

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

The Treachery of Images: Objects, Pictures, Words and the Role of Affordances in Similarity Judgements

Permalink

<https://escholarship.org/uc/item/9sz5t2gn>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 44(44)

Authors

Wilford, Robyn Erica
Raja, Vicente
Hershey, Meghan
[et al.](#)

Publication Date

2022

Peer reviewed

The Treachery of Images: Objects, Pictures, Words and the Role of Affordances in Similarity Judgements

Robyn E. Wilford (rwilfor@uwo.ca)

Department of Psychology, Western University, 1151 Richmond St,
London, ON, Canada, N6A 3K7

Vicente Raja (vgalian@uwo.ca)

Rotman Institute of Philosophy, Western University, 1151 Richmond St,
London, ON, Canada, N6A 3K7

Meghan Hershey (mhershey@fandm.edu)

Department of Psychology, Franklin & Marshall College, 637 College Ave,
Lancaster, PA 17603, USA

Michael L. Anderson (mande54@uwo.ca)

Rotman Institute of Philosophy, Department of Philosophy, the Brain and Mind Institute,
Western University, 1151 Richmond St, London, ON, Canada, N6A 3K7

Abstract

Categorization is a fundamental cognitive strategy employed to ease information processing and to aid memory formation. Past research on how humans categorize objects has used images of objects as experimental stimuli. Concurrently, studies in the past 10 years have found differences in the processing of images as compared to real-world objects. One proposed explanation is that these results are due to differences in the affordances of images versus objects. Using a similarity judgement paradigm, we explored the effect of affordances in a categorization task including words (object names), images, and objects. Consistent with previous research, we found significant differences in how participants made similarity judgements of images and objects. Moreover, we found that similarity judgments using object names were much more similar to the judgments of pictures than of objects. An exploratory cluster analysis opens the possibility of framing such differences as affordance driven. These results suggest a need for more ecologically valid categorization tasks, more conservative inferences when using images as stimuli in these tasks, and the need for further exploring the role of affordances in categorization.

Keywords: categorization; affordances; similarity judgment; real-world object

Introduction

Categorization is the fundamental cognitive ability of grouping different objects or events under abstract, general classes or types on the basis of their shared properties or their similarity (for example, the grouping of different objects in a house under the concept of “chair”). Categorization eases information processing and helps guide new information into memory (Bornstein & Arterberry, 2010), and there is evidence that people may categorize objects based on abstract linguistic categories (Mur et al., 2013; Schmitt & Zhang, 1998).

The similarity judgement paradigm, in which objects that are judged to be more similar are grouped together and those that are judged to be less similar are placed apart, has been used to study categorization. For reasons of convention and convenience, experimenters typically use pictures of objects and not real-world objects in their categorization studies. They then generalize their findings to the case of real-world objects, sometimes even referring to them interchangeably (e.g., Chao et al., 1999; Haxby et al., 2001; Mur et al., 2013; Valyear et al., 2006; Van Weelden et al., 2011). There is, however, reason to believe that this inference may be inappropriate.

Research in the last decade has shown that we treat objects and pictures of objects differently in a variety of psychological processes (for a review, see Snow & Culham, 2021). For example, there is a well-established effect that the hemodynamic response in the inferior-temporal cortex decreases with repeated presentation of 2D visual stimuli, known as “fMRI adaption” or “repetition suppression” (Grill-Spector & Malach, 2001; Valyear et al., 2012). Snow et al. (2011) found that when the stimuli used were real-world objects (i.e., 3D visual stimuli) rather than pictures (i.e., 2D visual stimuli), the classic repetition suppression effect was not observed. Other examples include differences in memory retention (e.g., real world objects are remembered better than pictures; see Snow et al, 2014), differences in attention (e.g., graspable objects grab attention more readily than images; see Gomez et al., 2018), and differences in decision-making (e.g., willingness-to-pay and satiety expectations are greater for real food compared to images of food; see Romero et al., 2017).

A proposed reason for the differences found between our treatment of and responses to real-world objects and pictures is that they differ in their affordances (Gibson, 1979/2015; Turvey, 1992; Chemero, 2003; Heras-Escribano, 2019).

