Coordinating Actions and Models

Daniel L. Schwartz

School of Education
Stanford University
Stanford, CA 94305-3096
daniel.schwartz@stanford.edu

Abstract

Physical action can improve people's ability to complete analog inferences about distal events. For example, if without vision, people pull a string that turns a spool, this movement improves people's ability to imagine the rotation of a block on top of the spool. Similarly, tilting a glass can help people imagine the behavior of water in that glass, even if their eyes are closed and there is no actual water. The dominant model for this facilitation effect is that people map distance information from their movements into their mental updates through feed forward or feedback mechanisms. The current paper offers new evidence that people use the timing of their movement rather than its distance to drive their qualitative reasoning about the effects of action. This evidence suggests four constraints on the design of qualitative reasoning engines that coordinate with physical action and that attempt to maintain psychological fidelity.

From catapults to hand drills, a distinctively human talent is the construction and use of multipart tools. The purpose of this paper is to explore a competence that may be responsible for this talent, namely, people's ability to model the environmental consequences of their actions. Actions constitute an important component of people's qualitative understanding. For example, after several days of wearing inverting spectacles people eventually see the world up right again. Kohler (1964), who studied the effects of prismatic glasses, noted that people's recovery of simple abilities, like drinking a cup of water, often preceded their ability to see the object of action as upright. "We stand in the visual world not only with our eyes but also with our hands, feet, and shoulders. It is just for this reason that anyone who wants to *see* correctly must first be able to *manipulate* correctly" (p. 163). Although there are limits to the claim; for example, color perception may not depend on manipulation, it seems clear that action takes a position of prominence in people's grasp of physical reality. Because of this prominence, we propose that people have representations that respond to action. These representations have developed to support physically situated cognition and tool use, and they keep people's knowledge of the material world in concert with their physical activity.

The representations we propose are tied to perceptual activity, and therefore they include sufficient metric information to support relatively precise spatial anticipation. At the same time, they include qualitative information that determines how objects and limbs interact with one another using non-Newtonian representations of force and movement. In earlier work (Schwartz & Black, 1996a), we developed an object-oriented model, called a *depictive model*, for how people might complete perceptually precise, qualitative inferences that use analog representations of space and force. In this paper we focus on the connection between people's models and their actions. Our work has not sufficiently progressed to make a computational model of this connection without too many unconstrained guesses. Therefore, we focus on evidence that identifies four core capacities of an eventual model: (1) Changing how a representation responds to action based on learning; (2) Updating to the timing of action; (3) Converting actions of one form into representational updates of another; and, (4) coupling different models to the same action. In the following, we begin with a general overview of our hypothesis about how people coordinate qualitative inferences with action. To do this, we compare our hypothesis to the "mapping theories" that dominate current thinking. Afterwards, we return to the four capacities.

Timing-Responsive Representations

The representations that we propose are timing-responsive representations (TRs). TRs model the distal environment and update according to the timing signals that mark change. For example, we showed people two glasses of identical height with the same level of water (Schwartz & Black, 1999). The glasses had different diameters. We asked if the two glasses would pour at the same angle or if one would pour sooner. Nearly everyone answered incorrectly. Yet, when people tilted each glass in turn, without vision until they imagined the water just touching the rim, nearly everyone correctly tilted the narrow glass farther than the wide one. People were extremely accurate, even when there was no actual water and they had to represent its presence. Our explanation is that people represented the water as an analog image and their imagery was timing responsive. Without the timing of action, their water image was difficult to transform, and therefore, people made static comparisons between the glasses using discrete quantitative reasoning (e.g., "the two glasses are the same height and therefore…"). In contrast, when people tilted the glasses, the timing of their movements drove the update of their water image.

