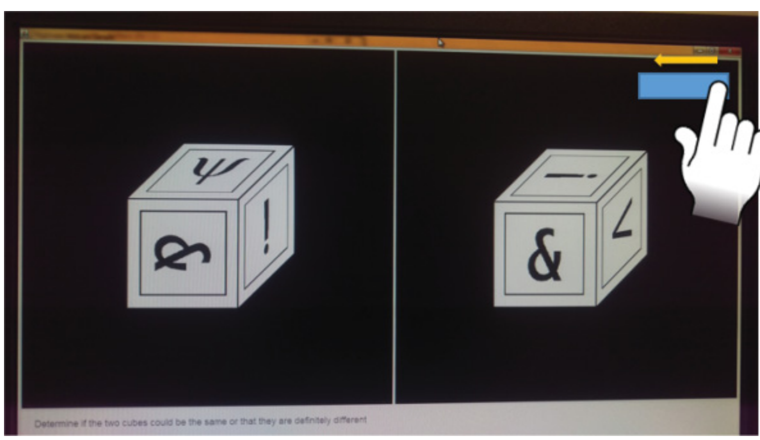

Fig. 8. Touchscreen interaction, close up – cube rotation control via finger drag on cube hot spots. The blue line illustrates a direction of finger drag.

Touchscreen interaction: The TS interaction is shown in Figures 7 and 8. Participants rotated the non-stationary (right) cube along all three Cartesian axes (both clockwise and counterclockwise) by locating their finger on the cube hot spots and dragging. Hot spots were located as follows:

(1) The horizontal axis hot spot was located directly above the cube 45 pixels high from the bottom of the cube image.
(2) The vertical axis hot spot was located directly on the right of the cube 180 pixels high from the bottom of the cube image.
(3) The line-of-sight axis hot spot was located directly to the upper right-hand corner of the cube 65 pixels high from the bottom of the cube image.

In terms of Norman Stages 3 and 4, the action sequence to rotate the cube was to align the participant's arm and hand to the hot spot on the cube and drag his/her hand in the desired direction. TS was further characterized as follows:

(1) *Engagement of real-world knowledge*: In terms of naïve physics knowledge, when participants dragged the on-screen cube with their finger, the cube reacted as if

it had mass and generally rotated in the direction of the user action. The one exception to this was the line-of-sight axis. When a user made a side-to-side motion, on the line-of-sight hot spot, the cube rotated around the line-of-sight axis. By contrast, side-to-side motion on the vertical hot spot generated a vertical axis rotation. In addition, TS engaged body awareness knowledge, whereas participant body movements yielded expected cube outcomes. For example, if participants wanted to rotate the on-screen object to the left around the vertical axis, they knew they needed to move their hand to the left across the vertical hot spot. Finally, when participants engaged the on-screen cube, they aligned themselves with the cube, permitting use of environmental awareness knowledge. Participants, by virtue of connecting their body to the object, established the primary frame of reference of the on-screen object as their own, rather than that of the display.

(2) *Multi-sensory properties*: In terms of mental representation of task, intention, and relevant multi-sensory perceptual inputs, TS participants received information from the visual presentation. The TS participants potentially gained a sense of the cube size by having their hands on the on-screen cubes. We expected that the mental representation of the task would be enriched, although limited, as the "cube" that they manipulated with their hands had neither solid nor haptic properties.

(3) *Ergonomic properties*: In general, it is well known that touchscreen users can face ergonomic challenges, such as "fat finger" (see also Lee and Zhai [2009] and Moscovich [2009]) and difficulties in finding hot spots. The current TS presentation might have posed these challenges for some users. In addition, for this task, the actual drag motions differed by rotation type (up/down, right/left) and might have offered unique ergonomic differences.

(4) *Instrumental Interaction Model*:

   (a) Degree of indirection: The interactive elements of TS were spatially close to the on-screen object as the hot spots were positioned directly on the edges of the cube.
   (b) Degree of integration: On the screen, the operational instruments of TS were essentially sliders; this type of instrument offered one degree of freedom. The touchscreen controller was aware of motion on two dimensions. Therefore, the ratio of degree of integration for the TS was 1/2.
   (c) Degree of compatibility: We considered TS to be intermediately compatible as the action of "touch and drag" did partially resemble the actual action of rotating a three-dimensional object for horizontal and vertical rotations. For line-of-sight axis rotations, the user dragged sideways to effect a rotation around an imaginary third axis. In addition, compatibility was limited by lack of haptic feedback,

In conclusion, TS had some properties of post-WIMP UIs but was not a post-WIMP UI.

Tangible/Post-WIMP interaction: The TG interaction, shown in Figures 9 and 10, was a post-WIMP UI. The participant rotated a 2×2 inch square white wooden cube in one of six directions along 90 degree rotation patterns. The participant immediately saw the corresponding result of their rotation on the (right) cube on their screen. To implement TG, we attached topcodes [Horn 2012] on each face of the physical cube. Users actions were tracked using a camera, with corresponding actions translated into rotations of the on-screen cube. Users were required to keep the physical cube inside of a 2×2 inch square on the tabletop surface to guarantee that the camera would correctly extract the topcodes. Additionally, users were limited to 90 degree rotations across the three Cartesian axes. In the TG condition, functional and operational knowledge of the UI were similar.

Fig. 9. Post-WIMP interaction, cube rotation control via rotation of physical cube.

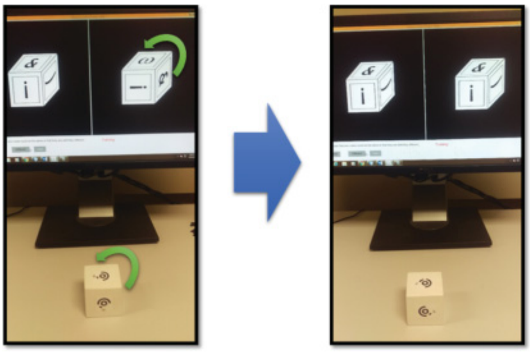

Fig. 10. Post-WIMP interaction, close up – cube rotation control via rotation of physical cube.

In terms of Norman Stages 3 and 4, if a participant's intention was to rotate clockwise along the vertical axis, the action sequence was to grab the physical cube and, using their hand, rotate the physical cube in the desired direction. TG, unlike M or TS, did not require localizing a pointing device (mouse pointer or finger) to a hot spot on the screen.

(1) *Engagement of real-world knowledge*: When participants moved the physical cube, the on-screen cube reacted as if it had mass and sensitivity to gravity. In other words, when a person rotated the physical cube, the on-screen cube correspondingly rotated in the direction of the physical cube movement. The physical cube also provided a sense of the tangible properties of the cube, such as the amount of effort it would require for the user to manipulate the cube, drawing on naïve physics knowledge. Secondly, when participants moved the physical cube, they drew on their previous experience of manipulating physical objects with their hands, engaging body awareness knowledge. Finally, when participants focused on the tangible cube, the orientation of the physical cube was in the egocentric frame of reference, yet the on-screen cube not precisely aligned with them. If participants focused on the on-screen cube, the physical cube was no longer aligned with them. Thus, we did not expect that participants' environmental awareness knowledge of the relation of their current orientation relative to the on-screen cube enhanced their representation of UI.

