University of Nevada, Reno

Physical Affordances Capture Attention

A thesis submitted in partial fulfillment of the
requirements for the degree of Master of Arts in
Psychology

by

Michael A. Gomez

Dr. Jacqueline C. Snow/Thesis Adviser

August, 2016

THE GRADUATE SCHOOL

We recommend that the thesis
prepared under our supervision by

MICHAEL A. GOMEZ

Entitled

Physical Affordance Captures Attention

be accepted in partial fulfillment of the
requirements for the degree of

MASTER OF ARTS

Jacqueline C. Snow, Ph.D., Advisor

Lars Strother, Ph.D., Committee Member

David Feil-Seifer, Ph.D., Graduate School Representative

David W. Zeh, Ph.D., Dean, Graduate School

August, 2016

Abstract

The opportunity an object presents for action is known as an 'affordance'. Studies using two-dimensional (2D) images of graspable objects demonstrate that affordances are processed automatically and have a powerful effect on attention and behavior. An important assumption in previous research, however, is that images elicit comparable effects on selection and filtering as real-world graspable objects. We tested this assumption by comparing interference effects on a flanker task using displays consisting of real-world tangible objects versus matched 2D (Experiment 1) or three-dimensional (3D) stereo images (Experiment 2) of the same items. As expected, flanker effects were observed for all types of stimuli. Critically, however, real objects elicited greater interference on responses overall, and irrelevant real flankers produced greater interference than planar and stereo image displays. Our results raise fundamental questions about the extent to which images are suitable proxies for real tangible objects in psychological research.

Acknowledgements

This research was supported by grants to JCS (NIH R01 EY026701-01; NIGMS P20 GM103650). We thank Andrew Delloro for help developing the stimuli.

Table of Contents

	Page
ABSTRACT	i
ACKNOWLEDGEMENTS	ii
LIST OF FIGURES	iv
LIST OF TABLES	v
INTRODUCTION	1
METHOD (EXPERIMENT 1)	4
Participants	4
Stimuli and Apparatus	4
Procedure	7
RESULTS (EXPERIMENT 1)	9
Response Times	9
Error Rates	11
METHOD (EXPERIMENT 2)	12
Participants	12
Stimuli, Apparatus, and Procedure	12
RESULTS (EXPERIMENT 2)	13
Response Times	13
Error Rates	15
EFFECT OF HANDLE ORIENTATION ON RTS FOR REAL OBJECTS	15
DISCUSSION	16
REFERENCES	21

List of Figures

Figure		Page
1	Stimuli and Trial Sequence in Experiments 1 and 2	6
2	Results of Experiment 1	10
3	Results of Experiment 2	14

List of Tables

Table		Page
1	Mean error rates (SEM) for each condition in Experiment 1	12
2	Mean error rates (SEM) for each condition in Experiment 2	15

Introduction

Natural environments contain many different objects, some of which are relevant for attention and others that are not. Our ability to direct attention to some objects over others is influenced by the opportunities those objects present for action (Humphreys et al., 2013), sometimes referred to as *affordances* (J. J. Gibson, 1979). Current models of attention argue that affordances (such as whether or not an object is graspable or is aligned with the observers' effectors) are learned via associative relations between object properties and our actions towards them, and that these high-level associations in turn trigger visuo-motor responses that modulate attention and behavior (Humphreys et al., 2013). Laboratory studies using images of graspable objects generally support this idea by showing that objects such as tools that have strong action-associations can be processed automatically (Makris, Grant, Hadar, & Yarrow, 2013; Murphy, van Velzen, & de Fockert, 2012), influence selection (Riddoch, Humphreys, Edwards, Baker, & Willson, 2003) and manual response times (Masson, Bub, & Breuer, 2011; Tucker & Ellis, 1998, 2004), and can stimulate neural activity in the dorsal motor (Grezes & Decety, 2002; Lewis, 2006) and ventral perceptual processing streams (Roberts & Humphreys, 2010).

To date, however, empirical studies have relied almost exclusively on two-dimensional (2D) images of objects to investigate affordance effects on attention. There are several potential problems with this approach. Theoretically, the notion that images are appropriate proxies for real objects diverges from classic theories of affordances which were envisioned in the context of real objects (J. J. Gibson, 1979). Ontologically, the practice of using images to study affordances overlooks the fact that images do not afford action, yet having the potential to physically interact with an object could have a powerful influence on attention. From an evolutionary perspective, the human brain has evolved presumably to allow us to perceive and interact with real objects and environments (Heft, 2013), and as such, images may constitute an atypical class of stimuli with which to characterize mechanisms of naturalistic vision. Indeed, there is a surprising paucity of studies in human adults that have examined whether real objects that afford physical actions are processed or represented differently than representations of objects. Differences in neural responses to stereo versus real environments have recently been revealed in the rat brain (Aghajan et al., 2015). In humans, whereas repetition of images or photographs of objects leads to a characteristic reduction in fMRI responses (fMRI-adaptation), the same effect is weak if not absent for real objects (Snow et al., 2011). Priming effects may not decay over time for real objects, as is often observed with pictorial primes (Squires, Macdonald, Culham, & Snow, 2015). Real objects are also more memorable than matched photographs or line-drawings of the same items (Snow, Skiba, Coleman, & Berryhill, 2014), can be recognized more accurately in patients with visual agnosia (Chainay & Humphreys, 2001; Humphrey, Goodale, Jakobson, & Servos, 1994), and may be perceived as being more valuable than images (Bushong, King, Camerer, & Rangel, 2010). In contrast to the mainstream empirical tradition of studying image perception in vision-related research, developmental psychologists have long recognized images as an unusual class of stimuli that must be learned to be fully understood (DeLoache, Pierroutsakos, Uttal, Rosengren, & Gottlieb, 1998), and have underscored the importance of physical exploratory behavior for normal perceptual development (E. J. Gibson, 1988).