The TR metaphor is simple: each timing signal is a small catalyst (or neural firing or computer interrupt) that causes a TR to transform. We use the expression "timing responsive" instead of "change responsive," because we believe that a TR can respond similarly to rates of change though the content of the changes may differ. For example, if people sidestep in a circle to face a target or swivel at the waist, they will update their mental map of their relative heading the same, although one motion is discrete and the other continuous. By being responsive to higher-order timing information, TRs permit cross-modal activation. For example, people's representation of a subway car's relative distance might update to an approaching rumble as well as to their own steps down the tracks. We refer to the timing signals generated by self-movement as though they were discrete signals of varying frequency and strength with the understanding that they may take many different forms (e.g., waves of varying amplitude, gradients, etc.).

The TR Hypothesis versus Mapping Theories

TRs may help explain findings that show action can improve people's ability to imagine and anticipate physical change (e.g., Rieser, Garing, & Young, 1994; Simons & Wang, 1998). For example, if people try to imagine a block rotation without vision, manually turning the block increases the speed they can imagine the rotation (Schwartz & Holton, 2000). The TR hypothesis proposes that actions produce strong timing signals that cause the block image to update; for example, each signal might cause one update of degree *x*. This explanation differs from several variants of mapping theory that also explain how action could facilitate cognitive updates. "Mapping theory" is a general label for those theories that assume people match representations of their proximal action (e.g., a hand movement) to their representation of the distal situation (e.g., the block).

One instance of mapping theory comes from feed-forward models. Feed-forward models propose that the planning component of an action facilitates a mental update. Imagine that people plan to move a block an inch to the left with their hand. Their motor plan specifies spatial information that they can map to their block representation to anticipate its subsequent position and appearance (an inch to the left). Another instance of mapping theory comes from feedback models. People feel the extent of their hand movement (an inch to the left), and they use this information to update their image of the block. Both of these models have merit, and as one might expect, there are hybrid models that include feed-forward and feedback mechanisms.

Mapping approaches, like the feed-forward and feedback models, typically have three characteristics: direct spatial mapping, non-concurrent updates, and the representation of movement. These three characteristics help illuminate what is unique about the TR hypothesis in contrast.

Depictive Models versus Direct Mappings. Mapping theories often presume a direct spatial mapping between a movement and an imagined update. Thus, a clockwise hand movement facilitates an imagined clockwise block rotation and interferes with a counter-clockwise one (Wexler & Klam, in press; Wohlschlåger & Wolschlåger, 1998). Yet, if actions always yielded spatially isomorphic representational changes, tool use would be nearly impossible. When people turn the steering wheel of their car, they would anticipate a barrel roll instead of a right turn. Moreover, in our research, we have found that spatial mapping can be violated and people still show facilitating effects of action on the imagination. For example, Figure 1 shows a block on a spool. When people pull the string, it improves their ability to imagine the rotation of the block (Schwartz & Holton, 2000). Notice that the motor plan and feedback specify a linear motion, while the imagery update is a rotation.

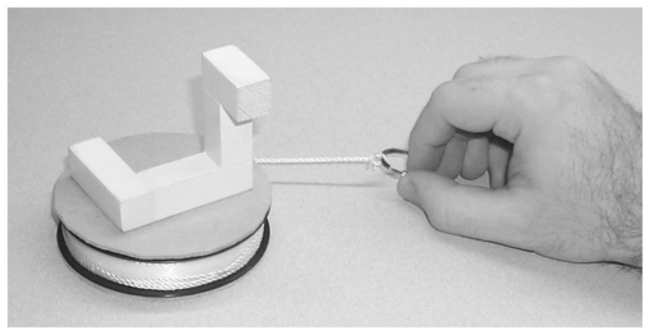

Figure 1.Pulling the string helps people imagine the block rotation

Spatial mapping could accommodate the spool example by allowing that people can insert a mental transformation matrix that converts the linear hand movement into an imagery rotation. With this amendment, action causes people to update their imagery, but the spatial content of the action does not determine the extent or direction of update. This content comes from people's transformation matrix, or as we prefer to call it, people's depictive model of the situation. The TR hypothesis goes further by assuming that depictive models do not require spatial input to model a spatial update. Mapping theories assume that actions generate specific spatial information that maps into a specific spatial update. The TR hypothesis assumes that actions can dramatically under-specify the trajectory of an update. Timing signals only trigger the update; it is the job of one's model to determine what update to complete. As a consequence, the timing generated by a repeated key press or a sound may facilitate an imagery rotation, if people have an appropriate model in mind (Holton, 2001).