Table III. Summary of Analyses of UI Conditions

UI conditions	RBI knowledge engaged	Multi-sensory properties	Ergonomic properties	Instrumental Interaction Model analysis
Mouse (M)	Minimal	Visual	Minimal	Low moderator properties
Touchscreen (TS)	Naive physics, body awareness, environmental awareness	Visual, via touch, possible sense of cube size	Drag motion differed by axis. Possible “fat finger”	Intermediate moderator properties
Tangible (TG)	Naive physics, body awareness, environmental awareness	Visual, from physical cube (touch, proprioception)	Two visual foci, cube rotated within a constrained space	High moderator properties

(2) *Multi-sensory properties*: TG was multi-sensory with information coming from both the visual presentation and the physical cube; the cube engaged visual, touch, and proprioceptive sensations. The representation of task should be highly enriched.
(3) *Ergonomic properties*: Users have two foci of attention: the cube for their hand and the display for their eyes. In addition, if participants rotated the physical cube with one hand, it was possible that one or more of those rotational motions might be more difficult to accomplish than others. Finally, rotating the tangible cube within the constrained space might require minor manual dexterity and minor motor skill.
(4) *Instrumental Interaction Model*:
    (a) Degree of indirection: TG was quite distant as the input device (the tangible object) was positioned in real space, away from the virtually manipulated cube.
    (b) Degree of integration: TG differed from the other two interactions (M and TS) in that it had a 3/3 (1) integration ratio; both the virtual and physical cube rotated through three degrees of freedom. In other words, even though the physical cube was rotated within a physical space, 90 degree rotations on the three Cartesian axes were possible.
    (c) Degree of compatibility: The user actions with TG were quite close to those of the virtual object, as rotations with the tangible object were directly mapped onto rotations with the virtual object, within the constraints of the rotation axes and directions.

    We concluded that TG was a reasonable post-WIMP UI for the CCT.

Summary of UI conditions: Our three UIs, selected from across a continuum of UIs for the CCT, differed in terms of the real-world knowledge engaged and their multi-sensory properties, issues that we have argued, in our review of background literature, contributed to the richness of users’ mental representations of UI and task. Our three interactions had unavoidable ergonomic differences as well. Within our comparison of our interactions, we also identified differences via the Instrumental Interaction Model. These differences corresponded in part to differences delineated by the RBI model. Particularly in the case of TS, the Instrumental Interaction framework described the way that implementation of the tool for line-of-sight rotations differed from the tools for the other axes. The Instrumental Interaction analyses additionally confirmed our assessment of the relative “post-WIMPness” of the three UI conditions. We summarize our analyses of the UI conditions in Table III.

4.1.2. Independent Variable 2 – Spatial Ability. SA has historically referred to skill in representing, transforming, generating, and recalling symbolic, non-linguistic information.

In their reviews of the psychometric literature on SA, Caroll [1993] and Lohman [1996] listed three primary SA factors: *Visualization*, the ability to manipulate visual patterns; *Speeded rotation*, the speed of manipulating simple patterns by means of a single rotation; and *Perceptual speed*, the speed in finding a known visual pattern (e.g., the letter A) or in comparing two patterns. For our study, we measured SA using two standardized tests of visualization and speeded rotation (described below). We note that while some studies of mental rotation have found differences by gender, we chose to focus on more direct measures of SA, that is, scores on standardized SA tests. In past, we found that on the interactive CCT, there were no significant interactions between UIs and gender (e.g., Poor et al. [2013]).[5]

4.1.3. Dependent Variables 1 and 2 – Measures of Richness of Mental Representation of UI and Task. To test our theoretical assertions relating mental representations, the Norman model and UIs, it was essential that we considered measurements of the richness of these representations. In our prior studies of the CCT, we found that user choices of axes of rotation revealed deeper behavioral phenomena. Thus, we incorporated two dependent measures, based on user choice of axes of rotation, that we operationalized as richness of representations.

Operationalization of richness of representation of UI: *measuring rotation counts on the one-rotation problems*. Prior research has indicated that difficulty in mental rotation tasks increases linearly with the angular disparity of the two objects. In the CCT, the notion of disparity translates to the number of 90 degree rotations needed to align the two cubes. A problem type used in the study (shown in Figure 5, as type “b”) involves only one rotation to align the two cubes. The findings of Just and Carpenter [1985] and Klopfer et al. [2007] confirmed that this problem type likely involved only one mental rotation. Further Just and Carpenter [1985] and Klopfer et al. [2007] indicated that this problem type was not sensitive to differences in SA and that users universally had the same task intention of making the one rotation. As such, differences in strategy for this problem were likely due to differences in mental representation of UI and related to differences in RBI knowledge.

We generated one-rotation problem rotation counts as follows. For each user, for each trial of the one-rotation problem type, we counted the number of rotations they made to acquire a match, if the user rotated the interactive cube to match the non-movable cube. Trials that showed an entirely mental process, with no interactive rotations, and trials that did not end in a match were not included in our count. Ideal UI behavior, reflecting a rich mental representation of UI, should be *one* rotation on average per trial for this problem type. We operationalized superior mental representation of UI as rotation counts on the one-rotation problem that were closest to one.

Operationalization of richness of representation of task: measuring choice of rotation by axis. In the current study, in all three UI conditions, participants could rotate the cubes across the three Cartesian axes of horizontal, vertical, and line of sight. No other rotational axes were available. By definition, when users rotated, they chose rotations across these three axes. Following earlier research, users who experienced the cubes as most realistic should “see” the on-screen cubes as real objects and have the richest mental representations of the task. Prior research has indicated that when people experience the on-screen cubes as real, they select rotations around the vertical axis. In terms of Stage 2 of the Norman model, if the UI enriched the user’s mental representation of task, then the user’s choice of task intention should favor vertical rotations (see

[5]We recognize that spatial ability is not a true independent because we cannot vary it systematically across the study. It is an individual difference covariate. In spite of this, for simplicity, we have referred to it as an independent variable.

also Shiffrar and Shepard [1991] and Jaffee et al. [2013]).[6] We operationalized richer mental representation of task as higher percentages of vertical rotations as compared to the other two axes.

4.1.4. Dependent Variables 3 and 4 – General Performance Measures. We measured two general performance measures, accuracy and reaction time, to verify that users successfully solved the CCT problems in our study. Accuracy was defined as percentage of trials where the user correctly identified if the two cubes were the same or different. Reaction time was defined as the time (measured in ms) from when the user first saw the two cubes to the time that they moved to the next trial. Extremely poor performance in any of the UI conditions would suggest a complete inability of users to solve the CCT problems, making all other results suspect. However, if users showed evidence of rich representations, as measured by dependent variables 1 and 2, but low accuracy and/or slow reaction time, ergonomic problems in the UIs might be in play.

In prior studies of the interactive CCT, different UIs yielded no differences in accuracy level but differences in reaction time (e.g., Jaffee et al. [2010, 2013]). Unlike non-interactive CCT results, Jaffee et al. [2010] found no differences in accuracy by SA on the interactive CCT, suggesting that for performance measures, the interactive experience can bridge the deficiencies in mental representations of task and/or UI endemic in low SA individuals (see also Cooper and Shepard [1975]).

4.2. Specific Predictions for Hypotheses H1–H3 as Reflected by the Dependent Measures

(1) Hypothesis H1 stated that *engaging RBI knowledge should enrich mental representations of UI*. Participants in the TS and TG conditions should engage RBI knowledge to a greater degree than the M condition. Our dependent measure of *rotation counts on one-rotation problems* was our measure of the quality of mental representation of UI. As a result of enhanced representations of UI, we expected that TS and TG users would have rotation counts on the one-rotation problem near the ideal of one.
(2) Hypothesis H2 stated that *using multi-sensory UIs should enrich mental representation of task*. The TG condition involved multi-sensory inputs to a greater degree than the other two conditions. As such, we expected that mental representations of task for TG participants to be richer and more similar to the representations for problems involving physical cubes. Our dependent measure of *choice of rotation by axis* was our measure of the quality of mental representation of task. As a result of enhanced representations of task, we anticipated that persons in TG should favor rotations on the vertical axis.
(3) Hypothesis H3 stated that *persons of lower SA should have less rich mental representation of task*. Our dependent measure of *choice of rotation by axis* was our measure of the quality of mental representation of task. As a result of enhanced representations of task, we anticipated that high SA participants should have greater preference for vertical rotations.

