Experiencing Objects: the Role of the Body

Ettore Ambrosini (e.ambrosini@unich.it)
Department of Neuroscience and Imaging, University G. d'Annunzio
Via dei Vestini, 33, 66013, Chieti, Italy

Claudia Scorolli (claudiascorolli@gmail.com)
Department of Psychology, EMbodied COgnition Lab, University of Bologna
Viale Berti Pichat, 5, 40100, Bologna, Italy

Anna M. Borghi (annamaria.borghi@unibo.it)
Department of Psychology, EMbodied COgnition Lab, University of Bologna
Viale Berti Pichat, 5, 40100, Bologna, Italy

Marcello Costantini (marcello.costantini@unich.it)
Department of Neuroscience and Imaging, University G. d'Annunzio
Via dei Vestini, 33, 66013, Chieti, Italy

Abstract

In this study we investigated whether the kind of action (i.e. manipulation vs. function) elicited by a visually presented object is modulated by its location in space, that is, its being within or outside the perceivers' reach-ability. Second, we investigated whether reach-ability mainly relies on the actual motor potentialities of individuals rather than their cognitive estimates of motor possibilities. Results showed that the activation of the potential actions to perform with objects is modulated by object accessibility. They also showed that accessibility is exclusively linked to the actual rather than the estimated reaching ability of the perceiver. The framework emerging from our results suggests that our cognitive capabilities are built online, via the current information in an implicit way that emerges in behavior, and are not necessarily reflected in explicit estimates or conscious representations. We believe this has interesting implications for embodied cognition theories, as it helps us to better qualify the notion of objects and of body/embodiment.

Keywords: Embodied Cognition; Affordance; Peripersonal space; reaching; reachability.

Introduction

Embodied and grounded theories, according to which cognition cannot be studied without considering that we possess a body endowed with specific sensorimotor capacities, have gained a wide and increasing consensus within cognitive science - neuroscience to robotics, cognitive linguistics, anthropology, philosophy (Chatterje, 2010; Gentner, 2010). In the last 10-15 years an impressive number of evidence has provided support to this perspective. Still, a lot of theoretical issues are unsolved, and embodied theories of cognition have a number of challenges to face. The aim of this work is to investigate the extent to which the way we perceive the relationship between the objects and our own body is dependent on the spatial context in which we operate. Specifically, we intend to verify whether objects evoke action possibilities (affordances) independently from their current location with respect to our own body. In addition, we intend to verify whether the way we explicitly represent our bodily capabilities in interacting with objects corresponds to the way we effectively exert them. We believe addressing these issues might have interesting implications for embodied theories, as it will help us to clarify to what extent our representations of objects and our own body are built online, and are contextually dependent.

Recent studies reveal that seeing objects activates the neural representation of their affordances (Gibson, 1979; Ellis & Tucker, 2000). A variety of studies on compatibility effect showed that observing pictures of objects or real objects activates the common reaching and grasping actions we typically perform with them. This effect refers to a decrease of reaction times when the subject executes a motor act which is congruent with that afforded by a seen object. For example, observing a handled cup leads to the activation of the movements aimed at reaching for its handle and the grip adequate to grasp it in order to drink (e.g., Tucker & Ellis, 1998; 2001; Ellis et al., 2007).

The activation of such motor information has also been shown by a recent study. Costantini and colleagues (Costantini et al. 2010) investigate whether and to what extent the effective processing of the affordances of an observed object might depend on its spatial location. The results showed that a compatibility effect occurred only when the visually presented object falls within the reaching space of an onlooker endowed with motor abilities which allows her to skilfully interact with the seen object, and that the first and more basic of these motor abilities is the reach-ability, i.e, the participant's motor potentiality of using her own body to physically reach and interact with the object.

