Viewing and performing actions can change what you see

Alexia Toskos Dils (atoskos@stanford.edu)
Stephen J. Flusberg (sflus@stanford.edu)
Lera Boroditsky (lera@stanford.edu)
Stanford University, Department of Psychology
Jordan Hall, 450 Serra Mall, Building 420, Stanford, CA 94305 USA

Abstract

Previous research has demonstrated a tight link between object perception and action: viewing an object primes the action needed to interact with it, while priming an action can affect the speed and accuracy with which we perceive the object. However, it is not yet known whether motor information can qualitatively change what object we actually perceive. We investigated this issue by having participants view or perform an action before viewing an ambiguous object. Results showed that viewing an action (a picture of a hand displaying a power or precision grasp) biased participants to interpret the ambiguous object as congruent with the action prime (Experiments 1 and 2). Conversely, performing an action (moving small or large balls from one tray to another) biased participants to interpret the object as incongruent with the motor action. Together, these results suggest viewing and performing actions can actually change what we see.

Keywords: Object perception**;** Action; Embodiment

Background

Can our actions influence how we perceive the world and affect the very contents of our visual awareness? Though perception and action have traditionally been studied independently in the cognitive sciences, in our everyday experience of the world they are dynamically linked. For example, many of the objects we look at are also the objects we grasp and manipulate. More generally, our movements and actions in the environment alter what perceptual information we have access to, and these changes in perceptual stimulation consequently influence how we traverse our surroundings and what actions we choose to take. For reasons such as these, ecologically orientated psychologists have argued that we perceive the world in terms of how it affords action (Gibson, 1979).

In recent years, researchers have gathered evidence in support of this view, showing tight links between perception and action across a wide range of cognitive and behavioral tasks (e.g., Witt & Proffitt, 2005; Bhalla & Proffitt, 1999; Witt, Proffitt, & Epstein, 2004). Other researchers have examined the role that motor actions play in object perception (e.g., Borghi et al., 2007; Bub et al., 2008; Chao & Martin, 2000; Helbig, Graf, & Kiefer, 2006; Tucker & Ellis, 1998, 2001; Witt & Brockmole, in press; Witt, Kemmerer, Linkenauger, & Culham, 2010). For example, Tucker and Ellis conducted a series of studies to test whether people automatically generate a motor representation in response to the visual presentation of an object, even when there is no intention to act on the object (Tucker & Ellis, 1998; 2001). In one experiment, participants made a left or right-handed button press to indicate whether an image of an object on the screen was upright or inverted. The objects were chosen to have a clear right or left-handed affordance (e.g., a frying pan with a handle oriented to the left affords a left-handed grasp). Participants responded faster and made fewer errors when their responding hand was congruent with the (task-irrelevant) affordance of the object on the screen.

Additional work has found that the relationship between motor actions and object perception is *functional* and not merely epiphenomenal. For example, Borghi et al. (2007) found that participants were faster to respond a picture of an object when it was preceded by a picture of a hand displaying an action that was congruent with the object. The authors concluded that visually priming an action facilitates object recognition (see also Helbig et al., 2006, Witt & Brockmole, in press). This suggests that preventing someone from engaging in an action should impair object recognition in a parallel fashion. Indeed, Witt et al. (2010) showed that participants were slower and less accurate when responding to a picture of a tool if the handle in the picture was oriented towards the participant's hand that was busy squeezing a rubber ball.

Taken together, these studies suggest that motor information can play a significant role in object perception by affecting the speed and accuracy with which we perceive an object. However, it is unclear just how deeply motor information can penetrate into our visual perception of objects. For instance, can viewing or performing a particular action *qualitatively* affect this perceptual process and change what object we actually see?

We investigated this possibility across three experiments. In Experiments 1 and 2, participants first viewed an image of a hand depicting a particular action (one of two specific grasp types). They then saw an image of an ambiguous object and had to indicate what they perceived it to be. Participants were biased to interpret the object as congruent with the action prime.