In two experiments, we investigated whether real objects exert a stronger competitive influence on attention and manual responses compared to computerized images of objects. We measured the influence of display format on behavior using a classic flanker paradigm (Eriksen & Eriksen, 1974). When multiple objects offer the potential for interactive behavior, a number of motor action alternatives are computed simultaneously and compete for selection (Cisek, 2007). Resolution of the competition between different action alternatives produces a processing cost that slows response times (Jax & Buxbaum, 2010). In the flanker task, observers take longer to respond to a central target object when it is flanked by distractor objects that would elicit a different (i.e., incongruent), versus similar (i.e., congruent), manual response (Eriksen & Eriksen, 1974). The magnitude of this interference effect reflects the extent of processing of the 'to-be-ignored' distractors. Critically, we compared interference effects for real objects with 2D images (Experiment 1), and with 3D images that convey additional stereo depth cues (Experiment 2). Viewing time was controlled in all conditions and the stimuli were matched closely for apparent size and distance. A critical test of current affordance-based models of attention (Cisek, 2007; Humphreys et al., 2013) is to determine whether real graspable objects trigger more robust action plans and therefore compete more strongly for attention and manual responses, compared to images of the same items. In the context of a flanker paradigm, this leads to several important predictions. First, observers should take longer to respond on real object versus image trials, because the motor plan triggered by the stimulus (i.e., a grasp) conflicts with the required motor response on the task (i.e., a button-press). Second, interference from irrelevant flankers should be greater for the real objects compared to the image displays. To the extent that the effect of real objects on attention is attributable to stereo depth information present in the stimuli, then any differences we observed in flanker interference between real objects versus the 2D images should disappear in the context of stereo displays.

EXPERIMENT 1

Method

Participants

Forty undergraduate students (32 females; Age: *M* = 20.75, *SD* = 4.27), all of whom were enrolled in an undergraduate research methods class, participated in Experiment 1 in exchange for course credit. All participants reported having normal or corrected-to-normal vision, and were right-handed. All participants provided informed consent and the protocols used throughout the studies reported here were approved by The University of Nevada Reno Institutional Review Board (IRB).

Stimuli and Apparatus

In Experiment 1, we compared target response time (RT) and error rates elicited by stimuli in two different *Display Formats:* real objects versus 2D images of the same items. The stimuli in the real object displays were three white plastic spoons (Figure 1A). The spoons were 12.2 cm from tip to handle, and 2.61 cm at the widest point, and at a viewing distance of 60 cm, each spoon subtended 13.7° x 2.9°. One spoon (the *target*) was centered at the vertical and horizontal midpoint of the display. The *flankers*, which were also centered along the horizontal midline of the display, were positioned 3.66° above and below the target. The spoons were mounted on a black vertically-oriented display board made of composite wood material. The display board was 62.5 x 37.0 cm in size (63.2° x 40.3°), matching the outer dimensions of the LCD monitor used on the image trials (described below). The spoons were held in position during each trial using magnets, which were fixed to both the convex side of the spoon, and the rear surface of the display board. Small adhesive tape markers were positioned on the display board to ensure accurate stimulus alignment. The markers, which were occluded by the spoons when they were mounted, were not visible to participants during the trials.

On half of all trials, the target spoon was oriented with the handle facing rightward, and on the remaining trials the handle was oriented leftward. By virtue of the relative orientation of the target, and the flankers, we manipulated target-flanker *Incongruency*. On half of the trials, the flankers were oriented so that their handle faced the same direction as the target ('congruent' flanker trials); on the remaining trials, the flankers were oriented with their handles facing the opposite direction to the target ('incongruent' flanker trials). The combination of each target handle orientation (left / right) and flanker handle orientation (left / right) yielded a total of four unique display configurations, which were used to generate stimuli for the image displays (described below). A separate LCD monitor, positioned behind the subject, was used to display an image of the stimulus configuration required for the upcoming trial. The testing room was illuminated from above with in-ceiling fluorescent lights.