In the first part of this study, we intend to extend previous finding by verifying whether the kind of action elicited by objects (i.e. manipulation vs. function) is modulated by their location in space, that is, their being within or outside the perceivers' reach-ability. To this purpose, we presented participants with verbs referring to function, manipulation and observation (e.g., “to drink”, “to grasp”, “to look at”) and were required to judge if the verb they read was compatible with a previously presented object. Hence we used response times to linguistic stimuli in order to understand which kind of information is activated while observing objects presented within (50 cm) or outside (170 cm) the participants’ reach-ability.

However, a distinction can be made between perceived reach-ability and actual reach-ability. A relatively common finding among studies of perceived estimates versus actual movement is the observation of an overestimation bias in reachability at midline positions (Bootsma et al., 1992; Mark et al., 1997; Fischer, 2000). That is, individuals exhibit a general tendency to perceive that they can reach objects that are actually out of grasp. Explanations for this bias have focused predominately on issues related to perceived postural constraints (Carello et al., 1989; Heft, 1993; Robinovitch, 1998) and the general notion that individuals ‘preconceive’ and calibrate such actions via ‘whole body engagement’ (e.g., Rochat & Wraga, 1997).

Thus, the second aim of the current study is to investigate, taking advantage of these naturally occurring differences between perceived reachability and actual reachability, whether the spatially constrained perception of an affording feature (i.e. the handle of a mug) mainly relies on the actual motor potentialities of individuals rather than their cognitive estimates of motor possibilities. To this aim in experiment 2 we employed the same task as in experiment 1 but presented the objects at four different distances, namely near reaching space (30 cm), actual reaching space (corresponding to each participant’s arm length), perceived reaching space (corresponding to each participant’s estimation of her own reaching space) and non reaching space (170 cm).

Experiment 1

Method

Participants 32 healthy subjects (17 males, mean age 33.5 years) took part in the experiment. All participants were native Italian speakers, had normal or corrected-to-normal visual acuity and were right-handed according to self report. They were naive as to the purpose of the experiment and gave their informed consent.

Materials We selected 12 critical Manipulation Verb - Function Verb - Object triples from a sample of 30 triple groupings. In order to perform the selection, we asked 48 Italian participants (22 males, mean age 30.9 years) to judge how compatible each verb was with each object. They were required to provide ratings on a 0-100 visual-analogical scale (Not compatible - Very Compatible), by making a cross on a line. We selected the triples (Manipulation Verb - Function Verb – Object) with highest compatibility scores. That is, for each object we had a highly compatible manipulation and function verb. As far as the Observation verbs are concerned, we used only four different verbs, due to the difficulty in finding a higher number of different verbs.

The experimental stimuli were images and verbs. Images consisted of red/cyan anaglyph stereo pictures depicting a 3D room displaying a table with an object placed on top of it. Twelve common objects were used (see Table 1). All of the objects used would normally be grasped with a power grip and were presented with the handle or the graspable part towards the right. Images were created by means of 3D Studio Max™ and StereoPhoto Maker. Using red/cyan anaglyph stereo pictures allowed us to present the objects either within the peripersonal (50 cm) or extrapersonal (170 cm) space of participants (See Fig 1, panel A). Verb stimuli consisted of three lists of Italian verbs in the imperative form. The three lists refer to function, manipulation and observation verbs, respectively (see Table 1). Each verb was matched with only one object, with the exception of the Observation verbs.

Table 1: Objects and verbs used in experiment 1
(in bold objects used in experiment 2)