What cognitive mechanisms might underlie this effect? One possibility is that viewing the hand action prime led participants to imagine or simulate performing that action themselves (Parsons, 1987; Rizzolatti & Craighero, 2004). Then, when they viewed the ambiguous image, participants saw the object they were prepared to interact with because of this active motor state (Hommel et al., 2001). On this view, perceived events and planned actions share a common representational medium to the extent that they share common (abstract) features. Alternatively, this effect may have simply been a result of purely visual or semantic priming due to the association between certain grasp types and certain objects.

To distinguish these possibilities, in Experiment 3 participants engaged in an actual manual motor action (moving small or large balls from one tray to another) while naming pictures displayed on a computer screen, including the ambiguous object used in Experiment 2. A visual priming account would predict that performing an action should have no effect on this task as long as participants cannot see their own hands as they engage in the action. A semantic priming account would predict that no matter how the action concept is activated (e.g. viewing an action, talking about an action, performing an action), the results should yield the same facilitation effect observed in Experiments 1-2. Conversely, the common coding approach would predict that performing an action should actually *interfere* with a participant's ability to perceive an action-congruent object, which will therefore lead them to perceive an action-*incongruent* object (Hommel et al., 2001). In this study, participants were actually biased to interpret the object as incongruent with the motor action they performed. This suggests that viewing and performing actions are supported by the same underlying representations.

Experiment 1

Can viewing an action change what object we see?

Methods

Participants 815 individuals were recruited to participate in this study through the amazon.com Mechanical Turk website in exchange for payment.

Stimuli & Procedure The stimuli for this experiment consisted of four photographs of hands and an ambiguous object line drawing created by the authors (Figure 1). The four hand photographs showed either left or right hands in either a power or precision grasp. Pilot testing suggested that the ambiguous object could be interpreted as an object that afforded a power grasp (*flashlight*) or as an object that afforded a precision grasp (*screw/bolt*). The drawing could also be interpreted as an object that afforded a right-handed functional grasp (e.g., the *flashlight* as oriented in Figure 1) or as an object that afforded a left-handed functional grasp (e.g., the *screw/bolt* as oriented in Figure 1).

Figure 1. In Experiment 1, participants viewed one of the four hand images on the left, and then viewed the ambiguous object on the right.

One of the four hand images was randomly selected for each participant and displayed on the screen for three seconds. Next, the ambiguous object drawing was displayed at 56% the size of the hand for three seconds. The left/right orientation of the drawing was counterbalanced across participants. After this, participants were asked to identify the object in the line drawing that they had just seen. They were then asked to identify whether the hand they had seen was a left or right hand. Finally, they were asked if they had any additional interpretation of the object and indicated whether they were left-handed, right-handed, or ambidextrous.

Results

The data from 179 participants were removed from analysis because they failed to respond to the test questions appropriately (e.g., did not provide an interpretation of the ambiguous object), because they took the survey more than once, or because they responded incorrectly to the question of whether the hand prime they saw was a left or right hand. This last question was used as manipulation check to ensure that participants were looking at and paying attention to the experimental stimuli.

Testing for effects of grasp type. For the remaining 636 participants, we coded their initial ambiguous object interpretation as *congruent* if it matched the hand prime they saw (i.e., power grasp hand and flashlight interpretation or precision grasp hand and bolt/screw interpretation). Responses were coded as *incongruent* if they did not match the hand prime (i.e., power grasp hand and bolt/screw interpretation or precision grasp hand and flashlight interpretation). 19 participants came up with both interpretations for the ambiguous object and were therefore removed from further analysis.

Of the 617 participants in this final set of data, nearly 61% (N=374) gave *congruent* responses, while 39% (N=243) gave *incongruent* responses (Figure 2). A chi-square goodness of fit test revealed that this difference was highly significant, $\chi^2 = 27.4$, $p < 0.0001$.