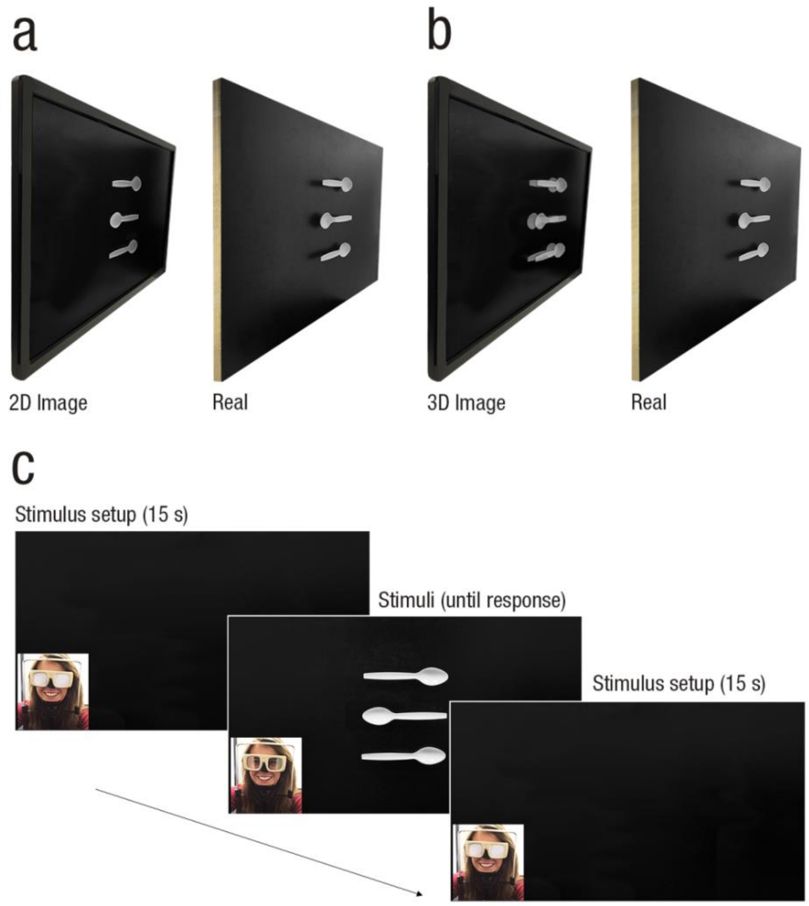

Figure 1. Stimuli and trial sequence in Experiments 1 and 2. (a) In Experiment 1 the displays consisted of either real spoons that afford grasping (shown right) or 2D computerized images of the spoons (shown left). All stimuli were presented within reach and were matched closely for background, apparent size and monocular depth cues. (b) In Experiment 2 we compared interference effects from real spoons whose 3D structure was defined using stereo images of the spoons. The stereo spoons were matched closely to the real objects for monocular and binocular depth cues, but did not afford genuine grasping. (c) Trial sequence in Experiments 1 and 2. After a 15 s setup period, the stimuli were presented and remained visible until response. Stimulus viewing time was controlled using PLATO LCD glasses (see inset, lower left), or via NVIDIA active shutter glasses on stereo trials in Experiment 2. White noise was played during the first 14 s of the setup period on all trials.

Stimuli in the image displays consisted of high-resolution color photographs of the stimuli in each of the four real object configurations. We used a Canon Rebel T2i DSLR camera with constant F-stop and shutter speed to photograph the real object displays. The stimuli were photographed separately for each display configuration, under identical lighting conditions. Image size was adjusted using Adobe Photoshop so that the stimuli in the 2D images were matched in size to the real spoons. Apart from re-sizing, the images were not cropped or otherwise adjusted, and the resulting stimuli preserved other monocular depth cues, including shadows and specular highlights that were present in the real object displays. The images were presented on a 62.5 x 37.0 cm, 27-in. LCD monitor (resolution: 1920 x 1080). Timing of events was controlled by a PC (Intel Core i7-4770 CPU 3.40 GHz, 16 GB RAM, NVIDIA Quadro GPU) running Matlab (Mathworks, USA). Stimulus viewing time in both the real object and image conditions was controlled using PLATO liquid crystal occlusion glasses (Translucent Technologies, Toronto, Canada) that alternate between opaque (closed) and transparent states (open). A chin rest was used to control viewing distance and stabilize the head between all trials.

Procedure

During each trial, participants made a two-choice response as to the orientation (left versus right) of a central target object (a spoon). The target was flanked above and below by two identical distractors whose handle orientation was the same (i.e., congruent) or opposite to (i.e., incongruent) the target, thereby eliciting a competing response. We did not include neutral flankers (stimuli that never appear as targets and are therefore not associated with a motor response) in our assessment of distractor processing. Flanker studies have generally found no consistent differences between RTs on Congruent and Neutral flanker trials (Lavie & de Fockert, 2003), and it is customary to compare interference from incongruent flankers with either congruent or neutral flanker trials (Avital-Cohen & Tsal, 2016). In the current study, we used Congruent flankers to ensure that our manipulation of *Incongruency* did not vary the number of response alternatives (Lavie & de Fockert, 2003). The stimuli were positioned so that the center of the display was aligned at eye-level when viewed from straight ahead in the chin mount. Each trial started with a 15 sec wait period during which the PLATO glasses were closed (Figure 1B). White noise was played during the first 14 seconds of the initial wait period on *all* trials to mask any sounds generated during mounting of the stimuli. Next, the PLATO glasses opened to reveal the stimuli, and remained in the transparent state until the subject's response. Responses were entered on a standard computer keyboard using the index finger of the left or right hand. The keys 'A' and 'L' were mapped to left and right hands, respectively. An auditory tone provided feedback on incorrect trials.

Participants completed two blocks of trials, separately for each *Display Format*. The order of blocks was counterbalanced within and between observers. The ordering of the target/flanker configurations was randomized within each block. Stimuli in each configuration were presented ten times; with four unique configurations and two display formats, this yielded a total of 80 trials. The entire experiment took ~30 minutes to complete. Participants were instructed to respond as quickly and accurately as possible whether the handle of the central target was oriented leftward or rightward. Participants were advised that the flankers were irrelevant to their task, that they held no predictive information about the orientation of the target, and that they should be ignored.

Results

Trials in which RTs were greater than 2 standard deviations from the mean were removed from all analyses, separately for each *Display Format*.