	Verb		
Objects	***Function***	***Manipulation***	***Observatio n***
Ball	Gioca (to play)	Colpisci (to hit)	Osserva (to watch)
Bottle	Versa (to pour)	Tappa (to plug up)	Guarda (to look)
Brush	Pettina (to comb)	Stringi (to hold)	Fissa (to gaze)
Controller	Premi (to push)	Appoggia (to put)	Vedi (to see)
Fork	Mangia (to eat)	Raccogli (to pick up)	Osserva (to watch)
Funnel	Travasa (to pour)	Prendi (to bring)	Guarda (to look)
Hammer	Batti (to hammer)	Impugna (to clasp)	Fissa (to gaze)
Mug	Bevi (to drink)	Prendi (to bring)	Vedi (to see)
Pan	Cucina (to cook)	Lava (to wash)	Osserva (to look)
Pen	Scrivi (to write)	Sposta (to move)	Guarda (to look)
Screwdriver	Avvita (to screw)	Posa (to put down)	Fissa (to gaze)
Shovel	Scava (to dig)	Afferra (to grasp)	Vedi (to see)

Procedure Participants sat in front of a computer screen at a distance of approximately 57 cm, wearing anaglyph 3D glasses. Each trial consisted of the presentation of an object for 500 ms followed, after a delay of 50 or 100 ms, by a verb presented at the center of the screen and lasting 1500 ms (see Fig 1, panel B). Each trial began with the subject resting the right index finger on a response button. Participants were instructed to respond if the object-verb combination was appropriate, and to refrain from responding if the object-verb combination did not make sense (Catch trials). Catch trials were created by combining objects with verbs related to other objects (e.g. Object/Verb: Ball/To plug up; Ball/To drink). Responses were made by lifting the finger from the response button and then making an unspecified grasping movement toward the computer screen. During the inter-trial interval, a white fixation cross was presented for 1000 ms. The presentation of the stimuli and the recording of the participants' responses were controlled by a custom software (Galati et al., 2008), implemented in MATLAB, using Cogent 2000 (developed at FIL and ICN, UCL, London, UK) and Cogent Graphics (developed by John Romaya at the UCL, London, UK).

For every object, all of the three types of verbs were presented twice in both peripersonal and extrapersonal space; therefore there were 24 trials per condition for a total of 144 trials plus 48 catch trials (25%), lasting approximately ten minutes. At the end of the experiment participants were requested to estimate the distance of the objects in relation to their body. The stimuli presented in the peripersonal and extrapersonal spaces were judged as being at a distance of 50 ± 14 cm and 190 ± 42 cm from the participants.

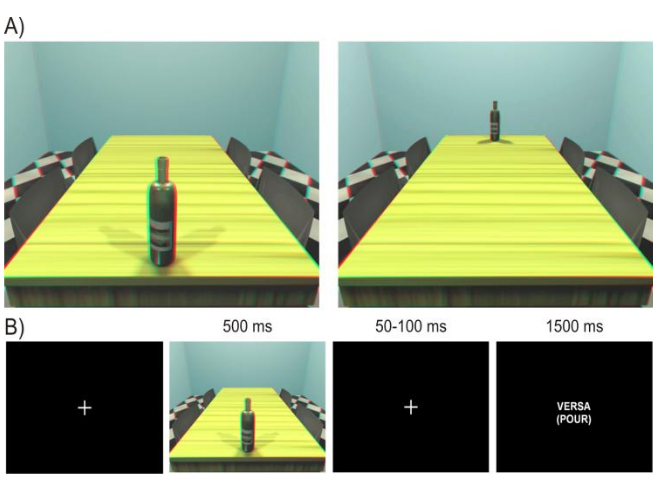

Fig 1: Example of experimental stimuli. Red/cyan anaglyph stereo pictures were used, allowing presenting the objects either within the peripersonal (50 cm) or extrapersonal (170 cm) space (panel A). Experimental timing (Panel B).

Results

Trials in which participants failed to respond (9.1%) were excluded from the analysis on response times (RTs). The mean RTs were calculated for each condition; responses longer than 2 standard deviations from the individual mean were treated as outliers (4.6%). Data were entered in a two-way ANOVA with Location of the object (Peripersonal vs. Extrapersonal space) and Verb (Function vs. Manipulation vs. Observation) as within-subjects factors.