It may seem strange to propose that timing can cause representations to change without specifying what change to make. Yet, it may be useful to begin with this minimal assumption about the informational content of action. It allows us to see how much we can load into qualitative models rather than direct environmental specification, and still maintain coordination between action and inference.

"Feed-During" versus Non-Concurrent Updating. A second characteristic of mapping theories is that representational updates occur before or after the action actually takes place, hence the names "feed-forward" and "feedback". Of course, feed-forward and feedback updates can occur throughout a relatively long motion. Regardless, information about a sub-movement within a long motion maps into a representation before or after that sub-movement takes place. These models are about the predecessors and consequents of motion, but not the dynamics of motion per se. TRs offer a different model that might be playfully described as "feed during." A timing signal causes a representation to change in real time.

Feed-forward, feedback, and feed-during models can all predict updating differences between action and no action. The feed-during property also predicts differences between types of action, like jumping and stepping to the same target (Schwartz & Williams, 2001). Stepping presumably generates more timing signals than jumping and therefore should cause more updates. We demonstrate this below.

Unrepresented Timing Signals versus Represented Movements. The third characteristic of mapping theory is that people match a representation of their movement to a representation of the situation. This means that people must represent their movement for it to have any effect on imagery. This again differs from the TR hypothesis, which assumes that material (unrepresented) timing signals during movement cause symbolic updates.

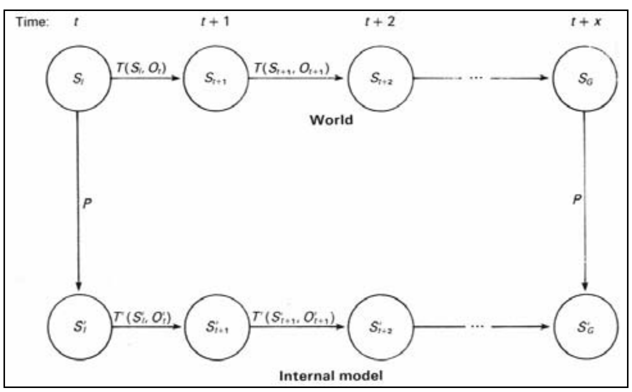

Figure 2. The Mapping Model according to Holland et al., 1985.

The conversion of physical reality into a symbolic form so it can affect processing is characteristic of many cognitive models that subscribe to the mapping theory. Holland et al. (1985) provide a schematic of the mapping theory in Figure 2. The vertical arrows represent people mapping the environment into symbolic representations through a process of recognition. Once converted, people complete a set of symbolic transformations, represented by the lower horizontal lines, and the world completes a set of physical transformations, represented by the upper lines. Sometime afterwards, people then try to recognize whether their mental transformations correspond to changes in the environment. In the mapping model, the physical world never causes mental transformations directly. The material world of causality and the syntactic world of symbols run in parallel. Shepard (1994), for example, has proposed that imagery evolved symbolic constraints that are isomorphs of physical constraints to help ensure imagery stays parallel to the environment during transformations. In contrast, for the TR hypothesis, one would need to add diagonal lines to Figure 2 so that material changes could directly regulate representational changes. We might call these diagonal lines the "direct to representation" component of the TR hypothesis.