4.2.1. Additional Predictions for General Performance Measures. In our study, we collected general performance measures of accuracy and reaction time. Based on our prior studies of the interactive CCT, we anticipated that people in all UI conditions should manifest high accuracy. Touchscreens in general are faster than other interactions (e.g., Sears and Shneiderman [1991]), but may limit user speed through ergonomic limitations such as "fat finger" (see also Lee and Zhai [2009] and Moscovich [2009]). Thus, TS

[6]Marr and Nishihara [1978] and others have theorized that extraction of major axes is an important step in object perception and thus in formation of mental representations of objects.

might be slower than M. Similarly, moving the tangible cube in a space might require manual dexterity and minor motor skill and be slower than M.

4.3. Methods

4.3.1. Procedure. Participants arrived at the laboratory and first completed the standardized tests of SA, followed by 18 CCT training trials in their assigned UI condition.[7] The purpose of the training trials was to familiarize the participants with their UI conditions as well as the CCT itself. The training trials were followed by 72 CCT trials. The trials consisted of 36 "same" (i.e., where the left and right cubes were the same) and 36 "different" (i.e., where the left and right cubes differed) trials with 12 problems from each of the six problem types (shown in Figure 5). The participants all saw the trials in the same random order. Prior CCT research with the materials from this study have showed no order effects (see Klopfer et al. [2007]). Within our CCT stimuli, cube pairs (and problems) were evenly distributed by axis and direction.

4.3.2. Spatial Ability. Participants completed two standardized tests of SA: Card Rotation Task (a measure of speeded two-dimensional rotational ability) and Paper Folding Task (a measure of three-dimensional visualization ability) (see also Ekstrom et al. [1976]). From the two scores for each participant, we generated a composite standardized z-score. Using norms of standardized scores from our laboratory data of over 300 participants and a median split, we designated each participant as a SA level of "High (H)" or "Low (L)."[8]

4.3.3. Participants. A total of 67 persons participated, with volunteers drawn from a number of undergraduate computer science classes. Participants were given extra credit for participating as part of a set of choices of extra credit events. Four participants did not produce viable data, leaving 63 participants. By UI condition, participants were distributed as follows: M, n = 24 (H n = 16, L n = 8); TS, n = 22 (H n = 14, L n = 8); TG, n = 17 (H n = 8, L n = 9). Data was collected by UI conditions. Unfortunately, factors outside of the experimenters' control accounted for differing numbers of participants per condition. The participants were treated in accordance with the Declaration of Helsinki and the research was approved by the local institutional human subjects review board.

4.3.4. Materials/Apparati. The software for the task was built with Java and Java3D running on Windows XP using a 19inch flat screen Dell display. Participants performed the task on a variety of machines and graphic cards, due to the nature of lab machines; all display refresh times were systematically backed out of measures of reaction time. For each participant, for each trial, reaction time was calculated as total time minus the number of interactions times refresh rate.

4.3.5. Specific Description: Experimental Design. Our design consisted of two independent, between participant variables: interaction (UI) condition and SA.[9] In addition, as all participants were free to rotate around the three Cartesian axes, our independent variable of axis of rotation was a within participant variable. Our dependent measures included average rotation counts for the matched one-rotation problem type, percentage

[7]A "trial" in the CCT corresponds to one pair of cubes, where the task is to determine if the cubes could be the same or are definitely different.
[8]We chose to use the norms from the large set of data from our laboratory rather than using norms from only the current participants. We felt that the norms from the larger sample were more likely to be stable as compared to the smaller sample.
[9]Limitations in our research environment made it impossible for us to run the three UI conditions as within participant.

of rotations by axis, average accuracy, and average adjusted reaction time (adjusted for the refresh rate of the machines used in the study).

5. RESULTS

We present the results in three subsections. In the first subsection, we describe our analyses of averages of rotations for the one-rotation problems as a test of hypothesis H1. In the second subsection, we report on analyses of percentages of rotations along each of the three rotational axes as a test of hypothesis H2. In the second subsection, we also analyze the percentage of rotations by axis, as impacted by SA, a test of H3. In the third subsection, we report on the two measures of performance: accuracy and reaction time. For our analyses of the two measures of rotations, we included all trials where the correct answer was "same," regardless of whether the participant had the correct answer.[10,11,12] For both accuracy and reaction time, we follow the protocol of Just and Carpenter [1985] and analyzed the dependent measures for trials in which the correct answer was "same." For reaction time, we further limited analyses to trials that the user successfully answered the problem.

5.1. Rotation Counts on One-Rotation Problems – Test of Hypothesis H1

The mean number of rotations on one-rotation problems that resulted in rotation to a matched state was 1.545 rotations per trial. The means for M, TS, and TG were 2.042, 1.280, 1.186, respectively. We found a significant main effect of UI (F (2,57) = 10.667, $p < 0.001$). We found no significant effects of SA or the interaction of UI and SA. Tukey pairwise tests of differences between M and TS and M and TG were significant (mean difference = 0.761, $p < 0.001$ and mean difference = 0.855, $p < 0.001$, respectively); Tukey pairwise test of differences indicated that the TS and TG were not significantly different. To summarize, the measured rotation counts for TS and TG were close to one, suggesting that users were able to translate their mental task intention into UI actions isomorphically. M users required an extra rotation on average, per trial, beyond the ideal. This finding supported our prediction for hypothesis H1 that the post-WIMP UI participants had enhanced mental representations of UI.

5.2. Choice of Rotation by Axis – Tests of Hypotheses H2 and H3

For these analyses, we calculated percentages of axis rotations and collapsed the percentages across problem types. We excluded the identity problem type.[13] In rotations drawn from the remaining five problem types, we found no significant main effects of UI, axis, or SA, but we did find a significant interaction between axis and UI (ggF (3.352, 93.857) = 3.704 $p < 0.013$ $\eta_p^2 = 0.117$).[14] The percentages are listed in Table IV and shown in Figure 11.

5.2.1. UI Conditions and Choice of Rotation by Axis – Test of Hypothesis H2. A Tukey pairwise test of differences between M and TG for the percentages of rotation around the vertical axis was significant (mean difference = −8.379, $p < 0.0233$). Additionally, a Tukey pairwise test of differences indicated that the TS and TG were significantly different

[10]Studies of mental rotation often focus on same trials (see also Just and Carpenter [1985] and Shiffrar and Shepard [1991]).
[11]Unless the main effects or interactions were significant, following standard reporting protocols, we do not report outcomes of pairwise tests.
[12]We used between and within ANOVA to analyze main and interaction effects.
[13]In identity problems, shown in Figure 2 as problem type "a," users saw two identical cubes; we counted essentially no rotations for identity problems.
[14]"ggF" refers to a Greenhouse–Geisser adjustment to repeated measures degrees of freedom when assumptions of sphericity are not met.