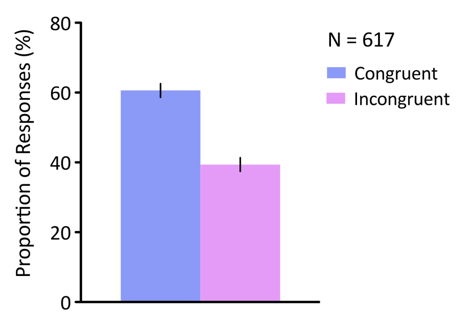

Figure 2. Results from Experiment 1, showing proportion of congruent and incongruent object interpretations. Error bars represent the standard error of the proportion.

One possible explanation for these results is that our study design was transparent and therefore our participants simply told us what they thought we wanted to hear. If our results were caused by this demand characteristic, then participants presumably perceived *both* interpretations for the ambiguous object and selected the interpretation they believed would make us happy. We tried to account for this possibility by asking participants if they had any additional interpretation of the object as one of our follow-up questions. In fact, 126 of our 617 participants provided additional interpretations of the object. Among the remaining 490 participants who only perceived one object interpretation, the results mirrored our previous analysis: nearly 61% (N=297) gave *congruent* responses, while 39% (N=193) gave *incongruent* responses. A chi-square goodness of fit test revealed that this difference was highly significant, $\chi^2 = 21.7$, $p < 0.0001$.

Testing for effects of orientation. We also asked whether the laterality of the action prime or participants' own handedness biased perception of the ambiguous object. To test these possibilities, participants' interpretations were coded as *leftward* if they saw the object whose handle (i.e., the head of the "bolt" or the barrel of the "flashlight") pointed to the left, and *rightward* if they saw the object whose handle pointed to the right. Neither the laterality of the action prime (51% congruent, N=249; 49% incongruent, N=241; $\chi^2 = 0.1$, $p > 0.5$), nor the handedness of the participant (51% congruent, N=238; 49% incongruent, N=237; $\chi^2 < 0.01$, $p > 0.5$), predicted whether subjects made a leftward or rightward interpretation of the object.

Discussion

In this experiment we asked whether viewing an action would influence what participants saw when they looked at an ambiguous object. We found that when participants were primed with a hand displaying a power grasp they were more likely to interpret an ambiguous drawing as an object that required a power grasp (*flashlight*). Conversely, when they were primed with a hand displaying a precision grasp, they were more likely to interpret the drawing as an object that required a precision grasp (*screw/bolt*). These results remained even after we removed participants who provided multiple interpretations of the ambiguous object, which helps to rule out an explanation based on demand characteristics. These findings suggest that viewing an action can qualitatively affect our perception of an object.[1]

However, manual actions are complex, and grasp type is just one dimension out of many that might affect object perception. In Experiment 1, we also tested whether priming an action with a right or left hand, irrespective of whether it displayed a power or precision grasp, would influence whether people perceived a leftward or rightward-facing object. We also reasoned that action simulations might be constrained by the idiosyncrasies of an individual's own motor system, so we tested whether the handedness of each participant, irrespective of the action prime, caused them to see a leftward or rightward-facing object. In our task, neither the laterality of the action prime nor the handedness of the participant affected what the ambiguous object appeared to be.

Why might the type of grasp displayed by a hand affect object perception, but not the laterality of the grasp or handedness of the participant? Perhaps some features of actions become privileged over others because they are more reliably associated with specific objects. Whether an object requires a power or precision grasp, for example, depends largely on the object's size, and for artifacts like flashlights and bolts, size is relatively constant across instances. The hand we use to grasp these objects, however, varies considerably depending on what we intend to do with the object and what else our hands are busy doing. By pitting various features of manual action against one another, we might have limited the likelihood that weaker effects of laterality and handedness would materialize. Exploring this possibility with objects that are ambiguous on one dimension only is the subject of future work.