The direct-to-representation component of the TR hypothesis may help explain effects in addition to those of timing. In our research, we have found that people directly rely on gravity to regulate their imagery and that they cannot represent it (Schwartz, 1999). We asked people to solve the pouring task with imagined water as described above, but with a small change. People held each glass sideways instead of upright. We told them to imagine that gravity was operating sideways (or that they were on their side and the glasses were upright). People were unable to represent gravity, and instead, their representation responded to real gravity so they could not complete the task accurately. People kept imagining that the water was pouring from the glasses once they began to tilt them. In general, it seems unnecessary to assume that people must represent gravity for it to control their representations. Gravity is ubiquitous, and therefore it is not necessary to represent it. Instead, representations can directly depend on gravity for their operation.

Similarly, timing is a ubiquitous aspect of action, and therefore, representations may have evolved to respond to it rather than represent it. To further clarify this claim we can compare it to temporal mapping models (as opposed to the spatial-only mapping models above). A large body of research shows that people use timing to regulate navigation, reinforcement, music, and motor activity (Rosenbaum & Collyer, 1998). Most theories that explicitly consider how time influences cognition describe how organisms operate over a stored representation of time. The representations of time are often scalar values collected by counting the cycles of an internal oscillator (e.g., Wing & Kristofferson, 1973). The internal oscillator does not represent time. It generates a periodic change that marks time. But, when the periodic change is mapped into a storage variable, time becomes representational. For example, imagine the task of determining how far one has moved from a starting point. According to the common dead reckoning model of navigation (e.g., Gallistel, 1990), people store the duration of their overall movement by tallying the number of cycles completed by their internal oscillator. They also compute the velocity of their movement by tallying the duration it took to travel a sample distance (perhaps by mapping the length of their strides). Given these two pieces of information, people "deductively reckon" how far they have traveled by multiplying duration and velocity.

In the dead reckoning model, the estimate of elapsed time reflects the duration of an action, but the time course of a mental update does not. Computing position based on a long duration would take the same amount of time as computing position based on a short duration, assuming people multiply equally fast in each case. Although the model explicitly considers the importance of time, it maps real time into symbolic time before computing updates to a representation.

Temporal mapping does not predict that action would affect imagery (unless it invokes spatial mapping assumptions). Given that timing information must be represented to have an influence, one could simply represent duration (or wait a few seconds to store an estimate of time), and action would be superfluous. In contrast, the TR hypothesis directly predicts an effect of action, because TRs depend on a material timing signal to propel each update.

TR Summary

In summary, the TR model proposes that people have representations that respond to action. This responsiveness helps "time lock" (Edelman, 1992) symbolic updates with corresponding material changes. To accomplish this time locking, TRs do not represent time, but instead they allow the timing of action to drive their updates. For example, each timing signal might cause a representation to update one step. Because TRs do not represent time, they need "real-time" signals to drive their transformations. An excellent source of timing information comes from the physical actions people use to change their situation. Notably, the timing signals that arise from action do not necessarily specify how a representation should change (although they can be concomitant with relevant information). The same actions can cause different representational updates depending on the model people have in mind. This isolates the importance of action per se on people's depictive models, regardless of the spatial content the action may or may not convey. All told, the TR approach offers a highly flexible and adaptive mechanism for the coordination of material and symbolic activity.

The Four Capacities

The TR hypothesis is meant to explain how people draw inferences as a consequence of action. These inferences may be distinguished from more explicit, propositional forms of reasoning. The water pouring is one example in which people can draw an inference about a physical behavior during action, even though they do not know the correct answer explicitly. Another example, comes from research in which we asked people to reason about chains of gears (Schwartz & Black, 1996b); for example, if the gear on the farthest left of a five-gear chain turns clockwise, what will the gear on the farthest right do? At first, people rely on their hand movements to infer the relative movements of adjacent gears. But, after solving a few problems, they learn a parity rule (e.g., odd gears turn the same direction), and they stop using their hands. Evidently, people did not know the qualitative behaviors of the gears explicitly at first, and they needed to rely on their own actions to gain access to this knowledge. Our hypothesis is that their qualitative representations are timing responsive, and they need timing signals to manifest their knowledge. For TRs to support this level of implicit inference they need four capacities.