RTs analysis revealed a significant main effect of object location ($F_{(1,31)} = 19.8$; $p < 0.001$), with higher RTs on extrapersonal trials (M = 798 ms) than peripersonal trials (M = 770 ms). The main effect of Verb was also significant ($F_{(2,62)} = 24.9$; $p < 0.001$). Post-hoc analysis (Newman–Keuls) revealed RTs to Function trials (M = 737 ms) being faster than both RTs to Manipulation (M = 792 ms) and Observation trials (M = 823 ms), which in turn did differ from each other. It is important to note here that the main effect of Verb is unlikely to be due to differences in the frequency of use. Indeed, we checked for it (DeMauro et al., 1993) and we found the following words frequencies: Function = 20; Manipulation = 19; Observation = 98. Thus, although Observation verbs had the highest frequency of use they had the slowest RTs.

Crucially RTs analysis revealed a significant Location by Verb interaction ($F_{(2,62)} = 7.4$; $p < 0.01$; Fig 2). Newman–Keuls post-hoc showed that while RTs to Observation verbs were comparable in the peripersonal and extrapersonal space (mean RTs: 822 vs. 823 ms), they were faster on peripersonal than extrapersonal space for both Function (mean RTs: 711 vs. 763 ms) and Manipulation verbs (mean RTs: 775 vs. 809 ms). Moreover, within the peripersonal space RTs to function verbs were faster than RTs to manipulation verbs ($p < 0.01$). Finally, RTs to Function verbs in the extrapersonal space were faster than RTs to Observation verbs in the same space.

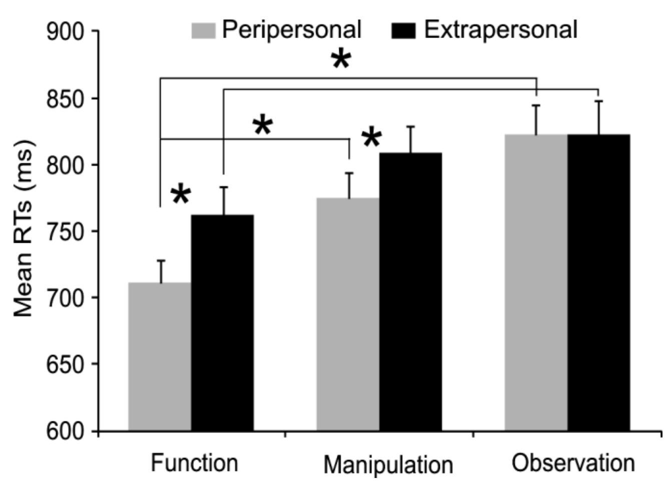

Fig 2: Mean reaction times in the experimental conditions. Error bars indicate standard errors.

Experiment 2

Method

Participants Fifteen healthy subjects (8 males, mean age 25.5 years) participated in the experiment. All participants were native Italian speakers, had normal or corrected-to-normal visual acuity and were right-handed according to self report. They were naïve to the hypotheses under investigation and gave their informed consent.

Materials The experimental stimuli were similar to those used in Experiment 1. The seven Manipulation Verb - Function Verb - Object triples were selected from Experiment 1.

As far as image stimuli are concerned, now there were seven objects presented at four different distances, with the closest and the most distant of them maintained fixed at 30 and 140 cm, respectively. The two intermediate distances, instead, varied for each subject and corresponded to her actual and perceived maximum reach range (see Procedure). In this way, the well-documented overestimation bias in perceived reachability (see for example Heft, 1993; Rochat and Wraga, 1997; Linkenauger et al, 2009) allowed us to present the objects either within (30 cm and actual reach range) or outside (perceived reach range and 140 cm) the reaching space of participants. As expected, participants systematically perceived the limit of their grasping space at farther distances than it actually was. In fact, participants estimated their reaching limit to be 71.5 cm, whereas the actual reach span was 61 cm (two-tailed t-test: $t_{(28)} = 4.98$; $p < 0.0001$).