Table IV. Percentage Rotations By UI and Axis

	Horizontal	Vertical	Line of sight
Mouse	30.687%	31.161%	38.151%
Touch	33.862%	36.379%	29.738%
Tangible	26.752%	38.242%	35.006%

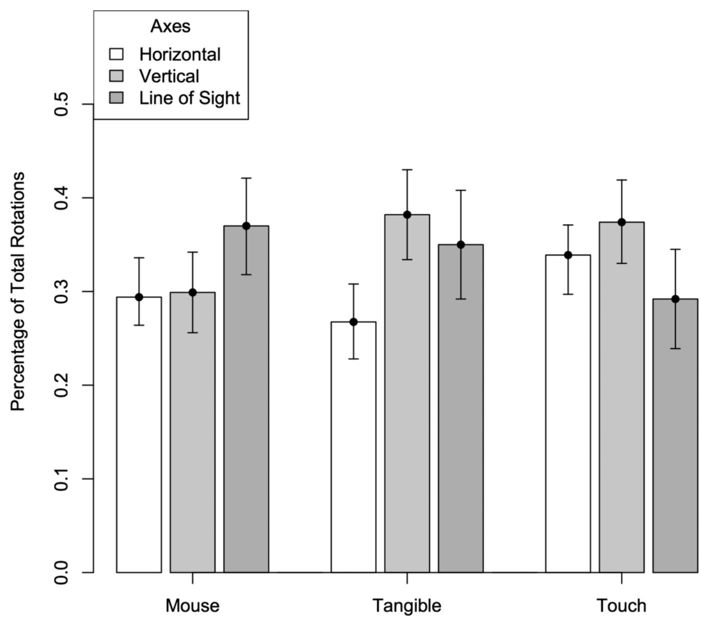

Fig. 11. Percentages of rotations by axis and UI (the error bars indicate the 95% confidence intervals).

from each other (mean difference = $7.130, p < 0.027$) for rotations around the horizontal axis. No other pairwise differences by axis and UI conditions were significant at the 0.05 level.

In conclusion, hypothesis H2 asserted that persons who had the highest multi-sensory experiences in their UI condition (specifically TG) should have richer mental representations of the CCT. As such, these persons should display rotational patterns most like those of persons rotating real cubes, favoring rotations around the vertical axis. TG users showed this pattern to a significantly higher extent as compared to M. The significant pairwise favoring of vertical axis rotations for TG users supported our prediction for hypothesis H2 that TG yielded the richest and most realistic mental representations of task and the objects contained within the task. While not conclusive, this result suggested that persons in the TG conditions experienced the cube interaction as more real and their mental representations of task as richer. Thus, we found support for hypothesis H2.

5.2.2. SA and Choice of Rotation by Axis – Test of Hypothesis H3. Hypothesis H3 postulated that low SA participants, would have limited mental representations of task. We found no differences for axis preference by SA; our results did not support hypothesis H3.

5.2.3. Choice of Rotation by Axis – Additional Findings. We found significant additional preferences for the horizontal control for TS. This result could be ergonomic. Our Instrumental Interaction Model and RBI analyses pointed to potential challenges in use of the line-of-sight axis for the TS conditions. The TS users could have favored rotations around the horizontal axis to avoid the rotations around the line-of-sight axis.

5.3. Accuracy and Reaction Time

To verify that our observations of the strategic-dependent variables described above were not due to simple failures in problem solving, we analyzed two performance measures of accuracy and reaction time. For analyses of accuracy and reaction time, we collapsed results across the six problem types shown in Figure 5. Overall, we found that average accuracy was 95.399%, with no significant differences in accuracy due to UI or SA conditions. The average reaction time was 10420.360ms. We found a significant effect of UI on reaction time (F (2,60) = 16.948, $p < 0.001$). TG was slower overall (13641.102ms) as compared to M (7630.549ms) and TS (9989.449ms). Tukey pairwise tests of differences by UI showed significant pairwise differences between M vs. TG and TS vs. TG (mean difference = −6113.773, $p < 0.001$ and mean difference = −3841.973, $p < 0.001$, respectively). We found no significant differences by SA for either performance measure, nor significant interactions. In summary, all UI conditions yielded high accuracy, suggesting that users were able to successfully solve the CCT problems. Breakdowns in ergonomic challenges and/or persistent unfamiliarity likely led to slower performance for the TG condition.

6. SUMMARY AND DISCUSSION OF FINDINGS

The research question in our study was to determine if UI conditions, ranging from traditional WIMP to post-WIMP, actually impacted users' mental representations of UI and task. In our study, we utilized a measure of richness of mental representation of UI, a measure of richness of mental representations of task, and two general performance measures. We found that the touchscreen and tangible UIs yielded richer representations of UI as compared to the mouse UI, in support of hypothesis H1. Our results also supported hypothesis H2 that users of the post-WIMP UI (TG) had enriched mental representations of task compared to the WIMP UI (M). Hypothesis H3 states that the mental representations of task for low SA individuals were less rich than the high SA users was not supported by our results. Finally, we confirmed that participants were highly accurate in all UI conditions. The TG users were slower than M participants. Our observed results, relative to our hypotheses were therefore not due to deficiencies in problem solving, although they may have reflected ergonomic challenges. Reviewing our results, we found some confirmation of our original predictions:

(1) Persons in the TS and TG UI conditions were near maximally efficient in their translation of a single-axis rotation intention to UI action. M users on average required an extra rotation, per measured trial (support for hypothesis H1).
(2) Persons in the post-WIMP, TG UI condition favored the vertical axis of rotation as compared to those in the WIMP condition of M (support for hypothesis H2).
(3) We found no differences by SA for the measure of richness of representation of task (no support for H3).
(4) All of the UI conditions led to high accuracy. Observed results were not due to deficiencies in problem solving. TG users had longer reaction time than those in the other two conditions and we could not rule out ergonomic challenges in that condition.

The Norman model predicted that users who experienced higher quality mental representations of UI and task should also experience reduced gulfs of execution and evaluation. Our research questions explored these assertions for a UI conditions along a continuum from WIMP to post-WIMP. In our study, we concluded the following. TS and TG users had enhanced representations of UI via RBI knowledge. TG users experienced enhanced representation of task, by virtue of the multi-sensory properties of the UI; M users showed neither enhancement. Additionally, high SA participants did not show evidence of enhanced representation of the task.

Prior applications of the Norman model to explain improved performance by UI conditions have attributed improvements to enrichment of the two representations in the model. We have shown that improvements to those two representations are in play in post-WIMP interactions. Moreover, we have identified mechanisms that account for the improvements. Finally, we have demonstrated that the act of interacting itself may have a feedback loop to the representation of task. Thus, our results are consistent with predictions of the Norman model *and* enhance the theory.

6.1. Additional Discussion of Approach and Results

Our strategy within this work is as follows:

(1) We identified a conceptual framework that allowed us to understand the apparent near universal performance advantages of post-WIMP UIs as compared to other interaction styles. We chose the Norman model as our theoretical base. We then identified recent work that suggested that the two mental representations (UI and task), central to the Norman model were enhanced with post-WIMP interactions.
(2) We formulated a specific research question: *Do post-WIMP UIs, by virtue of knowledge engaged and multi-sensory properties, lead to enhanced mental representation of UI and task?*
(3) We used our research question to derive our three stated hypotheses.
(4) We operationalized our richness of representation of UI and task. We operationalized points on this UI continuum by selecting three UI conditions. The WIMP condition had very limited multi-sensory aspects and was poorly informed by RBI knowledge. The post-WIMP condition had clear multi-sensory properties and was informed by RBI knowledge. The TS condition lie somewhere in between.
(5) We collected and analyzed our data. Our results supported our theoretical assertions of the interplay between our post-WIMP condition and our measures of richness of representations. The relationship between the post-WIMP condition and the representation of task was predictable from research outside of HCI. By finding that the relationship between the post-WIMP interaction and the representation of task exists, we have extended our understanding of the Norman model as well.