The Ability to Learn New Updates to Action

A future computational model must allow that people learn to execute new updates for a given action. People regularly change the coupling between their actions and their imagined updates; for example, when learning to use a bigger tennis racket, a heavier bowling ball, or a spool with a greater diameter. This phenomenon is called recalibration or adaptation. Figure 3 shows a turntable apparatus we have used to explore recalibration. People step in a circle on the turntable. As they step in one direction, the turntable rotates in the opposite direction. People have to do twice as much work to travel at their usual rate with respect to the environment. The subjective experience is the rotational equivalent of walking the wrong direction on a moving sidewalk. After a few minutes, people automatically recalibrate the coupling between their steps and the mental maps they use to monitor their position (Rieser, Pick, Ashmead, & Garing, 1995). For example, after recalibration, people might leave the turntable and try to turn one revolution without vision. People unwittingly turn too far. Although they know they are standing on stationary ground, they update their mental maps considering the progress they would have made on the moving turntable. Our assumption is that recalibration changes the amount people change their representation for each timing signal of action. (The timing signals do not change because these are not represented.) Among other things, this explains why people are unaware they have been recalibrated (which is why they turn too far). Because people do not represent time, they cannot store and reflect upon changes to the coupling between timing and movement.

Responsiveness to Material Timing Signals

The TR hypothesis proposes that people update to the material timing of action. This is a difficult claim to show empirically, because actions typically confound timing and spatial movement. The recalibration paradigm helps begin teasing apart the temporal and spatial aspects of action (Schwartz & Williams, 2001). According to the TR hypothesis, recalibration changes the amount that an image updates to the timing of an action. This implies that if people can complete a spatial task that requires a mental map but does not require on-going action, there should be a minimal effect of recalibration because there would be few timing signals to drive the update of the map. The turntable in Figure 3 helped test this prediction.

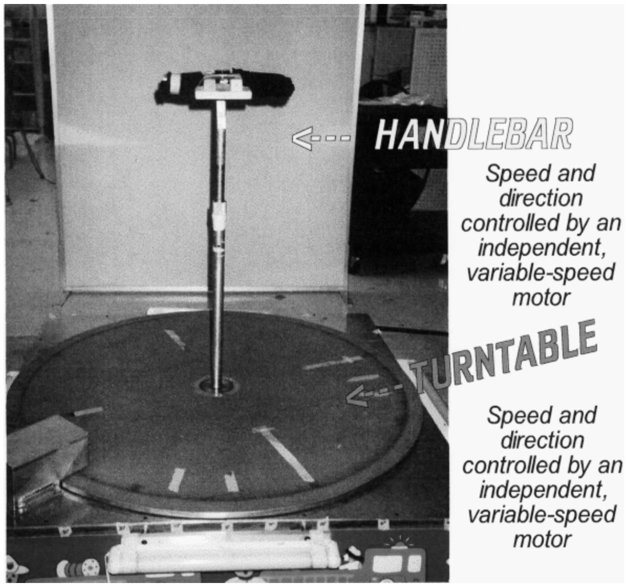

Figure 3. The turntable apparatus.

We asked people to complete a common set of tasks before and after they were recalibrated on the turntable. People memorized targets placed in a circle around them. Without vision or hearing, they oriented toward a named target using different motions. For the Leg conditions, people sidestepped in a circle moving slowly, normally, quickly, or they jumped to face the target. In the arm conditions, people held an outstretched arm rigidly in front of them. They swiveled at the waist to point to the target moving slowly, normally, or quickly. To parallel the ballistic case of jumping, they also raised their arm from their side to point directly at the target. The question of interest is how far they overshoot the targets after recalibration. Figure 4 plots how far people overshot before and after recalibration. After recalibration, when people stepped or swiveled, they overshot each target a greater distance than when jumped or directly pointed, and they showed greater effects of recalibration the more slowly they moved. The TR hypothesis provides an explanation.