Affordances are defined as opportunities for action that organisms can perceive in their environment. For instance, a floor affords “walkability” to an adult human being and “crawlability” to a baby. Since affordances were first introduced by Gibson (1979/2015) they have been described as dispositional properties of the environment (Turvey, 1992; Heras-Escribano, 2019), as perceivable relationships between animals’ abilities and environmental properties (Chemero, 2003), or even as a form of action-oriented mental representation (Cisek, 2007; Clark, 2015; Pezzulo & Cisek, 2016). Snow and Culham (2021) point to this definitional ambiguity and advocate for the term “actability” to best describe the potential for genuine interaction between an organism and its environment. Their definition of actability is consistent with the Chemero (2003) definition of affordance and is the one we will use in this paper. Despite these diverse ontological characterizations, affordances are generally understood as a primary object of perception and that perception is fundamentally devoted to the control of successful action.

Experimental research on affordances has explored the way affordances are perceived to guide one’s interaction with the environment (for a review, see Fajen et al. 2009; Dotov et al. 2012). But if affordances are a primary object of perception, one would expect they also play a role in the way we categorize objects and events in our environment. And if affordances are a primary object of perception, one would expect they would be part of the explanation for the observed differences in the neural and psychological effects of objects vs. pictures. After all, objects and pictures of those objects afford very different actions—e.g., a hammer affords hammering while a picture of a hammer does not. Yet, there is very little research into the possible effects of affordances in this context and, crucially, in the similarity judgement paradigm for categorization tasks. For instance, while Holler and Snow (2020) recently found differences in how images and real-world objects are categorized, they did so without paying attention to the hypothesis that affordances may be responsible for those differences. Also, Castellini et al. (2011) found that affordances potentiated object recognition but did not investigate the effect on similarity judgements. So, considering the literature on the differences in how we treat real-world objects and pictures of those objects, and the fact that each afford different actions, will these differences in affordances influence the way objects and pictures are categorized?

To lay out a framework to start answering this question, we developed an experiment in which we asked participants to make similarity judgments of either real-world objects, pictures of those objects, or the names of those objects. We predicted that the similarity judgements would be different for objects, pictures, and words. We also predicted that, considering the evidence that pictures of objects are often categorized based on abstract linguistic categories (e.g., Mur et al., 2013), the similarity judgements for words would more closely resemble those for pictures than those for objects. We expected the greatest differences to be found between objects

and words. Finally, we preliminarily explored the hypothesis that the different affordances of objects with respect pictures and words at least partially explain these differences.

Methods

Data collection took place in the Spring of 2016 at Franklin & Marshall College (Lancaster, PA, USA). The study was conducted in accordance with Franklin & Marshall’s Research Integrity Policy and the protocol was approved by the Institutional Review Board.

Participants

48 undergraduate students from Franklin & Marshall College were recruited to participate in the experiment (N = 48). The sample consisted of 17 males and 31 females (1:1.8).

Experimental Paradigm

We used a categorization task based on the one used in Mur et al. (2013). There, participants were asked to make similarity judgments by arranging pictures on a computer screen such that pictures judged more similar were closer together on the screen, and those less similar were placed further apart. We moved this procedure into the real world and, in addition to the picture condition, included real objects and object names as additional experimental conditions in a between-subjects design. The degree of similarity between the two stimuli was represented by the geometrical distance between them after placement.

Experimental Set-up

The experiment took place in an empty room with a plain desk. A Canon Vixia hfs20 HD video camera was mounted on the ceiling directly above the desk. The stimuli consisted of 9 familiar objects: an apple (red delicious); a full roll of duct tape; a leather glove; a carpenter’s hammer; a lollipop; a pinecone; a wooden ruler; a seashell; and a tablespoon. Importantly, these objects could each be categorized using multiple criteria (e.g., typical location, colour, size, shape, typical usage, etc.) giving the participants multiple degrees of freedom in making their judgments. Three different representations of the 9 objects were used: the real-world objects (the Object Condition, Figure 1A); the pictures of those objects reflecting the same size and viewing angle as the real objects (the Picture Condition, Figure 1B); and the names of those objects (the Word Condition, Figure 1C). The pictures were printed on white letter-size card stock paper, and the names of the objects were printed in plain black lettering on the same card stock paper used for the pictures. We used a between-group experimental design: each participant experienced the same 9 different stimuli, but only in one of the three conditions; therefore, each condition contained 16 participants.