It is also worth noting that the object we used in Experiment 1 was a *tool* under all possible interpretations, which might further limit our ability to generalize the effects of viewing actions to all graspable objects. Would the patterns we found in Experiment 1 with the flashlight/bolt image extend to graspable objects whose *primary* affordance is related to eating and not grasping (e.g., fruit)? Furthermore, the flashlight/bolt image remains an abstract, ambiguous, unrealistic line drawing. Would a photorealistic image in which the ambiguity of the object was less obvious show similar effects from viewing actions? To test these possibilities, we replicated this study in Experiment 2 using a photorealistic image of an object that could either be seen as an apple or a cherry.

Experiment 2

In Experiment 1 we found that viewing an action influenced what participants saw when they looked at an ambiguous object. However, it remains unclear whether these results will generalize to more realistic-looking objects that are not in the tool category. To address this issue, we created a new ambiguous object, the photorealistic image depicted in Figure 3 that can be interpreted as an apple (power grasp affordance) or a cherry (precision grasp affordance). We then replicated Experiment 1 using this new object.

Methods

[1] The results of Experiment 1 replicate a pilot version of this study reported at an earlier meeting of the Cognitive Science Society conference (Flusberg, Toskos Dils, & Boroditsky, 2010). Though the main effect in that study was nearly the same as in Experiment 1, the nature of the ambiguous object we used (a line drawing that could be perceived as a football or a nut) limited how we could interpret the results. First, this object elicited much more varied interpretations than the stimuli used in the present set of studies, suggesting that it may have been perceived as an extremely abstract figure rather than a concrete object. Second, a greater proportion of participants had multiple interpretations of the object than we see in the present study.

Participants 353 individuals were recruited to participate in this study through the amazon.com Mechanical Turk website in exchange for payment.

Stimuli & Procedure The stimuli and procedure for this experiment were identical to Experiment 1, with the exception of the ambiguous object, which was the cherry/apple picture depicted in Figure 3 presented at 29% the size of the hand.

Figure 3. The ambiguous object created for Experiment 2. It can be interpreted as an apple, which affords a power grasp, or a cherry, which affords a precision grasp.

Results

The data from 22 participants were removed from analysis because they failed to respond to the test questions appropriately, because they took the survey more than once, or because they responded incorrectly to the question of whether the hand prime they saw was a left or right hand. For the remaining 335 participants, we coded their initial ambiguous object interpretation in the same way we did in Experiment 1. Four participants came up with both interpretations for the ambiguous object and were therefore removed from further analysis.

Of the 331 participants in this final set of data, 58% (N=192) gave *congruent* responses, while 42% (N=139) gave *incongruent* responses (Figure 4). A chi-square goodness of fit test revealed that this difference was highly significant, $\chi^2 = 8.16$, $p < 0.005$. Once again, we used responses to our follow-up question to help rule out a demand characteristic account of these results. 110 participants provided additional interpretations of the object. Among the remaining 221 participants who only perceived one object interpretation, the results mirrored our previous analysis. Nearly 62% (N=136) gave *congruent* responses, while 38% (N=85) gave *incongruent* responses. A chi-square goodness of fit test revealed that this difference was highly significant, $\chi^2 = 11.32$, $p < 0.001$. The pattern of results produced by the *cherry*/*apple* in Experiment 2 did not differ reliably from the pattern produced by the *flashlight*/*bolt* from Experiment 1, $\chi^2 = 0.03$, $p > 0.5$.

Discussion

The results of Experiment 2 replicated what we found in Experiment 1 using a photorealistic ambiguous object that was in a very different category from the tool image used in the previous study. Taken together, these experiments demonstrate that viewing an action can qualitatively change how people perceive an object.