For jumping and directly pointing, once people start to move, the use of their mental map is done and irreversible (e.g., they do not suddenly twist in mid-air during a jump). This allows us to determine whether planning the distance of the jump leads to recalibration effects. We also gave people a chance to adjust their position once they landed from their jump (not shown in the figure). This allows us to determine whether spatial feedback from the movement leads to the recalibration effect. We predicted there would be minimal effects of recalibration in these conditions, because people update to the timing signals of action not the distance, and jumping and directly pointing are ballistic movements that do not generate on-going timing signals. In contrast, for the stepping and swiveling conditions,

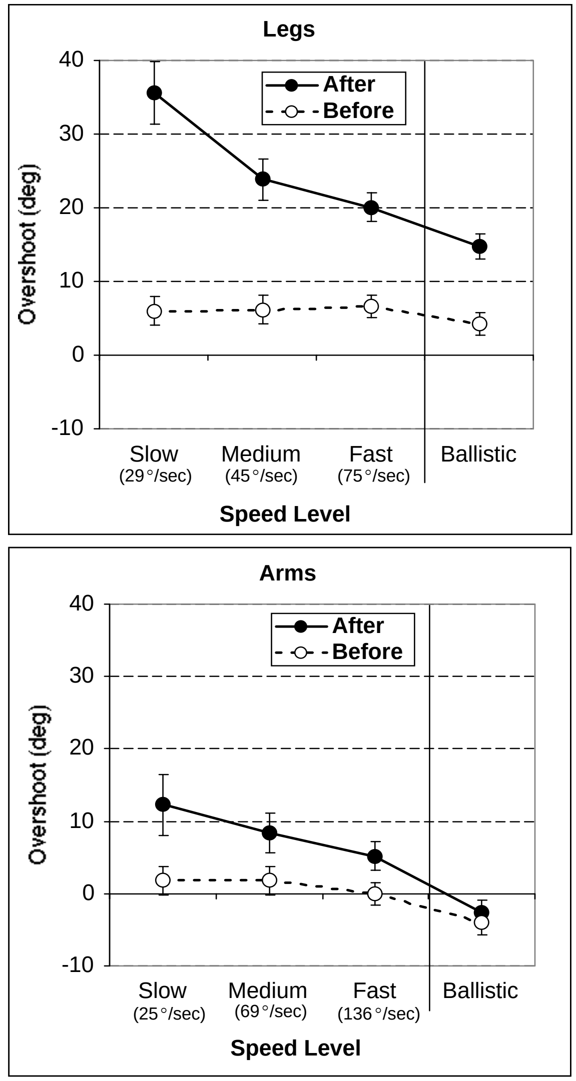

Figure 4. The effects of different movements and speeds on recalibration.

people generate timing signals that drive their on-going updates. The results showed that sidestepping and swiveling typically doubled the recalibration effect of jumping and pointing, and that people did not adjust their landing position after jumping. These results indicate that the spatial feed-forward before an action and the spatial feedback after an action are not fully responsible for the recalibration effect. In contrast, the results for the arm condition are nearly ideal for the TR hypothesis. Swiveling at the waist led people to point beyond the target on the posttest, even though they were accurate on the pretest and simply raising an arm to the target showed no effect of recalibration at all. Evidently, the inferences that people draw through action can be different from the deliberative inferences they draw before and after action.

The study also shows that changing the duration or number of hypothesized timing signals influences the recalibration effect. People sidestepped and swiveled at different speeds. Before recalibration, different speeds of movement had little effect. Presumably, the participants had previous experiences with turning at different rates and therefore had well-tuned TRs. After recalibration, people showed greater effects of recalibration the more slowly they moved. Analyses of the participants' actual movements showed that the speed and duration of their movements correlated with the size of the recalibration effect whereas the number or size of their steps did not. This suggests that people's spatial representations were updating with respect to temporal information rather than the distance information available in their movements.