Thus, we followed a standard approach to theory-driven research. However, like all theory-driven research, it is always possible to draw the wrong conclusion. We see three specific limitations on our work, as summarized above. First, it was possible that our operationalization of WIMP and post-WIMP and our subsequent choices of actual UIs may not have been fully representational, thus our results may not generalize. Second of all, as we alluded to above, it was possible that alternative explanations of our findings may have been just as feasible as the ones that we chose. For example, near ideal performance on the one-rotation problem type for TS and TG could be explained by high compatibility within the Instrumental Interaction Model as opposed to improvements in mental representations. Preference for rotations on the vertical axis by the TG participants could reflect ergonomic avoidance of other axes. Finally, we had no direct measures of gulfs of execution and evaluation and whether they were actually influenced by the richness of mental representations. In fact, one performance measure, reaction time, suggested a possible longer gulf of execution for TG participants. In the next section, we revisit the data and explore the possibilities of alternative explanations of results.

7. REVISITING THE DATA

To shed possible light as to whether our interpretation of results from Section 5 had validity, we present a second set of analyses. These analyses are of patterns of rotations rather than of single rotations. We admit, in advance, that the patterns of rotations that we consider are not clean, distinguishable measures of richness of mental representations of task or UI. Thus, these analyses are exploratory.

7.1. Unintended/Intended Rotation Patterns

Early in our lab's studies of the interactive CCT, we observed that participants occasionally rotated the cube and then "undid" the rotation (e.g., Klopfer et al. [2007]). This rotation pattern could indicate that the user

(1) had a limited mental representation of task and were unsure of their task intention until they saw the outcome (Norman model Stage 2);
(2) had a limited mental representation and mistranslated their task intention into UI actions (Norman model Stage 3);
(3) made an ergonomic error relative to the physical operation of the user interface (Norman model Stage 4);
(4) had a limited mental representation of task and were uncertain of their task intention in the context of the observed state of the UI (Norman model Stage 7);
(5) engaged in an epistemic action. Epistemic behaviors in the CCT would be manipulations of the problem environment to get a better mental representation of the task and UI, though the action was not directly related to problem solving (see also Kirsh and Maglio [1994]).

Unfortunately, without extensive computational simulation and/or verbal protocols, we could not directly determine which "do-undo" rotation pattens were related to activities within Norman stages versus those that were epistemic. If we assumed that the majority of "do-undo" patterns of rotations were epistemic in nature, we could infer that either the user's mental representation of task limited their ability to predict or evaluate the outcome of a task intention or to predict the result of a UI rotation (the Norman model Stage 3). In either of these two episodic rotation examples, the "do-undo" rotation pattern would indicate that the user rotated to explore the task and the UI. On the other hand, if "do-undo" rotations indicated problem solving, ergonomic, or evaluative difficulties, they would also indicate limitations in mental representation. Therefore, though we could not distinguish episodic rotations from those relating to challenges in problem solving (Norman model Stages 2 and 3), ergonomics (Norman model Stage 4), or evaluation (Norman model Stage 7), "do-undo" rotation patterns potentially reveal limitations in users' mental representations.

We defined these patterns as *unintended/intended rotation patterns*. "Unintended rotations" meant a rotation on a given axis and direction that was followed by a rotation on the same axis in the opposite direction. In other words, an "unintended rotation" was one that was done and then undone. The rotation that followed, if there was one, was defined as the intended rotation. In prior CCT studies, we observed three common intended/unintended rotation patterns [Poor et al. 2013]:

(1) *Unintended rotation at the end of the UI inputs*: This rotation pattern occurred when the user rotated and then "undid" the rotation at the end of their UI input stream. The occurrence of extra rotations at the end of the UI inputs suggested that the user was unsure that they were finished rotating the cube. High frequencies of this type of pattern might indicate that the user's mental representation of task was limited as the user was unable to determine if the UI state matched their task

intention, lengthening the gulf of evaluation. High frequency of this pattern on only one axis for a given UI condition also might indicate ergonomic problems.

(2) *Unintended/intended rotations on the same axis*: In this case, the unintended rotation was on the same axis as the intended rotation, varying only by direction. Here, the user had found the UI control for the intended axis of rotation. The user's intention had gone awry when choosing the direction of rotation. As such, the user had either misidentified a UI action or made an error in execution (Norman model Stages 3 and 4). High frequency of this pattern on only one axis for a given UI condition also might indicate an ergonomic issue, not a problem-solving error. Occurrence of this pattern would lengthen the gulf of execution.

(3) *Unintended/intended rotation on a different axis*: Here, the unintended and intended rotations were on different axes. This rotation pattern appeared to be due to either a malformed task intention and/or a failure of UI action to meet task intention. Malformed task intention would reflect poor mental representation of the task (Norman model Stage 2). Failure to generate a UI action sequence to meet task intentions could reflect a poor mental representation of UI (the Norman model Stage 3). Occurrence of this pattern would lengthen the gulf of execution. High frequency of this pattern on only one axis for a given UI condition also might indicate ergonomic problems but could also point to difficulties in visualizing rotations for a given axis.

To quantify unintended/redone rotation patterns, we extracted frequency measures of these rotations as compared to all rotations. The frequency measures were calculated as the ratio of unintended rotations to total rotations, standardized to the various lengths of the unintended rotations. As such, the frequency measures varied from 0 to 100%. Note that our unintended/intended metric did not measure the user's deviation from an optimal rotation sequence; rather, it used the user's own actions to indicate when their UI action deviated from an intended rotation, was recognized and was redone.

7.2. Predictions for Patterns of Unintended/Intended Rotations

Overall, the unintended/intended rotation patterns constituted 13.370% of all rotations, with no significant differences by UI or SA. However, we anticipated differential results for each of the three rotation patterns. Based on our results from Section 5, we believed that persons in TG, with enhanced RBI knowledge and multi-sensory experiences, had well-developed mental representations of UI and task. We further thought that TS users had well-developed mental representations of UI. In terms of unintended/intended rotation patterns, we expected

(1) M participants, with less rich task representations, should have less well-formed task intentions. Unintended/intended rotations at the end of UI inputs point to limitations in task intention as well, during the evaluative Stage 7 of the Norman model. M users, with less detailed mental representations of the task and less clear task intention, should have a higher occurrence, compared to TG users.

(2) TS and TG users, with enriched mental representations of UI should find it easier to plan UI actions. If unintended rotations along the same axis related to actions at Stage 3 of the Norman model, M users should have a higher occurrence than either TS or TG.

(3) On the other hand, TS users faced potential ergonomic challenges (e.g., "fat finger," difficulty in finding the touch "hot spots," and difficulties with the line-of-sight axis, as predicted by the Instrumental Interaction model). TS participants might have had a higher percentage occurrence of unintended rotations along the same axis during the Norman model Stage 4. Similarly, TG users may have faced ergonomic

challenges in manipulating the free moving cube in a constrained space and might also have a high occurrence of unintended rotations on the same axis.

(4) Unintended/intended rotations by axis could be due to limitations in mental representations at any stage of the Norman model, making it difficult to make predictions relating to UI. However, one group should show significantly higher percentage of unintended rotations by axis. Low SA persons *should* have less rich mental representations of task than high SA individuals. While we found no evidence of a significant SA effect in Section 5 on our measure of representation of task (preference for vertical axis rotations), we expected to see higher percentage of unintended rotations by axis for low SA users.