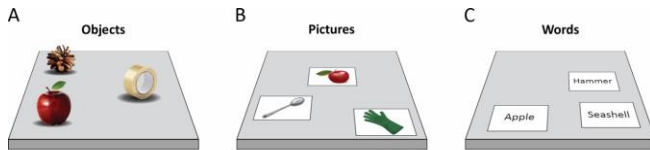


Figure 1: Diagram of the experimental setup for the three conditions. Participants were asked to place (A) real objects, (B) pictures, or (C) word names of the object on the table. They placed them at a distance that reflects how similar they judge the objects to be—the closer, the more similar.

Procedure

A bucket with (initially) 48 slips of paper each bearing the number 1, 2, or 3 (16 slips each) was prepared in advance. Each participant selected a slip of paper to determine the condition they were placed in, and that slip was discarded. Before each participant began their first trial, the experimenter shuffled a set of cards on which all possible triples of the 9 stimuli were written to determine the order of the trials. Thus, both the condition and the trial order were arbitrary for each participant. For each of the 84 trials, the experimenter placed the three stimuli indicated on the card in the centre of the table and instructed the participant: "Please arrange these so that the things that are more similar are closer together, and those that are less similar are further apart." The participant had as much time as they needed to arrange the stimuli, and importantly, they spontaneously made their similarity judgements without suggestion of how the stimuli should be categorized. After the participant was happy with their arrangement, a picture of the arrangement was taken using the camera mounted in the ceiling. This procedure continued until all cards in the set were used, concluding the 84th trial for the participant.

Measurement

The distances between stimuli were measured in pixels from the photographs taken after each trial using the software Greenshot (<https://getgreenshot.org>). Measurements were made from the geometrical centre of each stimulus. Each pair-wise distance was recorded in a Microsoft Excel spreadsheet, noting which trial and participant number the distance was taken from.

Data Preparation

We used Microsoft Excel to prepare our raw data for analysis. We first performed an outlier analysis on our data testing all pair-wise distances for each participant, looking for any pair-wise distance that was greater than 4 standard deviations above or below the mean pair-wise distance for that participant. In doing so, we found 4 outlying distances: 3 outliers for one participant in the picture condition, and 1 outlier for one participant in the word condition. For all 4 outliers, we replaced the outlying value with the average pair-wise distance for the respective participant.

Next, since relative rather than absolute distance between stimuli represents perceived similarity for our research question, we normalized our data (in two different ways to ensure our normalization procedure did not distort our results in undesired ways). First, we normalized every raw pair-wise distance for each participant using the following equation: $ND = D \cdot (MR/IR)$; where ND is the resulting normalized pair-wise distance, D is the raw pair-wise distance being normalized, MR is the mean range of pair-wise distances for the condition, and IR is the range of pair-wise distances for that individual. Second, we additionally normalized every raw pair-wise distance for each participant using the following equation: $ND = D \cdot (MD/TD)$; where ND and D are the same as above, MD is the mean pair-wise distance for the condition, and TD is the total distance (the sum of pair-wise distances) for the participant in question.

After normalizing our data, for each participant, we averaged the pair-wise distances for each pair of stimuli across trials and then constructed an individual 9x9 dissimilarity matrix using the averaged pair-wise distances. This procedure resulted in 48 individual 9x9 dissimilarity matrices – one for each participant, and 16 in each condition.

Analysis Methods

Using MATLAB 2020a, we performed a one-way ANOVA and followed-up with a Tukey's HSD post-hoc test of the mean pair-wise distances in our experimental conditions to confirm that they come from different populations.

For our main analyses, we also used MATLAB 2020a to design and implement a custom analysis program to test our hypotheses. First, from the individual 9x9 dissimilarity matrices, we created an average dissimilarity matrix for each of the three conditions (Figure 2). We then performed planned comparisons between each of our conditions: Object-Picture, Object-Word, and Picture-Word. For each comparison we performed a Mantel test and a Procrustes Analysis. The script for the Mantel tests was provided by Glerean et al. (2016). A Mantel test returns a measure of similarity between two matrices in the form of a Pearson correlation coefficient. A Procrustes test returns the measure of dissimilarity of two matrices by performing a linear transformation between the two. For the Mantel test, the higher the Pearson correlation coefficient the more similar the matrices; in contrast, the higher the Procrustes value the greater the transformation needed to line up the matrices, and thus the less similar the matrices are. Together these two tests provide complementary results of how similar the similarity judgements were across experimental conditions.