Procedure Before the experimental task we conducted a preliminary session in order to collect the perceived and actual maximum reach distance for each subject. Participants were seated on a chair at a uniformly white table. The distance from the subject's eye to the table border was 25 cm, in order to maintain the same perspective as the visual stimuli. The experimenter moved a small object at a slow speed of about 2 cm/s away from or toward the participant. Subjects were instructed to say "stop" when they thought they could barely grasp it with the right arm without moving their shoulders from the back of the chair. The experimenter then stopped moving the object and used a tape measure to determine the distance between the participant's eye and the object. The average of these two measures was approximated at the nearest even centimeter and is here referred to as the perceived reaching space. There was no practice, and the subjects were not allowed to try out their reaching ranges on the table surface. After the estimation, the experimenter assessed the actual reach range by asking participants to place the object as far as they can without leaning forward.

The procedure of the experimental task was the same as in Experiment 1. For every object, all of the three types of verbs were presented twice in each of the four distances, resulting in 14 trials per condition, for a total of 168 trials plus 42 catch trials (20%).

Results

Trials in which participants failed to respond (0.4%) were excluded from the analysis on response times (RTs). The mean RTs were calculated for each condition; responses more than 2 standard deviations from the individual mean were treated as outliers (4.1%). Data were entered in a two-way ANOVA with Object Location (Near-Reaching vs. Actual-Reaching vs. Perceived-Reaching vs. Non-Reaching space) and Verb (Function vs. Manipulation vs. Observation) as within-subjects factors. Whenever appropriate, post hoc comparisons were performed with the Newman–Keuls method. An alpha level of 0.05 was always used.

The analysis revealed a significant main effect of Object Location ($F_{(3,42)} = 6.36$; $p = 0.001$). Post-hoc analysis showed faster RTs for both Near-Reaching (M = 670 ms) and Actual-Reaching space (M = 670 ms) compared to those in both Perceived-Reaching (M = 701 ms) and Non-Reaching space (M = 713 ms). The ANOVA also revealed a significant main effect of Verb ($F_{(2,28)} = 33.39$; $p < 0.0001$). Post-hoc analysis revealed both RTs to Function and Manipulation trials (M = 653 and 675 ms, respectively) being faster than RTs to Observation trials (M = 738 ms). Moreover, the difference in RTs between Function and Manipulation trials (23 ms) reached the significance (p=0.046).

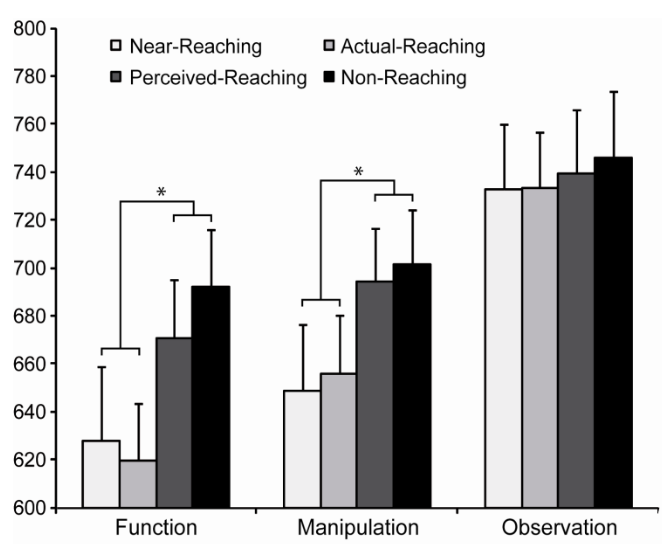

Fig 3: Mean reaction times in the experimental conditions. Error bars indicate standard errors.