Response Times

Only trials with correct responses were entered into the reaction time (RT) analysis. The mean RT data were analyzed using a repeated measures analysis of variance (RM ANOVA) with the factors *Display Format* (real objects vs. images) and flanker *Incongruency* (congruent vs. incongruent). As expected based on previous flanker-paradigm studies, there was a significant main effect of *Incongruency* (F(1,39) = 9.11, $p <$ .001); participants responded faster when the flankers were congruent (M = 441 ms, SE = .003) than when they were incongruent (M = 474 ms, SE = .004) with the target. Critically, however, we also observed a significant main effect of *Display Format* (F(1,39) = 73.58, p < .001, ηp^2 = .654); participants were faster to respond to the 2D image targets (M = 445 ms, SE = .004) than the real object targets (M = 471 ms, SE = .004). Longer RTs were observed for real objects versus image displays in both the congruent (t(39) = 6.77, p = < .001, d = 1.071) and incongruent (t(39) = 8.83, p = < .001, d = 1.40) display configurations. Moreover, we found a significant two-way interaction between *Display Format* and *Incongruency* (F(1,39) = 10.36, p = .003, ηp^2 = .21). Figure 2A displays mean RTs, separately in each *Display Format* and *Incongruency* condition.

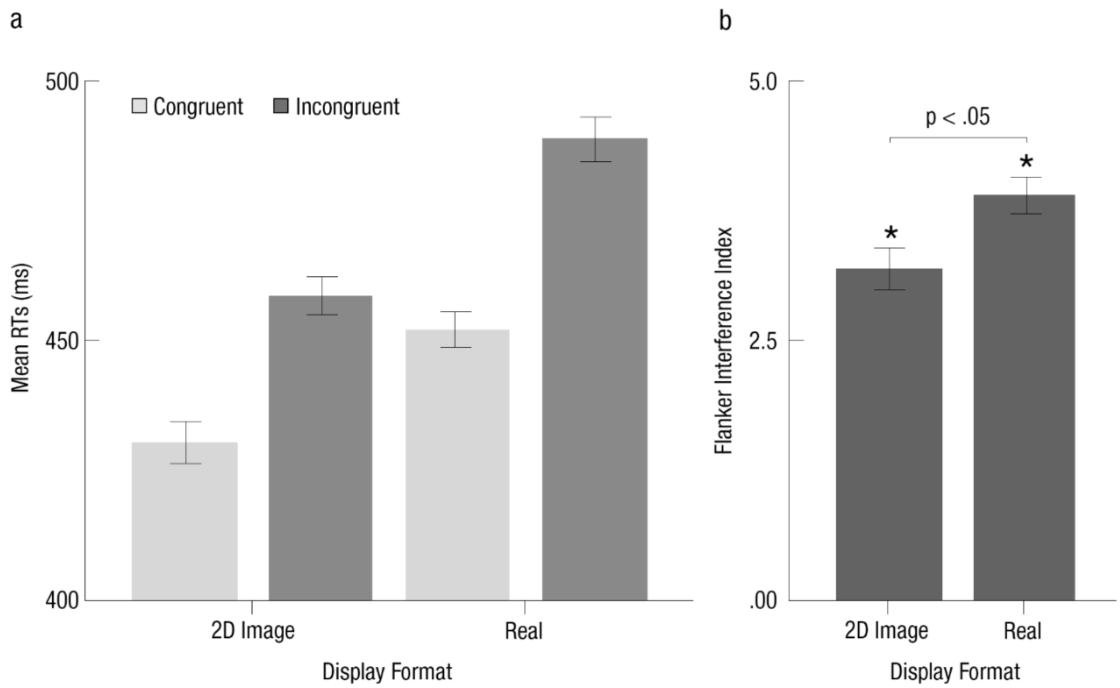

Figure 2. Results of Experiment 1. (a) Mean response time in each Display Format and target-flanker Incongruency condition. (b) Flanker Interference Index for stimuli in each Display Format. Error bars represent +/-1 *SEM*.

To break down this interaction, we quantified flanker effects using an interference index, which estimates the RT difference between the Incongruent and the Congruent conditions, and permits a direct comparison of the relative strength of interference effects across *Display Formats*. The index was defined using the formula:

$$\text{Interference Index} = ((RT_{incongruent} - RT_{congruent}) / (RT_{incongruent} + RT_{congruent}) \times 100)$$

where $RT_{incongruent}$ is the mean RT obtained on Incongruent flanker trials, and $RT_{congruent}$ is the mean RT obtained on Congruent flanker trials. Positive index values indicate longer RTs to Incongruent than to Congruent displays, negative values indicate the opposite, and values around zero indicate an absence of flanker interference (Figure 2B). One-sample t-tests against zero confirmed that flanker interference effects were significantly greater than zero for both the 2D image ($t(39) = 15.52$, $p < .001$), and the real object displays ($t(39) = 21.38$, $p < .001$). Critically, a paired-samples t-test revealed that the Interference Index for the real object displays was significantly greater than that of the 2D image displays ($t(39) = 2.35$, $p = .024$).

Error Rates

For consistency, we used the 2SD RT-filtered data set in the analysis of Error Rates, although the results for both experiments reported here were the same when all data were analyzed. Table 1 shows mean error rates and standard error of the mean (SEM) in each *Display Format* and *Incongruency* condition. A *Display Format* (real objects vs. images) x *Incongruency* (congruent vs. incongruent) repeated measures analysis of variance (ANOVA) performed on the mean error rates revealed no significant main effects or interaction (all *p*-values > .05).