Next, to provide the context to test for whether any differences found between our conditions could be considered significant, we performed a bootstrapping procedure to construct population distributions of Procrustes scores and Pearson correlations by scrambling our condition labels. This procedure is based on the representational similarity analysis permutation method for comparing two dissimilarity matrices (Kriegeskorte et al., 2008). We iterated through the 48 individual 9x9 dissimilarity matrices 10,000

times, each time selecting three sets of 16 matrices (random selection without replacement). After averaging the pairwise distances across the sets, we then performed a Procrustes test and Mantel test between each of the three resulting averaged matrices. This resulted in two distributions of 30,000 values each, one for Pearson correlations and another for Procrustes scores, that we used to determine whether our experimental conditions resulted in a significant effect.

Also using MATLAB 2020a, we then performed an exploratory hierarchical agglomerative cluster analysis to see how the sets of 9 stimuli were categorized in each condition. We used a linkage method (Ward’s minimized variance) and set the colour threshold (each node having a unique colour) for the resulting dendrograms at 60% the maximum linkage between nodes. Using this method, the distance between two clusters (i.e., linkage distance) is how much the sum of squares will increase when the two clusters are merged; therefore, the smaller the linkage distance, the less variability under the respective node. So, this method creates clusters where the variance within each cluster is less than the variance between clusters. In the context of similarity judgement measurements, this means that those stimuli clustered together are considered more similar than those in differing clusters, in turn reflecting the underlying category representations for objects, pictures, and words, respectively.

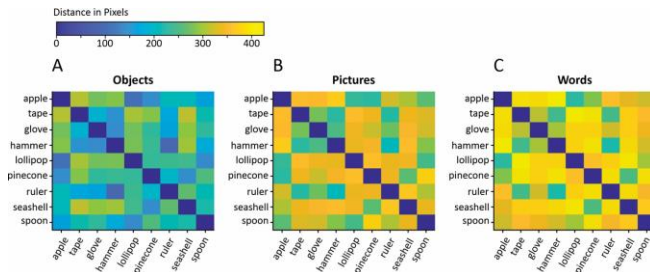


Figure 2: Dissimilarity matrices for Object Condition (A), the Picture Condition (B), and the Word Condition (C). Matrices depict all average pairwise distances (in pixels) between each pair of experimental stimuli.

Results

As noted in the Methods section, we normalized our data in two different ways. We performed the described analyses on both sets of normalized data and found the same results. So, for the sake of brevity, the results presented below are only for the data normalized using the first method described in the Methods (i.e., the range of pair-wise distances: $ND = D \cdot (MR/IR)$).

Differences in Mean Pair-wise Distances

We performed a one-way ANOVA testing for the effect of the condition on mean pair-wise distance and found a highly significant difference between at least two conditions, $F(2, 3453) = 652.01, p < .001$. A Tukey’s HSD post-hoc test revealed significant differences between all three conditions’ mean pair-wise distance (Figure 3). The mean distance

between object pairs ($N = 1,152, M = 220.11, SE = 2.58$) was significantly less than the mean distance between picture pairs ($N = 1,152, M = 313.33, SE = 2.58, q(3, 3453) = 41.41, p < .001$). The object mean distance was also significantly less than the mean distance between words ($N = 1,152, M = 347.13, SE = 2.58, q(3, 3453) = 57.60, p < .001$). And finally, the mean distance between pictures was significantly less than that for words ($q(3, 3453) = 16.19, p < .001$).

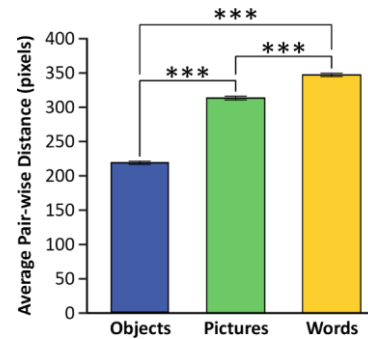


Figure 3: Mean pair-wise distance for each of the three experimental conditions: the Object Condition (blue bar), the Picture Condition (green bar), and the Word Condition (yellow bar). Each mean pair-wise distance was significantly different from the others ($*** = p < .001$). Error bars indicate the standard error of the means (SEM) for each condition’s mean pair-wise distance.