A curious dissociation helps to show how the TR hypothesis simplifies computation. The dissociation was that people's actual rate of movement was unaffected by recalibration (e.g., when told to turn slowly at pre- and posttest they turned at the same speed). There was a .9 correlation between movement speeds on pretest and posttest. Yet, recalibration led people to imagine they were covering less distance than they actually were; they overshot on the posttest. Evidently, people managed to keep the same physical speed even though they thought they were moving more slowly!

This dissociation between movement production and movement perception is problematic for temporal mapping models, because they presume that people compute and represent speed both to monitor position and to produce movement. The TR hypothesis is simpler. It explains this dissociation by assuming that people use an internal oscillator to control their rate of stepping. People might use one cycle per step to move quickly, and three cycles per step to move slowly. However, people determine their position without computing or monitoring the speed. To update their representations, people do not need to represent time. People simply attend to their imagined position as it updates one step to each timing signal produced by their muscles. Speed information may be implicit in these representational updates, but people do not represent the durations needed to compute it. Consequently, there is no conflict between movement production and movement perception.

Converting One Action into Updates of Another

A third important capacity for a qualitative model that depends on action would be the ability to convert changes of one form into updates of another. This is particularly important for tool use, because tools typically convert motion from one form to another. The spool example provides one instance. People's linear motion of pulling a string facilitated their ability to imagine an object rotating on top of the spool. Another example comes from the water-pouring task. When people rotated the glass using their wrist, it facilitated their ability to imagine the water "sliding up the side of the glass."

In each of these cases, people are able to use a depictive model to convert a spatial action of one trajectory into a spatial update of another trajectory. The changes that occur through action, however, are not purely spatial. They typically involve work, as in the case of pulling the string on a bow. A strength of the TR hypothesis is that it allows that people can also convert the timing of work into spatial updates. (TRs depend on rates of change rather than distances of change.) This is important for problems involving force. The water-pouring task again provides useful evidence along this line (Schwartz, 1999). We asked people to tilt the glasses several times, imagining a different water level each time. In one case, people tilted the usual wide and thin glass. In the other case, people tilted the two glasses with a small weight affixed to their bases. In the later case, people increasingly under-tilted the glasses the farther they had to turn them (due to the lower imagined water level). Our explanation is that the weights caused an increasing torque the farther the glasses tilted toward the horizontal plane. People perceived an increase in the rate of work they applied to overcome the torque, and they updated their representations in concert with this increase in work. This led them to update their representations more quickly compared to the non-weighted condition, and they imagined the water reached the rim sooner. As usual, people were unaware of this effect, because they were not representing the increased work, they were simply responding to the increased number of timing signals indicating a change was occurring.

Modeling Different Changes for the Same Actions

The final capacity we consider is the ability to swap in different models to mediate the effects of action on representational updating. This is an important capacity for tool use, because the same action can yield different consequences depending on the tool at hand. One example comes from the block on the spool. If people believed the string would turn the block clockwise, it facilitated their ability to imagine a clockwise rotation of the block on the spool, and interfered with their ability to imagine a counter-clockwise rotation. Similarly, if they believed the string turned the spool counter-clockwise, it facilitated a counter-clockwise block rotation and interfered with a clockwise one. This occurred even though they committed the same movement in each case, and they could not actually see or feel the spool turning. Evidently, the effect of action depended on their model of the situation.

Water pouring provides a second example of the ability to change the models that respond to action. We asked people to imagine that the glasses had molasses instead of water. In this case, their TRs updated more slowly in response to their tilting, and they tilted too far. This occurred even though they explicitly believed that molasses would not change the angle of tilt necessary to get the liquid to the rim. As fits our general story, they were reasoning based on the timing of molasses rather than distance of movement.

Conclusion

To make further headway on the TR hypothesis, it seems necessary to develop a clearer account of the timing signals that drive representations. This may be difficult. We have argued that the effect of timing information depends on people's model, and timing signals can appear in different modalities and surface forms. At this point, rather than developing a premature account of timing signals, it may be better to explore the range of qualitative inferences the TR hypothesis might explain or be tested against.