Table 1. Mean error rates (SEM) for each condition in Experiment 1.

	2D Image		Real	
Congruency	*M*	*SEM*	*M*	*SEM*
Congruent	.05	.002	.02	.015
Incongruent	.08	.003	.04	.025

EXPERIMENT 2

Method

Participants

Thirty-eight right-handed undergraduate students (26 females; Age: $M = 22.0$, $SD = 4.96$) with self-reported normal or corrected-to-normal vision, participated in the experiment in exchange for course credit. Our sample size in this experiment was selected to match the number of participants tested in Experiment 1.

Stimuli, Apparatus, and Procedure

The stimuli, apparatus, and procedure for Experiment 2 were identical to those of Experiment 1, with the following exceptions: Stimuli in the image Displays were 3D stereo images of the spoons in each of the four target/flanker display configurations. We created stereo images by photographing each of the real object displays with a forward-facing camera, positioned at 60 cm from the screen, and 3.2 cm to the left and right of midline, respectively. Participants viewed the stimuli through glasses in both *Display*

Formats. The real objects were viewed using PLATO glasses, as described in Experiment 1. In stereo trials, the stimuli were viewed binocularly through active shutter glasses (3D Vision 2, NVIDIA, USA) and displayed on an ASUS (VG278HE) LCD monitor (120 Hz) with a screen resolution of 1920 x 1080 pixels.

Results

Trials with incorrect responses and those with RTs > 2 SDs from the mean were removed from the RT analyses, separately for each *Display Format* and *Incongruency* condition.

Response Times

A RM ANOVA with the factors of *Display Format* (real objects vs. images) and *Incongruency* (congruent vs. incongruent) showed a significant main effect of *Incongruency* ($F(1,37) = 115.64$, $p < .001$), where participants responded faster when flankers were congruent ($M = 431$ ms, $SE = .008$) versus incongruent ($M = 458$ ms, $SE = .008$) with target orientation. We again observed a significant main effect of *Display Format* ($F(1,37) = 9.31$, $p = .004$, $\eta p^2 = .201$), where participants were faster to respond to the stereo targets ($M = 435$ ms, $SE = .009$) than the real object targets ($M = 455$ ms, $SE = .009$). Significantly longer RTs for real objects versus images were observed in the incongruent displays ($t(37) = 3.74$, $p < .001$, $d = 0.607$) and there was a similar trend for the congruent displays ($t(37) = 1.84$, $p = .073$, $d = 0.298$). Critically, we also found a significant two-way interaction between *Display Format* and *Incongruency* ($F(1,37) = 7.62$, $p = .009$, $\eta p^2 = .171$). Figure 3A shows mean response times separately in each condition. We examined the magnitude of the flanker effects by calculating Interference Indices for the stimuli in each Display Format, as described in Experiment 1 (Figure 3B). One-sample t-tests confirmed that flankers interfered with RTs in both the stereo ($t(37) = 6.83$, $p < .001$) and real object displays ($t(37) = 7.68$, $p < .001$). Critically, however, a paired-samples t-test revealed that the Interference Index for the real objects was significantly greater than for the stereo objects ($t(37) = 2.65$, $p = .012$).

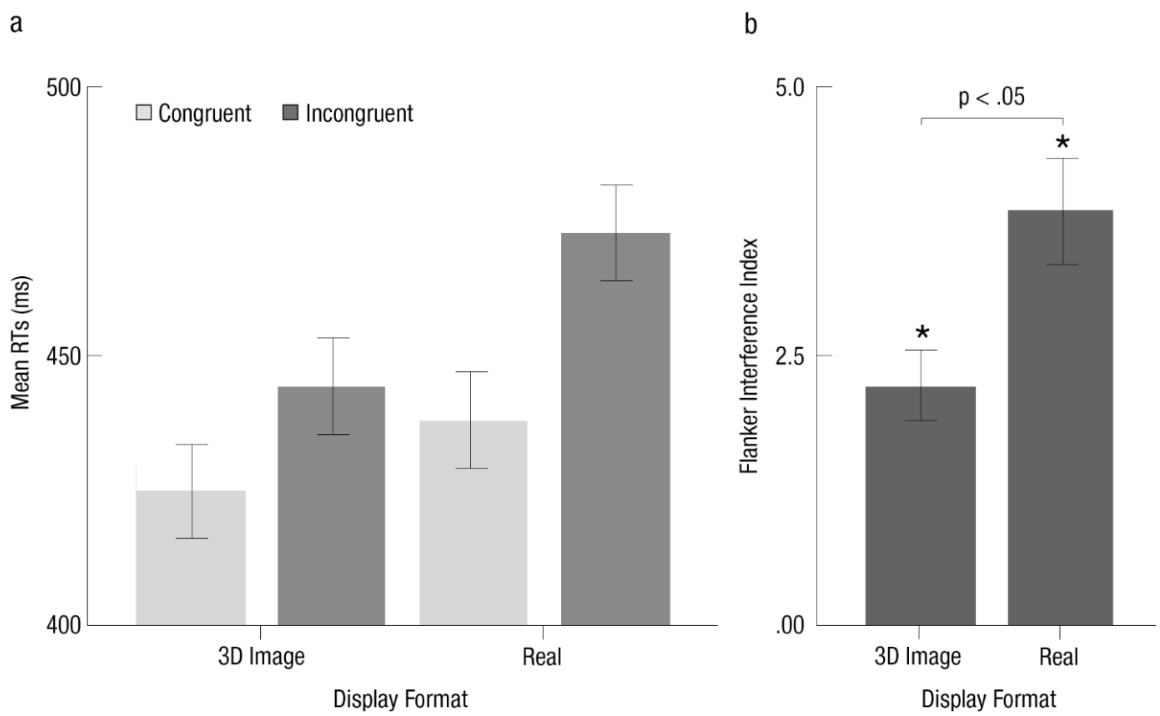

Figure 3. Results of Experiment 2. (a) Mean response time in each Display Format and target-flanker Incongruency condition. (b) Flanker Interference Index for stimuli in each Display Format. Error bars represent +/-1 Standard Error of the Mean (SEM).