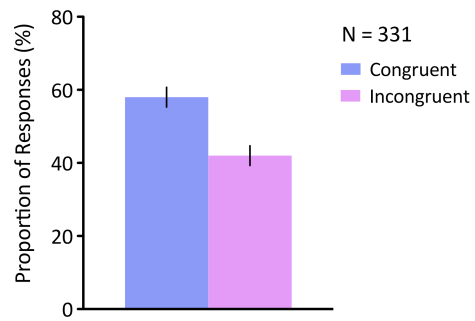

Figure 4. Results from Experiment 2, showing proportion of congruent and incongruent object interpretations. Error bars represent the standard error of the proportion.

What cognitive mechanisms might underlie this effect? One possibility is that the hand action prime led participants to simulate performing that action themselves (Parsons, 1987; Rizzolatti & Craighero, 2004). Then, when shown the ambiguous image, participants saw the object they were prepared to interact with because of this active motor state (Hommel et al., 2001). On this view, perceived events and planned actions share a common representational medium to the extent that they share common (abstract) features. Alternatively, this effect may have simply been a result of visual or semantic priming due to associations between grasp types and objects. Importantly, these accounts make three distinct predictions about how performing an action when participants cannot see their hands should affect object perception. A visual priming account would predict that performing an action should have no effect on this task as long as participants cannot see their own hands as they engage in the action. A semantic priming account would predict that no matter how the action concept is activated (e.g. viewing an action, talking about an action, performing an action), the results should yield the same facilitation effect observed in Experiments 1-2. Conversely, the common coding approach would predict that performing an action should actually *interfere* with a participant's ability to perceive an action-congruent object, which will therefore lead them to perceive an action-*incongruent* object (Hommel et al., 2001). Experiment 3 was designed to differentiate among these possibilities by having participants perform a manual motor action while they observed the ambiguous object used in Experiment 2.

Experiment 3

Can performing an action change what you see in the same way that observing an action does?

Methods

Participants 102 individuals were recruited to participate in this experiment from the Stanford community in exchange for course credit or five dollars.

Stimuli & Procedure When participants entered the lab, they were told they would be partaking in a study of multitasking. They were then seated at a desk and positioned with their head in a chin rest facing a computer screen (Apple iMac, 20" monitor). At this point they were given detailed instructions for how to proceed in the task.

The motor action participants performed consisted of picking up and moving balls located in a tray underneath the desk they were seated at (Figure 5). Participants picked up one ball in each hand from the lower tray and moved them simultaneously to the upper tray whenever they heard a beep coming from the computer. The apparatus was designed so that balls placed in the upper tray would fall back down into the lower tray. Importantly, with their heads in the chinrest, participants could not see this action as they performed it. Half of the participants were randomly assigned to pick up tennis balls, which require a power grasp action, while the remaining participants picked up small bouncy balls, which require a precision grasp action.

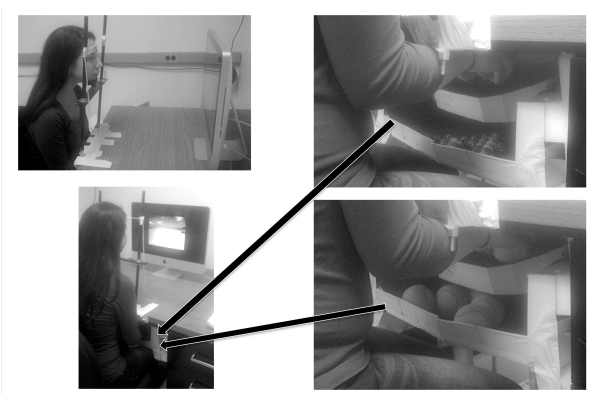

Figure 5. The laboratory setup used for Experiment 3. Half of participants moved bouncy balls in each hand (upper-right) while the remaining participants moved tennis balls in each hand (lower-right).

When the experiment began, the screen was black. A beep was played every 1.25 seconds, and each time it played participants engaged in the ball moving action. After 12.5 seconds, pictures started appearing on the screen one by one, each one remaining on the screen for 2 seconds, with an inter-stimulus interval of 500 milliseconds. While these pictures were appearing, the beeps kept playing at a rate of one every 1.25 seconds (twice per image).