7.3. Analyses of Unintended/Intended Rotations at the End of Inputs

Unintended/intended rotations at the end of inputs occurred when a user made a rotation, undid it and then gave their answer to the question of *same* or *different*. Overall, rotations in this pattern accounted for 1.902% of total rotations There was a significant effect of UI on the occurrence of this pattern ($F\,(2{,}60) = 13.220, p < 0.001$). M users exhibited this pattern at a significantly higher rate than TS or TG users (means M = 3.892%, TS = 0.727%, TG = 0.612%). Two Tukey pairwise comparisons by UI were significant: M vs. TS and M vs. TG (mean difference = 3.165, $p < 0.001$, mean difference = 3.280, $p < 0.001$, respectively).

We performed additional pairwise analyses to determine if there were axis and UI combinations that manifested this pattern more frequently than the others.[15] We found no significant effects for the TS nor TG conditions. However, paired comparisons by axis for the M condition indicated significantly more unintended/intended axis at the end of the inputs for the line-of-sight axes as compared to the horizontal axis (means horizontal = 0.590%, line of sight = 2.206%, $t = 3.389$, $df = 23$, $p < 0.001$). We interpreted this result to suggest that the M users had higher levels of uncertainty at the end of the input stream when making line-of-sight rotations.

We speculated that unintended/intended rotations at the end of the input stream indicated less rich representation of task. In our prior analyses, we had found indication that the M users had less rich representations of task, compared to the TG conditions. The analyses of unintended/intended rotations at the end of the stream pointed to the same result: The M participants had less rich representations of task. The new analyses indicated that M users were specifically challenged by rotations around the line-of-sight axis. Perhaps the lack of multi-sensory inputs, especially as related to the third dimension of the line-of-sight axis limited the representation of task for M.

7.4. Unintended/Intended Rotation on the Same Axis

Unintended/intended rotations on the same axis occurred when the user selected the intended axis and rotated the cube in an unintended direction. Once the user undid the unintended direction, they then rotated on the same axis in the intended direction. The overall percentage of rotations of this pattern was 4.862%. There was a significant effect of UI on the occurrence of this pattern (F (2,60) = 5.230, $p < 0.008$). TS and TG exhibited this pattern at a higher rate than M users (means M = 2.436%, TS = 5.915%, TG = 7.143%). We found significant Tukey mean differences for TS and TG compared to M (mean difference = 3.479, $p < 0.042$, mean difference = 4.487, $p < 0.012$, respectively).

We performed additional analyses to determine if there were specific axis and UI combinations that showed the unintended rotations by direction pattern. The means are shown in Table V. Using paired comparisons, we found no significant differences

[15] The additional paired comparisons were analyzed with paired t-tests.

Table V. Percentage Unintended/Intended Rotations Along the Same Axis by UI Condition

	Horizontal	Vertical	Line of sight
Mouse	0.4727%	0.695%	1.269%
Touchscreen	4.171%	1.448%	0.296%
Tangible	2.599%	2.973%	1.352%

by axes for either M or TG. For TS, two paired comparisons between axes pairs were significantly different, horizontal vs. vertical and horizontal vs. line of sight ($t = 1.932$, $df = 21$, $p < 0.033$, $t = 3.050$, $df = 21$, $p < 0.003$, respectively). In our analyses of Section 5, we found significantly higher rates of rotations around the horizontal axes for TS, as compared to TG. We found significantly more unintended rotations around horizontal as compared to vertical and line of sight for TS.

From the analyses of unintended analyses on the same axis in combination with our findings of Section 5 and of the patterns of unintended/intended rotations at the end of the input stream, we concluded that TS users ran into ergonomic difficulties (Stage 4 of the Norman model) rather than limitations of representations of UI. Both Instrumental Interaction Model and RBI analyses of the line-of-sight axis for TS users pointed to potential difficulties in forming and executing UI action sequences that involved the line-of-sight axis. Our findings in the Section 5 suggested that TS users had rich representation of UI and *knew how* to use the interaction to achieve their task intentions (Stages 3 and 4 of Norman model). The current finding that TS users made more unintended/intended rotations around the horizontal axis than the other two suggests some ergonomic challenges were in play for this UI condition. Perhaps TS users were simply avoiding the line-of-sight axis and, to a lesser extent, the vertical axis.

7.5. Unintended/Intended Rotations onto Another Axis

Overall, 6.489% of all rotations were involved in unintended/intended rotations onto another axis. There were no main effects by UI. Additional analyses pairwise by axis, within UI conditions were not significant. There was an overall main effect of SA (F (1,61) = 4.811, $p < 0.032$; high SA = 5.067%, low SA = 8.650%). These results indicated that low SA users had more limited mental representations of task than high SA persons, as suggested by hypothesis H3.

7.6. Summary of Analyses of Patterns of Unintended/Intended Rotations

We inspected patterns of unintended/intended rotations. Our goal was to better understand our results from Section 5, unfortunately our secondary findings were very limited. The measures of unintended/intended rotation patterns did not cleanly separate mental representation of task and UI from ergonomics and/or epistemic behaviors. Given that proviso, we saw some validation of our finding that the mental representations of task for M users may be more limited as compared to TG. Our analyses of unintended rotations on the same axis did not disconfirm that TS users had rich mental representations of UI. Thirdly, we saw some confirmation that our metric of SA was in fact related to richness of representation of task.

In terms of the validity of our measures from Section 5, the re-analyses of patterns of unintended rotations provided some interesting insights. Our original metric of richness of mental representation of task (choice of axis of rotation) and our alternative metric of unintended rotations at the end of inputs yielded similar results. Our original metric of richness of representation of UI, rotations for the one-rotation pattern, did not have a clear correlate in the alternate analysis. The alternative measure, unintended rotations onto another axis, confirmed the validity of the SA tests used to assess SA if not our Section 5 measure of mental representation of task.

8. LIMITATIONS AND OPEN QUESTIONS

Our work, couched in the Norman [1986] model, focused on ways that user knowledge, multi-sensory experiences, and corresponding mental representations distinguish WIMP and post-WIMP UIs. Many have argued that post-WIMP UIs offer opportunities to integrate real-world knowledge into mental representations of the UI. Post-WIMP UIs which offer multi-sensory experiences have the further advantage of enhancing users' task knowledge. However, as Jacob et al. [2008] reminds us, these advantages might come with costs, particularly relating to ergonomics.