We believe the range of phenomena is large because time is an inherent property of all change. Even people's symbolic thoughts take a time course. Therefore, TRs may be capable of responding to timing signals generated by changing representations, although these timing signals would presumably be weaker than those generated by action. For example, people can run internal simulations of external events and motor events (Decety, Jeannerod, & Problanc, 1989; Shepard & Cooper, 1986). Presumably, these mental simulations depend on the generation of timing signals to determine how fast things should change in one's depiction of the world (Schwartz & Black, 1996a). Although such timing signals would come from the rate of representational change, this does not mean that the timing information available to the simulation would be "represented time." A representation changes in real time. This makes it so the same TR can respond to internal and external sources of timing information, because both sources carry the same "real time."

References

Decety, J., Jeannerod, M., & Problanc, C. 1989. The timing of mentally represented actions. *Behavioural Brain Research:* 34, 35-42.

Edelman, G. M. 1992. *Bright air, brilliant fire: On the matter of the mind*. New York: Basic Books.

Gallistel, C. R. 1990. *The organization of learning.* Cambridge, MA: MIT Press.

Holland, J. H., Holyoak, K. J., Nisbett, R. E., & Thagard, P. R. 1986. *Induction: Process of inference, learning, and discovery.* Cambridge, MA: MIT Press.

Holton, D. 2001. *The influence of action-generated timing cues on mental imagery*. Unpublished Masters Thesis. Vanderbilt University, Nashville, TN.

Kohler, I. 1964. The formation and transformation of the perceptual world. *Psychological Issues*, Vol. 3 (4), Monograph 12. Translated by Harry Fiss.

Rieser, J. J., Garing, A. E., & Young, M. E. 1994. Imagery, action, and young children's spatial orientation. It's not being there that counts, it's what one has in mind. *Child Development*, 65, 1262-1278.

Rieser, J. J., Pick, H. L., Jr., Ashmead, D. H., & Garing, A. E. 1995. The calibration of human locomotion and models of perceptual-motor organization. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 480-497.

Rosenbaum, D. A., & Collyer, C. E. (Eds.) 1998. *Timing of behavior: Neural, psychological, and computational perspectives*. Cambridge, MA: MIT Press.

Schwartz, D. L. (1999). Physical imagery: Kinematic versus dynamic models. *Cognitive Psychology*, 38, 433-464.

Schwartz, D. L. & Black, J. B. 1996a. Analog imagery in mental model reasoning: Depictive models. *Cognitive Psychology*, 30, 154-219.

Schwartz, D. L. & Black, J. B. 1996b. Shuttling between depictive models and abstract rules: Induction and fallback. *Cognitive Science*, 20, 457-497.

Schwartz, D. L., & Black, T. 1999. Inferences through imagined actions: Knowing by simulated doing. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 25, 1-21.

Schwartz, D. L., & Holton, D. 2000. Tool use and the effect of action on the imagination. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 26, 1655-1665.

Schwartz, D. L., & Williams, D. B. 2001. *Real-Time Cognition: Coordinating Symbolic and Physical Activity*. Manuscript under review.

Shepard, R. N. 1994. Perceptual-cognitive universals as reflections of the world. *Psychonomic Bulletin and Review*, 1, 2-28.

Shepard, R. N. & Cooper, L. A. (Eds.). 1986. *Mental images and their transformation*. Cambridge, MA: The MIT Press.

Simons, D. J. & Wang, R. F. 1998. Perceiving real-world viewpoint changes. *Psychological Science*, 9, 315-320.

Wexler, M. & Klam, F. 2001. Movement prediction and movement production. *Journal of Experimental Psychology: Human Performance and Perception*. Forthcoming.

Wing, A. M., & Kristofferson, A. B. 1973. Response delays and the timing of discrete motor responses. *Perception & Psychophysics*, 14, 5-12.

Wohlschläger, A., & Wolschläger, A. 1998. Mental and manual rotation. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 397-412.