Error Rates

Table 2 shows mean error rates (and SEM) for each condition in Experiment 2. A repeated measures ANOVA for mean error rates, with the factors of *Display Format* (real objects vs. images) and *Incongruency* (congruent vs. incongruent) revealed no significant main effects or interactions (all *p's* > .05).

Table 2. Mean error rates (SEM) for each condition in Experiment 2.

	3D Image		Real	
Congruency	*M*	*SEM*	*M*	*SEM*
Congruent	.02	.001	.01	.001
Incongruent	.02	.001	.02	.001

Effect of Handle Orientation on RTs for Real Objects

The results of Experiments 1 and 2 demonstrate that real graspable objects compete more strongly for attention than do computerized images of graspable objects: RTs were slower overall, and flanker interference effects were greater, for real spoons versus both 2D and 3D images of the spoons. The key question is whether the greater interference observed for real objects is a function of the fact they afford manipulability or whether they require more detailed and richer visual processing (and incongruence in the visual domain results in greater competition). If the disproportionate effects for real objects reflect competitive effects of affordances per se on attention and action, then RTs for the real objects should be influenced *not only* by flanker orientation (*Incongruency*), but *also* by target handle orientation. Targets whose handle orientation is compatible with a grasp with the dominant right hand should trigger a stronger motor plan, and therefore elicit greater interference with button-press responses, than leftward-oriented targets (and note that the visual input is equivalent in these two cases). To confirm this final prediction, we combined the data for real object trials from participants in Experiments 1 and 2. The RT data were analyzed using a mixed-model ANOVA with the within-subjects factors of *Incongruency* (congruent vs. incongruent) and *Target Orientation* (left vs. right), and the between-subjects factor of *Experiment* (1 vs. 2). As expected, there was a significant main effect of *Incongruency* ($F(1,76) = 230.11$, $p < .001$), where participants responded slower on incongruent ($M = 481$ ms, $SE = .005$) versus congruent ($M = 446$ ms, $SE = .005$) flanker trials. Critically, we also observed a significant main effect of *Target Orientation* ($F(1,76) = 7.30$, $p = .008$, $\eta p^2 = 0.01$), where participants were slower to respond when the target's handle was oriented rightward ($M = 467$ ms, $SE = .005$) versus leftward ($M = 460$ ms, $SE = .005$), consistent with the notion that for right-handed observers, rightward- (versus leftward-) oriented handles should elicit more robust action plans that compete more strongly with task-related button-press responses. Apart from a marginal main effect of Experiment ($F(1,76) = 3.05$, $p = .080$, $\eta p^2 = 0.04$: Experiment 1 RT = 472 ms, Experiment 2 RT = 455 ms), there were no other significant higher-order interactions (all p-values > 0.15), confirming that the effect of target orientation was comparable in both Experiments.

Discussion

We examined whether real objects that afford action exert a stronger competitive influence on attention and responses than matched image representations of the same items. We hypothesized that because real objects (but not images) afford action, they should trigger more robust stimulus-specific visuo-motor plans that compete more strongly for attention and behavior than those elicited by representations of objects. In the context of a flanker task, we predicted that real objects should serve as stronger competitors for manual responses than images, as reflected by longer overall RTs. Second, we predicted that irrelevant real object flankers should compete more strongly for attention with real targets and therefore produce greater interference effects, than images. In line with these predictions, compared to both 2D and 3D images, RTs were slower overall for real objects, and 'to-be-ignored' flankers interfered more strongly with target RTs. Finally, to the extent that the effects of real objects on attention reported above reflect affordances, we predicted that we should observe greater interference for real targets whose orientation was compatible with grasping by the right (dominant) versus the left hand. The follow-up analysis of RTs real object trials across Experiments 1 and 2 confirmed that this was indeed the case: RTs were slower overall on target handle right versus left trials. Together, our results demonstrate for the first time that real graspable objects exert a more powerful influence on attention and behavior than representations of those objects. Importantly, our experimental approach precludes a number of alternative explanations. Because the stimuli were presented in the same display format on each trial, differences in interference effects cannot be due to the apparent depth or relative conspicuity of the target and flankers. The powerful effect of tangible objects on attention also cannot be attributed to differences in scaling that could arise as a result of longer RTs on real object trials, because absolute interference measures on congruent versus incongruent trials were scaled by overall response times, separately for each display format.