Our driving theoretical position has been that the nature of the UI impacts the mental representations of both UI and task. Our study of these assertions left a number of open questions:

(1) A post-WIMP UI is specialized to the task. There are many possible UIs that might have at least some sense of being post-WIMP for a given task, as the touchscreen condition was in our study. Did we select the best representative post-WIMP UI for the CCT? For that matter, did we select the best representative of WIMP UIs for comparison? Our analyses, using the Instrumental Interaction Model, suggested that there were clear distances between the UIs on the continuum of WIMP to post-WIMP. Whether those differences were significant is open for discussion.
(2) We chose an object manipulation task because many post-WIMP UIs are designed for tasks with significant object manipulation. However, was the CCT really representative of tasks that come into use in post-WIMP UIs? The CCT is known to be sensitive in some situations to SA. Would it have been more instructive, in the context of this work to select a task that is less sensitive to SA?
(3) The CCT is a specific form of object rotation task, involving only symmetric objects. It is conveniently tied to both real-world knowledge of objects and easy to control in an interactive form using multiple input modalities. Given the nature of objects in the CCT, it is easy to envision corresponding post-WIMP UIs. Should we have selected a task that offered more significant challenges in constructing purpose-built UIs, a limitation alluded to by Jacob et al. [2008]?
(4) Identifying dependent variables which measure aspects of mental representations is notoriously hard and made harder here as our dependent measures were by necessity tied to the task. As such, our measures do not easily transfer to other tasks outside of the CCT. Were there other and better ways to operationalize our concepts of richness of mental representations? Could we have identified measures that operationalized gulfs of execution and evaluation on a UI by UI and task by task basis?
(5) Our research was clearly theoretical in nature. Yet ultimately, work in HCI should be tied to practical design principles. We were able to demonstrate *why* post-WIMP UIs may be more useful for some tasks than other types of interaction styles. How does a designer translate our findings into quality designs?
(6) We assumed that the mental representations of UI and task were operative throughout the Norman model. Further research focusing on this perspective of the Norman model and its relationships to post-WIMP UIs would be of use. For example, Svendsen [1991] suggested that there might be many paths to enhanced mental representations. Multi-sensory inputs or engagement of RBI knowledge are possible strategies, but user reflection and planning may be another pathway; were these strategies in play in our study?

9. CONCLUSIONS

As we suggested in the Introduction, HCI happily links ideas from multiple disciplines. The proliferation of post-WIMP UIs offers new chances to use and enhance older theories. In this work, we attempted to both drive our research from theory and to use our results to enhance the theory. Past applications of the Norman model [Hutchins et al. 1986] have focused on the interplay between the UI and the mental representation of the UI. Yet little work involving the Norman model has focused on the influence of the UI on mental representation of task. Stated another way, little work has considered that within the Norman model, there may be feedback to Stage 2 from latter stages of the model. We believe that by adding additional focus on potential ways that the interface may impact the representation of task, we have enhanced the Norman model itself.

In the end, as Norman implied decades ago, understanding the details of human cognitive structures involved in interaction were key to explaining observed differences in performance. Over the last 30 years, the Norman model has been useful in explaining performance differences in UIs. We suggest that the Norman model, including the cognitive aspects of interaction, is a good framework to study post-WIMP UIs and their underlying cognitive processes.

ACKNOWLEDGMENTS

The authors would like to thank Dr Robert Jacob for his continued insight into RBI, its past, and its future. We would also like to thank Julie Barnes, Brianna Tomlinson, Darren Guinness, Scott Kelley, Tom Donahue, Martez Mott, Alyssa Tomlinson, Justin Gilkey, Randall Littlejohn, John Strobel, Alan Jaffee, Jeremy Athy, Lisa Weihl, Rachelle Hippler, Dina Battaglia, and the BGSU Joint Cognitive Forum. Finally, we would like to thank our editors and reviewers. Their inputs were invaluable to us as we developed and finalized our ideas.

REFERENCES

Christian Battista and Michael Peters. 2010. Ecological aspects of mental rotation around the vertical and horizontal axis. *J. Individ. Differ.* 31, 2 (2010), 110–113.

Michel Beaudouin-Lafon. 2000. Instrumental interaction: An interaction model for designing post-WIMP user interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 446–453.

Charles E. Bethell-Fox and Roger N. Shepard. 1988. Mental rotation: Effects of stimulus complexity and familiarity. *J. Exp. Psychol. Hum. Percept. Perform.* 14, 1 (1988), 12–23.

John B. Carroll. 1993. *Human Cognitive Abilities: A Survey of Factor-Analytic Studies*. Cambridge University Press.

William G. Chase and Herbert A. Simon. 1973. Perception in chess. *Cogn. Psychol.* 4, 1 (1973), 55–81.

Michelene T. H. Chi, Paul J. Feltovich, and Robert Glaser. 1981. Categorization and representation of physics problems by experts and novices. *Cogn. Sci.* 5, 2 (1981), 121–152.

Lynn A. Cooper and Roger N. Shepard. 1975. Mental transformation in the identification of left and right hands. *J. Exp. Psychol. Hum. Percept. Perform.* 1, 1 (1975), 48.

Andrea diSessa. 1986. Models of computation. In *User Centered System Design: New Perspectives on Human-Computer Interaction*, D. A. Norman and S. W. Draper (Eds.). Lawrence Erlbaum, Hillsdale, NJ, 201–218.

Thomas J. Donahue, G. Michael Poor, Martez E. Mott, Laura Marie Leventhal, Guy Zimmerman, and Dale Klopfer. 2013. On interface closeness and problem solving. In *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction*. ACM, 139–146.

Randolph D. Easton, Anthony J. Greene, and Kavitha Srinivas. 1997a. Transfer between vision and haptics: Memory for 2-d patterns and 3-d objects. *Psychol. Bull. Rev.* 4, 3 (1997), 403–410.

Randolph D. Easton, Kavitha Srinivas, and Anthony J. Greene. 1997b. Do vision and haptics share common representations? Implicit and explicit memory within and between modalities. *J. Psychol. Appl. Learn. Mem. Cognit.* 23, 1 (1997), 153.

R. B. Ekstrom, John W. French, Harry H. Harman, and Diran Derman. 1976. *Manual for Kit of Factor-Referenced Cognitive Tests*. Educational Testing Services, Princeton, NJ.

Ronald A. Finke. 1980. Levels of equivalence in imagery and perception. *Psychol. Rev.* 87, 2 (1980), 113–132.

Alinda Friedman and David Lawrence Hall. 1996. The importance of being upright: Use of environmental and viewer-centered reference frames in shape discriminations of novel three-dimensional objects. *Mem. Cognit.* 24, 3 (1996), 285–295.

Jennie J. Gallimore and Michael E. Brown. 1993. Visualization of 3-d computer-aided design objects. *Int. J. Hum.-Comput. Int.* 5, 4 (1993), 361–382.

Isabel Gauthier, William G. Hayward, Michael J. Tarr, Adam W. Anderson, Pawel Skudlarski, and John C. Gore. 2002. BOLD activity during mental rotation and viewpoint-dependent object recognition. *Neuron* 34, 1 (2002), 161–171.

James Jerome Gibson. 1979. 1986. *The Ecological Approach to Visual Perception* (1979). Houghton Mifflin, Boston.

Rachelle Hippler, Dale Klopfer, Laura Leventhal, G. Michael Poor, Brandi Klein, and Samuel D. Jaffee. 2011. More than speed? An empirical study of touchscreens and body awareness on an object manipulation task. *HCI International 2011*. Orlando, FL, 33–42.

James Hollan, Edwin Hutchins, and David Kirsh. 2000. Distributed cognition: Toward a new foundation for human-computer interaction research. *ACM Trans. Comput.-Hum. Interact.* 7, 2 (June 2000), 174–196. DOI:http://dx.doi.org/10.1145/353485.353487

Michael Horn. 2012. Topcodes: Tangible Object Placement Codes. (2012). Retrieved January 15, 2012 from http://users.eecs.northwestern.edu/∼mhorn/topcodes/.

Edwin L. Hutchins, James D. Hollan, and Donald A. Norman. 1986. Direct manipulation interfaces. In *User Centered System Design: New Perspectives on Human-Computer Interaction*. D. A. Norman and S. W. Draper (Eds.). ACM, 87–125.

Robert K. Jacob, Audrey Girouard, Leanne M. Hirshfield, Michael S. Horn, Orit Shaer, Erin Treacy Solovey, and Jamie Zigelbaum. 2008. Reality-based interaction: A framework for post-WIMP interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 201–210.