The most important result, however, was the significant interaction between Object Location and Verb ($F_{(6,84)} = 2.88$; $p = 0.013$). This interaction is illustrated in Fig. 3 and shows that on Function trials, RTs were faster when the object was presented in both reaching spaces (M = 628 and 620 ms for Near-Reaching and Actual-Reaching space, respectively) compared to both non reaching spaces (M=671 and 692 ms for Perceived- Reaching and Non-Reaching space, respectively). The same RTs pattern was observed for Manipulation trials (M=649, 656, 694 and 702 ms for Near-Reaching, Actual- Reaching, Perceived- Reaching and Non-Reaching space, respectively), whereas there were no significant differences between RTs for objects presented at different location on Observation trials.

General Discussion

According to an embodied perspective, a mental process (e.g., representation) is "embodied" if and only if features of an individual's body play a non-trivial role in an explanation of that process (Wheeler & Clark, 1999). Relative the notion of affordance, it is embodied if and only if the perception of an affording feature strictly depends on the motor abilities the perceiver is endowed with.

The aim of this study is two-fold. Firstly, we aimed at investigating whether the kind of action elicited by objects (i.e. manipulation vs. function) is modulated by their location in space, that is, their being within or outside the perceivers' reach-ability. Secondly, we aimed at investigating whether reach-ability mainly relies on the actual motor potentialities of individuals rather than their cognitive estimates of motor possibilities.

Relative to the first question, our results clearly show that the activation of the potential actions to perform with objects is modulated by the current context and by object accessibility. RTs for manipulation and function verbs differed depending on the object location in the peri- vs. extrapersonal space, whereas RTs for observation verbs did not differ depending on the distance of the object from the body. This suggests that objects are represented in a flexible way, and that motor information related to both manipulation and use of objects is more relevant when a physical interaction with an object is effectively possible.

Possibly, one could explain the difference between Function/Manipulation and Observation verbs on the basis of our design: each Observation verb was presented more frequently during the experiment compared to each Manipulation and Function verb. Moreover, there were not catch trials with Observation verbs, so they were always responded to. However, our results contrast a frequency based account: indeed, Observation verbs were responded to more slowly than both Manipulation and Function verbs. Most importantly, consider that our task required participants to respond if the object-verb combination was appropriate (catch trials were only 25%), and that we did not use different blocks for each kind of verb. Due to the mixed design we used it would be improbable that participants formed separate categories for each verb kind (Observation, Function and Manipulation) and decided to respond to Observation ones, but not to the other verbs. To accomplish the task it is much more probable that they simply responded to the sensibility of each combination. Relative to the second question our results show that the perception of an affording feature is exclusively linked to the actual rather than the estimated reaching ability of the perceiver.

Overall, our results support a strong version of embodiment, in which behavior is explained by the interaction between the current context, in a given moment, and the specific characteristics of our own body. They highlight the necessity of taking into account both the spatial and the temporal context. First, they show that affordances are modulated by the effective possibility we have to interact with objects. We do not represent them in terms of stable characteristics – the actions evoked by a given object vary depending on where it is located in a given moment. Second, they show that the way our body acts does not correspond to the way we explicitly represent our bodily capabilities. As Lakoff and Johnson (1999) put it, both metaphysical realism and radical relativism would fail to explain our results, which clearly show that affordances are relational (Chemero, 2009; Costantini & Sinigaglia, 2011), dependent from the current relation with our own body, and that our bodily representation does not necessarily corresponds to how our body really is (Linkenauger et al., 2009). **Future Neuroimaging studies may help us in revealing the neuronal counterpart of our behavioural effects.**

The framework emerging from both experiments suggests that our cognitive capabilities are built online, via the current information in an implicit way that emerges in behavior, and are not necessarily reflected in explicit estimates or conscious representations. We believe this has interesting implications for embodied cognition theories, as it helps us to better qualify the notion of objects and of body/embodiment. In the radical view we are endorsing, objects are what you can effectively do with them, not what you could do in principle, independently from the current context. Similarly, the body is what you can do with your body in a given moment, not how you represent it offline.