The finding that real flankers elicited greater interference effects than images provides compelling support for cognitive models that emphasize the importance of action constraints on attention (Humphreys et al., 2013) and underscore the notion that the goal of attention is to ensure that we *act upon* the right object and capacity limits reflect physical constraints imposed by the number of actions that can be performed coherently on an object at a time. Our results also support the ‘affordance competition hypothesis’ (Cisek, 2007), and raise tantalizing new questions about how real object stimuli are represented in the brain and the nature of the motor plans that are triggered by real objects versus ‘virtual’ objects (Wamain, Gabrielli, & Coello, 2016). For example, the question of whether real objects trigger a greater number of competing action plans, or whether the plans and their associated feedback gains are more highly elaborated (Gallivan, Logan, Wolpert, & Flanagan, 2016), awaits future investigation. Future studies will also reveal the conditions under which real objects compete for attention, such as whether or not they are within reach (Gallivan, Cavina-Pratesi, & Culham, 2009; Iriki, Tanaka, & Iwamura, 1996), are physically accessible (Bushong et al., 2010), or must be visible at the time of response (Squires et al., 2015; Tucker & Ellis, 2004).

The slower RTs for real objects versus images in our task are consistent with a response incompatibility effect that arises due to differences between the motor plans elicited automatically by the stimulus, versus the response required for the task (Hommel, Musseler, Aschersleben, & Prinz, 2001). The powerful effect we observed for real objects is consistent with eye-movement studies showing that primates produce a greater number of spontaneous saccades towards (Held, Birch, & Gwiazda, 1980), and fixate longer upon (Mustafar, De Luna, & Rainer, 2015), real objects versus matched 2D representations. Another possibility is that real objects take longer to process than images because they are visually 'richer' and more complex. Somewhat paradoxically, previous studies have found that two-dimensional photographs of objects are named more quickly than line-drawings that provide fewer shape and (monocular) depth cues (Humphrey et al., 1994; Salmon, Matheson, & McMullen, 2014). Salmon et al., (2014) argued that richer representations of graspable objects are 'more embodied' –in that they are more likely to activate motor areas as a part of their object representation. Patients with visual agnosia are also faster to name real objects than 2D photos of the same items (Chainay & Humphreys, 2001; Humphrey et al., 1994). Our results suggest that real objects activate embodied representations more efficiently than impoverished image representations, but (unlike naming) when these representations conflict with the manual response required on the task, they result in strong interference effects.

Experimental stimuli may be seen as falling along a continuum, from planar images (such as line-drawings, silhouettes, or photos) to real-world exemplars. Real-world objects convey absolute information to the visual system about object size, distance and geometric shape. Whereas real world objects have a physical distance from a perceiver, objects depicted in a 2D image do not, except for the physical distance of the image plane from the perceiver. Veridical information about object size and distance is critical, however, for planning and executing goal-directed actions with an object (Goodale & Milner, 1992). While three-dimensional (3D) stereo images approximate more closely the apparent size, distance and geometric structure of real objects (Sanchez-Vives & Slater, 2005), only tangible objects afford physical grasping and interaction. A critical question for future studies will be to examine the extent to which the reduced interference effects in the stereo displays (versus real objects) in our task reflects a lack of tangibility, versus the presence of cue-conflicts between stereopsis (which signal that the stimulus has depth) and of vergence and accommodation (which signal that the stimulus is flat). Taken together, our data suggest that real objects have a more powerful influence on attention and action than matched 2D and 3D images of the same objects. Our results serve as an important catalyst for future studies to investigate the cognitive and neural processes that unfold during naturalistic vision. Just as normal cognitive and behavioral development of humans (Heft, 1981; Kretch & Adolph, 2015) and other animals (Held & Hein, 1963) relies on having the ability to exert active control on the environment in response to visual inputs, studying real-world vision may yield new and as yet undiscovered insights into the important reciprocity between an animal and its physical environment (Proffitt, 2006).

References

Aghajan, Z., Acharya, L., Moore, J., Cushman, J., Vuong, C., & Mehta, M. (2015). Impaired spatial selectivity and intact phase precession in two-dimensional virtual reality. *Nature Neuroscience, 18*(1), 121-128.

Avital-Cohen, R., & Tsal, Y. (2016). Top-Down Processes Override Bottom-Up Interference in the Flanker Task. *Psychological Science, 27*(5), 651-658.

Bushong, B., King, L. M., Camerer, C. F., & Rangel, A. (2010). Pavlovian processes in consumer choice: The physical presence of a good increases willingness-to-pay. *The American Economic Review, 100*(4), 1556-1571.

Chainay, H., & Humphreys, G. W. (2001). The real-object advantage in agnosia: Evidence for a role of surface and depth information in object recognition. *Cognitive Neuropsychology, 18*(2), 175-191.

Cisek, P. (2007). Cortical mechanisms of action selection: the affordance competition hypothesis. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 362*(1485), 1585-1599.

DeLoache, J. S., Pierroutsakos, S. L., Uttal, D. H., Rosengren, K. S., & Gottlieb, A. (1998). Grasping the nature of pictures. *Psychological Science, 9*(3), 205-210.

Eriksen, B., & Eriksen, C. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics, 16*(1), 143-149.

Gallivan, J. P., Cavina-Pratesi, C., & Culham, J. C. (2009). Is that within reach? fMRI reveals that the human superior parieto-occipital cortex encodes objects reachable by the hand. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience, 29*(14), 4381-4391.

Gallivan, J. P., Logan, L., Wolpert, D. M., & Flanagan, J. R. (2016). Parallel specification of competing sensorimotor control policies for alternative action options. *Nature Neuroscience, 19*(2), 320-326.