Samuel D. Jaffee, Dena Battaglia, and Dale S. Klopfer. 2010. Tipping cubes: The effect of project on performance and strategy in a mental rotation task in a virtual environment. *Poster Session at the Annual Meeting of the Midwestern Psychological Association, Chicago, IL* (May 2010).

Samuel D. Jaffee, Laura M. Leventhal, Brandi A. Klein, and T. J. Donahue. 2013. Once again comparing cube rotations around axes inclined relative to the environment or to the cube: Shiffrar and Shepard (1991) revisited. *Paper Presented at the Annual Meeting of the Midwestern Psychological Association, Chicago, IL* (May 2013).

Marcel A. Just and Patricia A. Carpenter. 1985. Cognitive coordinate systems: Accounts of mental rotation and individual differences in spatial ability. *Psychol. Rev.* 92, 2 (1985), 137.

D. Kirsh and P. Maglio. 1994. On distinguishing epistemic from pragmatic action. *Cogn. Sci.* 18 (1994), 513–549.

Dale Klopfer, Jeremy Athy, and Laura Leventhal. 2007. Working memory: Just and Carpenter (1985) revisited. In *Proceedings of the 48th Annual Meeting of the Psychonomic Society*. ACM, 251–258.

Heidi Lam. 2008. A framework of interaction costs in information visualization. *Trans. Vis. Comput. Graphics*, 6 (2008), 1149–1156.

Seungyon Lee and Shumin Zhai. 2009. The performance of touch screen soft buttons. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 309–318.

David F. Lohman. 1996. Spatial ability and g. *Hum. Abilities., Nature Meas.* 97 (1996), 116.

David Marr and Herbert Keith Nishihara. 1978. Representation and recognition of the spatial organization of three dimensional shapes. *Proc. R. Soc. London* 200 (1978), 269–294.

Roxana Moreno and Richard Mayer. 2007. Interactive multimodal learning environments. *Educ. Psychol. Rev.* 19, 3 (2007), 309–326.

Tomer Moscovich. 2009. Contact area interaction with sliding widgets. In *Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology*. 13–22.

Martez Mott, Thomas Donahue, G. Michael Poor, and Laura Leventhal. 2012. Leveraging motor learning for a tangible password system. *Extended Abstracts on Human Factors in Computing Systems (CHI'12)*. ACM, 2597–2602.

Allen Newell and Herbert Alexander Simon. 1972. *Human Problem Solving*. Vol. 104. Prentice-Hall Englewood Cliffs, NJ.

Donald A. Norman. 1986. Cognitive engineering. In *User centered system design: New perspectives on human-computer interaction*. Donald A. Norman and Stephen W. Draper (Eds.). Lawrence Erlbaum Associates, Hillsdale, NJ, 31–61.

Donald A. Norman. 1990. The design of everyday things. (1990). Doubleday, New York, NY.

Nancy Pennington. 1987. Stimulus structures and mental representation in expert comprehension of computer programs. *Cogn. Psychol.* 19 (1987), 295–341.

G. M. Poor, Brianna J. Tomlinson, Darren Guinness, Samuel D. Jaffee, Laura M. Leventhal, Guy Zimmerman, and Dale S. Klopfer. 2013. Tangible or gestural: comparing tangible vs. Kinect^{TM} interactions with an object manipulation task. In *7th International Conference on Tangible, Embedded and Embodied Interaction*. Barcelona, Spain.

G. Michael Poor. 2008. *The effects of varying levels of reality-based interaction styles on a subject's ability to perform a 3d construction task*. Unpublished Doctoral Dissertation, Tufts University, MA.

G. Michael Poor, Laura Marie Leventhal, Scott Kelley, Jordan Ringenberg, and Samuel D. Jaffee. 2011. Thought cubes: Exploring the use of an inexpensive brain-computer interface on a mental rotation task. In *Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS'11)*. ACM New York, NY, 291–292.

G. Michael Poor, Guy Zimmerman, Dale S. fer, Samuel D. Jaffee, Laura Marie Leventhal, and Julie Barnes. 2013a. Mobility matters: Identifying cognitive demands that are sensitive to orientation. In *Human-Computer Interaction–INTERACT 2013*. Springer, 193–210.

G. Michael Poor, Guy W. Zimmerman, Dale S. Klopfer, Samuel D. Jaffee, Laura Marie Leventhal, and Julie Barnes. 2013b. Mobility matters: Identifying cognitive demands that are sensitive to orientation. (2013), 193–210. DOI:http://dx.doi.org/10.1007/978-3-642-40483-2_14

Andrew Sears and Ben Shneiderman. 1991. High precision touchscreens: Design strategies and comparisons with a mouse. *Int. J. Man-Mach. Stud.* 34, 4 (1991), 593–613.

Roger N. Shepard. 1984. Ecological constraints on internal representation: Resonant kinematics of perceiving, imagining, thinking, and dreaming. *Psychol. Rev.* 91, 4 (1984), 417–447.

Roger N. Shepard and Jacqueline Metzler. 1971. Mental rotation of three-dimensional objects. *Science* 171, 3972 (1971), 701–703.

Margaret M. Shiffrar and Roger N. Shepard. 1991. Comparison of cube rotations around axes inclined relative to the environment or to the cube. *J. Exp. Psychol. Hum. Percept. Perform.* 17, 1 (1991), 44–54.

Benjamin Shneiderman. 1982. The future of interactive systems and the emergence of direct manipulation. *Behav. Inf. Technol.* 1, 3 (1982), 237–256.

Gunnvald B. Svendsen. 1991. The influence of interface style on problem solving. *Int. J. Man-Mach. Stud.* 35, 3 (1991), 379–397.

Louis Leon Thurstone. 1938. Primary mental abilities. (1938). The University of Chicago Press, Chicago.

Brygg Ullmer, Hiroshi Ishii, and Robert J. K. Jacob. 2005. Token+ constraint systems for tangible interaction with digital information. *ACM Trans. Comput.-Hum. Interact. (TOCHI)* 12, 1 (2005), 81–118.

Manuel Vidal, Alexandre Lehmann, and Heinrich H. Bülthoff. 2009. A multisensory approach to spatial updating: The case of mental rotations. *Exp. Brain Res.* 197, 1 (2009), 59–68.

Florian Waszak, Knut Drewing, and Rainer Mausfeld. 2005. Viewer-external frames of reference in the mental transformation of 3-d objects. *Percept. Psychophys.* 67, 7 (2005), 1269–1279.

Mark Wexler, Stephen M. Kosslyn, and Alain Berthoz. 1998. Motor processes in mental rotation. *Cognition* 68, 1 (1998), 77–94.

Gunnar Wiedenbauer, Juliane Schmid, and Petra Jansen-Osmann. 2007. Manual training of mental rotation. *Eur. J. Cognit. Psychol.* 19, 1 (2007), 17–36.

Andrew T. Woods, Allison Moore, and Fiona N. NewellÙ. 2008. Canonical views in haptic object perception. *Perception* 37 (2008), 1867–1878.

Maryjane Wraga, Sarah H. Creem, and Dennis R. Proffitt. 2000. Updating displays after imagined object and viewer rotations. *J. Exp. Psychol. Learn. Mem. Cognit.* 26, 1 (2000), 151.

Yan Zhang. 2012. The impact of task complexity on peoples mental models of MedlinePlus. *Inf. Process. Manag.* 48, 1 (2012), 107–119.

Received January 2015; revised June 2016; accepted August 2016