Participants were instructed to name aloud the image on the screen as quickly as possible. There were 12 images in all, and the first 11 depicted objects or scenes that did not afford a particular manual grasp action (e.g., beach, house, etc.). The final image was the ambiguous cherry/apple object used in Experiment 2. The pictures were presented in the same order for all participants.

Results

The results from 2 participants were removed because they failed to complete the experimental task (i.e., they did not name every picture that appeared on the screen).

For the remaining 100 participants, we coded their response to the final picture (the ambiguous cherry/apple) as *congruent* if it matched the action they were performing (moving tennis balls and said apple, or moving bouncy balls and said cherry), and *incongruent* if it did not match the action they were performing (moving tennis balls and said cherry, or moving bouncy balls and said apple). 21 participants said both cherry and apple and were therefore removed from further analysis.

Of the 79 participants in this final set of data, 33% (N=26) gave *congruent* responses, while 67% (N=53) gave *incongruent* responses (Figure 6). A chi-square goodness of fit test revealed that this difference was highly significant, $\chi^2 = 8.56$, $p < 0.005$. This pattern differed reliably from the patterns observed in Experiment 1, ($\chi^2 = 20.87$, $p < 0.0001$), and Experiment 2, $\chi^2 = 18.06$, $p < 0.0001$.

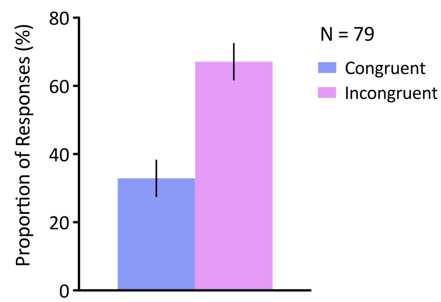

Figure 6. Results from Experiment 3, showing proportion of congruent and incongruent object interpretations. Error bars represent the standard error of the proportion.

Discussion

In this experiment we asked whether performing an action would influence what participants saw when they looked at an ambiguous object. We found that when participants engaged in power grasp action (moving tennis balls in each hand), they were biased to perceive an ambiguous object that was *incongruent* with that action (i.e., a cherry, which affords a precision grasp). Similarly, when they engaged in precision grasp action (moving small bouncy balls in each hand), they were also biased to perceive an ambiguous object that was incongruent with that action (an apple, which affords a power grasp).

Therefore, it seems that performing an action can change what people see when they look at an ambiguous object, and the direction of the effect suggests that it arises from overlapping representations between perception and action. Indeed, as predicted by the common coding account (and not by visual or semantic priming mechanisms), the specific pattern of results in Experiment 3 shows the *opposite* pattern from what we observed in Experiments 1 and 2 (when participants simply observed an action). In those experiments, viewing an action resulted in a *priming* effect, such that participants were biased to perceive an object that was congruent with the action they observed. In Experiment 3, on the other hand, performing an action resulted in an *interference* effect, such that participants were biased to perceive an object that was incongruent with the action they were engaged in.

However, there is one key difference between the experiments that may also have contributed to these divergent results. While participants in Experiments 1 and 2 only observed a single visual hand prime, participants in Experiment 3 engaged in a repetitive series of manual actions, moving balls from one tray to another 34 times. Behavioral repetition of this sort has been known to cause *adaptation* effects such that performance on a subsequent task is biased in the opposite direction of repeated behavior (e.g., Cattaneo et al., 2011). We are currently working on a new series of laboratory studies designed to tease apart the different possible mechanisms that may underlie the divergent patterns of results observed in these experiments.

General Discussion

We began this paper by asking whether viewing or performing an action could qualitatively affect how people perceive an object. That is, does our current motor state change how we see the world?