Gibson, E. J. (1988). Exploratory behavior in the development of perceiving, acting, and the acquiring of knowledge. *Annual Review of Psychology, 39*(1), 1-42.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Boston, MA: Houghton Mifflin.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neuroscience*, *15*(1), 20-25.

Grezes, J., & Decety, J. (2002). Does visual perception of object afford action? Evidence from a neuroimaging study. *Neuropsychologia*, *40*(2), 212-222.

Heft, H. (1981). An examination of constructivist and Gibsonian approaches to environmental psychology. *Population and Environment*, *4*(4), 227-245.

Heft, H. (2013). An ecological approach to psychology. *Review of General Psychology*, *17*(2), 162.

Held, R., Birch, E., & Gwiazda, J. (1980). Stereoacuity of human infants. *Proceedings of the National Academy of Sciences of the United States of America, 77*(9), 5572-5574.

Held, R., & Hein, A. (1963). Movement-produced stimulation in the development of visually guided behavior. *Journal of Comparative and Physiological Psychology*, *56*(5), 872.

Hommel, B., Musseler, J., Aschersleben, G., & Prinz, W. (2001). The Theory of Event Coding (TEC): a framework for perception and action planning. *Behavioral Brain Sciences, 24*(5), 849-878; discussion 878-937.

Humphrey, G. K., Goodale, M. A., Jakobson, L. S., & Servos, P. (1994). The role of surface information in object recognition: studies of a visual form agnosic and normal subjects. *Perception, 23*(12), 1457-1481.

Humphreys, G. W., Kumar, S., Yoon, E. Y., Wulff, M., Roberts, K. L., & Riddoch, M. J. (2013). Attending to the possibilities of action. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 368*(1628).

Iriki, A., Tanaka, M., & Iwamura, Y. (1996). Attention-induced neuronal activity in the monkey somatosensory cortex revealed by pupillometrics. *Neuroscience Research, 25*(2), 173-181.

Jax, S. A., & Buxbaum, L. J. (2010). Response interference between functional and structural actions linked to the same familiar object. *Cognition, 115*(2), 350-355.

Kretch, K. S., & Adolph, K. E. (2015). Active vision in passive locomotion: real-world free viewing in infants and adults. *Developmental Science, 18*(5), 736-750.

Lavie, N., & de Fockert, J. W. (2003). Contrasting effects of sensory limits and capacity limits in visual selective attention. *Perception & Psychophysics, 65*(2), 202-212.

Lewis, J. W. (2006). Cortical networks related to human use of tools. *Neuroscientist, 12*(3), 211-231.

Makris, S., Grant, S., Hadar, A. A., & Yarrow, K. (2013). Binocular vision enhances a rapidly evolving affordance priming effect: behavioural and TMS evidence. *Brain and Cognition, 83*(3), 279-287.

Masson, M. E., Bub, D. N., & Breuer, A. T. (2011). Priming of reach and grasp actions by handled objects. *Journal of Experimental Psychology: Human Perception and Performance, 37*(5), 1470-1484.

Murphy, S., van Velzen, J., & de Fockert, J. W. (2012). The role of perceptual load in action affordance by ignored objects. *Psychonomic Bulletin & Review, 19*(6), 1122-1127. doi:10.3758/s13423-012-0299-6

Mustafar, F., De Luna, P., & Rainer, G. (2015). Enhanced visual exploration for real objects compared to pictures during free viewing in the macaque monkey. *Behavioural Processes, 118*, 8-20.

Proffitt, D. R. (2006). Embodied perception and the economy of action. *Perspectives on Psychological Science, 1*(2), 110-122.

Riddoch, M. J., Humphreys, G. W., Edwards, S., Baker, T., & Willson, K. (2003). Seeing the action: neuropsychological evidence for action-based effects on object selection. *Nature Neuroscience, 6*(1), 82-89.

Roberts, K. L., & Humphreys, G. W. (2010). Action relationships concatenate representations of separate objects in the ventral visual system. *Neuroimage, 52*(4), 1541-1548.

Salmon, J. P., Matheson, H. E., & McMullen, P. A. (2014). Photographs of manipulable objects are named more quickly than the same objects depicted as line-drawings: Evidence that photographs engage embodiment more than line-drawings. *Frontiers in Psychology, 5*, 1187.

Sanchez-Vives, M. V., & Slater, M. (2005). From presence to consciousness through virtual reality. *Nature Reviews Neuroscience, 6*(4), 332-339.

Snow, J. C., Pettypiece, C. E., McAdam, T. D., McLean, A. D., Stroman, P. W., Goodale, M. A., & Culham, J. C. (2011). Bringing the real world into the fMRI scanner: repetition effects for pictures versus real objects. *Scientific Reports, 1*, 130.

Snow, J. C., Skiba, R. M., Coleman, T. L., & Berryhill, M. E. (2014). Real-world objects are more memorable than photographs of objects. *Frontiers in Human Neuroscience, 8*, 837.

Squires, S. D., Macdonald, S. N., Culham, J. C., & Snow, J. C. (2015). Priming tool actions: Are real objects more effective primes than pictures? *Experimental Brain Research*, 234(4), 963-976.

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

Tucker, M., & Ellis, R. (2004). Action priming by briefly presented objects. *Acta Psychologica (Amst), 116*(2), 185-203.

Wamain, Y., Gabrielli, F., & Coello, Y. (2016). EEG mu rhythm in virtual reality reveals that motor coding of visual objects in peripersonal space is task dependent. *Cortex*, *74*, 20-30.