Acknowledgments

Part of this work is in press: Costantini, Ambrosini, Scorolli & Borghi, When objects are close to me: affordances in the peripersonal space (in press), Psychonomic Bulletin and Review. The second and the third authors were supported by the European Community, project ROSSI: Emergence of communication in RObots through Sensorimotor and Social Interaction (Grant agreement n. 216125).

References

Bootsma, R. J., Bakker, F. C., Snippenberg, F. V., & Tdlohreg, C. W. (1992). The effects of anxiety on perceiving the reachability of passing objects. *Ecological Psychology*, *4*, 1–16.

Carello, C., Grosofsky, A., Reichel, F. D., Soloan, H. Y., & Turvey, M. T. (1989). Visually perceiving what is reachable. *Ecological Psychology*, *1*, 27–54.

Chatterje, A. (2010). Where is the knowledge we have lost in information? *International Journal on Humanistic Ideology*, *1*, 49–58.

Chemero, A. (2009). *Radical embodied cognitive science*. Cambridge: MIT Press.

Costantini, M., & Sinigaglia, C. (2011). *Grasping affordance: A window onto social cognition. In joint attention: New developments* (Axel Seemann ed). Cambridge: MIT press.

Costantini, M., Ambrosini, E., Tieri, G., Sinigaglia, C., & Committeri, G. (2010). Where does an object trigger an action? An investigation about affordances in space. *Experimental Brain Research*, *207*, 95–103.

DeMauro, T., Mancini, F., Vedovelli, M., & Voghera, M. (1993). *Lessico di frequenza dell'italiano parlato*. Milano: ETASLIBRI.

Ellis, R., & Tucker, M. (2000). Micro-affordance: The potentiation of components of action by seen objects. *British Journal of Psychology*, *9*, 451–471.

Ellis, R., Tucker, M., Symes, E., & Vainio, L. (2007). Does selecting one visual object from several require inhibition of the actions associated with nonselected objects? *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 670–691.

Fischer, M. H. (2000). Estimating reachability: Whole-body engagement or postural Stability*? Human Movement Science*, *19*, 297–318.

Galati, G., Committeri, G., Spitoni, G., Aprile, T., Di Russo, F., Pitzalis, S., et al. (2008). A selective representation of the meaning of actions in the auditory mirror system. *Neuroimage*, *40*(3), 1274-1286.

Gentner, D. (2010). Psychology in Cognitive Science: 1978–2038. *Topics in Cognitive Science*, *2*, 328–344.

Gibson, J. (1979). *The ecological approach to visual perception*. Boston: Houghton-Mifflin.

Heft, H. (1993). A methodological note on overestimates of reaching distance: Distinguishing between perceptual and analytical judgments. *Ecological Psychology*, *5*, 256–271.

Lakoff, G. & Johnson, M. (1999). *Philosophy In The Flesh: the Embodied Mind and its Challenge to Western Thought.* New York: Basic Books.

Linkenauger, S. A., Witt, J. K., Bakdash, J. Z., Stefanucci, J. K., & Proffitt, D. R. (2009). Asymmetrical body perception: A possible role for neural body representations. *Psychological Science*. *20*(11), 1373-1380.

Mark, L. S., Nemeth, K., Gardner, D., Dainoff, M. J., Paasche, J., Duffy, M., et al. (1997). Postural dynamics and the preferred critical boundary for visually guided reaching. *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 1365–1379.

Robinovitch, S. N. (1998). Perception of postural limits during reaching. *Journal of Motor Behavior, 30,* 352–358.

Rochat, P., & Wraga, M. (1997). An account of the systematic error in judging what is reachable. *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 199–212.

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

Wheeler M. & Clark A. 2008. Culture, embodiment and genes: unraveling the triple helix. *Philosophical Transactions of the Royal Society B: Biological Sciences, 363*, 3563–3575.