In Experiments 1 and 2, participants first viewed an action hand prime and then viewed an ambiguous object. They were biased to perceive the object as congruent with the preceding hand image. When the hand prime displayed a power grasp, participants were more likely to see an object that afforded such a grasp, like a flashlight or an apple. When the hand prime displayed a precision grasp, participants were more likely to see a bolt or cherry, which afford the same grasp type. In Experiment 3, participants performed a manual motor action while they interpreted the ambiguous object. When they were picking up tennis balls, which afford a power grasp, they were more likely to see an object that afforded a precision grasp (i.e., cherry). Similarly, when they were picking up small bouncy balls, which afford a precision grasp, they were more likely to see an object that afforded a power grasp (i.e., apple).

These results demonstrate that viewing or performing an action can in fact qualitatively change what an object is perceived to be, and the pattern of results across experiments suggests that this shift is subserved by shared representations between perception and action.

Acknowledgments

The authors would like to thank Laura Malkiewich for creating our cherple stimulus and help with earlier versions of these studies. We would also like to thank all members of the Cognation Lab. This research was supported by a McDonnell scholars grant & NSF BCS #1058119 to LB.

References

Bhalla, M. & Proffitt, D. R. (1999). Visual-motor recalibration in geographical slant perception. *JEP: Human Perception and Performance, 25*, 1076-1096.

Borghi, A.M., Bonfiglioli, C., Lugli, L, Ricciardelli, P., Rubichi, S., Nicoletti, R. (2007). Are visual stimuli sufficient to evoke motor information? Studies with hand primes. *Neuroscience Letters, 411(1)*, 17-21.

Bub, D. N., Masson, M. E. J., & Cree, G. S. (2008). Evocation of functional and volumetric gestural knowledge by objects and words. *Cognition, 106*, 27-58.

Cattaneo, L., Barchiesi, G., Tabarelli, D., Arfeller, C., Sato, M. & Glenberg, A.M. (2011). One's motor performance predictably modulates the understanding of others' actions through adaptation of premotor visuo-motor neurons. *Social Cognitive & Affective Neuroscience, 6(3)*, 301-310.

Chao, L. L. & Martin, A. (2000) Representation of manipulable man-made objects in the dorsal stream. *Neuroimage, 12,* 478-484.

Flusberg., S. J., Toskos Dils, A., & Boroditsky, L. (2010). Motor affordances in object perception. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society* (pp. 2105-2110). Austin, TX: Cognitive Science Society.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Lawrence Earlbaum: Hillsdale, NJ.

Helbig, H. B., Graf, M., & Kiefer, M. (2006). The role of action representations in visual object recognition, *Experimental Brain Research, 107*(2), 221-228.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences, 24*, 840-937.

Parsons, L. M. (1987). Imagined spatial transformation of one's body. *JEP: General, 19,* 178-241.

Rizzolatti, G. & Craighero, L. (2004). The mirror-neuron system. *Annual Reviews Neuroscience, 27*, 169-192.

Tucker, M. & Ellis, R. (1998). On the relations between seen objects and components of potential actions. *JEP: Human Perception and Performance, 24(3)*, 830-846.

Tucker, M. & Ellis, R. (2001). The potentiation of grasp types during visual object categorization. *Visual Cognition, 8(6)*, 769-800.

Witt, J. K., & Brockmole, J. R. (in press). Action alters object identification: Wielding a gun increases the bias to see guns. *JEP: Human Perception and Peformance.*

Witt, J. K., Kemmerer, D., Linkenauger, S. A., & Culham, J. (2010). A functional role for motor simulation in naming tools. *Psychological Science, 21,* 1215-1219.

Witt, J. K. & Proffitt, D. R. (2005). See the ball, hit the ball; Apparent ball size is correlated with batting average. *Psychological Science, 16*, 937-939.

Witt, J.K., Proffitt, D.R., & Epstein, W. (2004). Perceiving distance: A role of effort and intent. *Perception, 33*, 577-590.