Between-Condition Planned Comparisons

Mantel Test We performed a Mantel test between the averaged dissimilarity matrices for each of the conditions to assess their degree of similarity. The Mantel tests returned a Pearson correlation coefficient for each comparison: Object-Word, Object-Picture, and Picture-Word. As expected, given the nature of the stimuli, all the matrices were significantly correlated. However, each also differed from the others. The object and word dissimilarity matrices were the least similar ($r = 0.88, p < 0.001$). The object and picture dissimilarity matrices were the second least similar ($r = 0.90, p < 0.001$). And the picture and word dissimilarity matrices were the most similar ($r = 0.95, p < 0.001$). All correlations were statistically significant.

Situating the Mantel Test Results in the Mantel Test Distribution To test whether the observed differences were significant, we performed a bootstrapping procedure to build a population distribution of Pearson correlation coefficients in which we could situate our results (see Methods). The observed correlation coefficient for the Object-Picture comparison fell far into the tail of the distribution, with only 1% of the distribution showing lower values. The Object-Word comparison was even further into the tail, with only 0.02% lower values in the distribution. In contrast, the Picture-Word comparison was in the middle of the distribution, being higher than 39% of the values (Figure 4A). This suggests that the Picture and Word matrices do not

significantly differ from one another (their differences can be accounted for by chance). In contrast, the Object matrix, despite being significantly correlated with the other two, also significantly differs from them as shown by the improbability that the difference in conditions had no effect on the correlations.

Procrustes Analysis We also performed a Procrustes Analysis between the averaged dissimilarity matrices for the conditions. A Procrustes analysis returns a score that is a measure of dissimilarity between two matrices, in our case we once again made three comparisons: Object-Picture, Object-Word, and Picture-Word. The Procrustes scores were as follows: the most dissimilar was 0.08 for the Object-Word comparison, the next most dissimilar was 0.07 for the Object-Picture comparison, and the most similar was 0.02 for the Picture-Word comparison.

Situating the Procrustes Analysis Results in the Procrustes Score Distribution As with the Mantel results, the Procrustes scores showed that all three matrices are quite similar, but we wanted to know if the observed differences were significant. We performed the same bootstrapping procedure to build a population distribution of Procrustes scores (see Methods) and situated the Procrustes scores from our planned comparisons within that distribution. As with the Mantel results, the Object-Picture comparison fell far into the tail of the distribution, with only 0.02% of results being higher (more dissimilar) and the Object-Word comparison was the very last value in the distribution. In contrast, the Picture-Word comparison was in the middle of the distribution, with 57% of the comparisons being more dissimilar (Figure 4B). These results converge with the Mantel distribution results in showing that the similarity judgements in the object condition were significantly different from those in the picture and word conditions, with the greatest difference being between the object and word conditions.

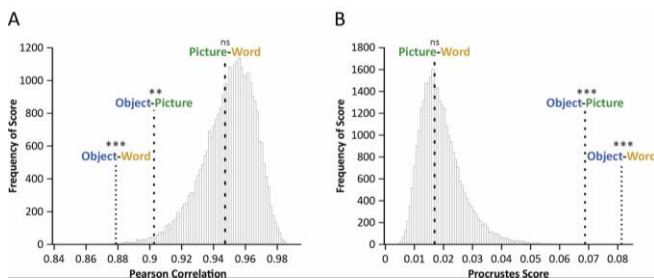


Figure 4. Mantel and Procrustes Analysis results situated in the bootstrapped distributions. (A) The histogram of the Pearson correlation coefficients (the outcome of the Mantel test). (B) The histogram of the Procrustes scores (the outcome of the Procrustes test). Dashed lines show the position of the correlations between the matrices of the experimental conditions in each distribution, indicating their significance. *** = $p < .001$; ** = $p < .01$; ns = not significant

Exploratory Cluster Analysis

We performed a hierarchical agglomerative cluster analysis (linkage method: Ward’s minimized variance) to explore the categorization of the 9 stimuli in each condition (see Methods). In the object condition, there were two main clusters, in the picture condition there were three main clusters, and in the word condition there were four main clusters. This indicated an increase in variability between categories when moving from objects to pictures to words being categorized.

The first distinct cluster in the object condition consisted of the seashell, pinecone, spoon, lollipop, and apple. Within this cluster there were subclusters of seashell and pinecone and spoon, lollipop, and apple, with the lollipop and apple being subsumed by the spoon-lollipop-apple cluster. The second distinct cluster in the object condition consisted of the tape, glove, ruler, and hammer objects, with a subcluster of glove, ruler, and hammer, with the ruler and hammer subsumed by the glove-ruler-hammer subcluster (Figure 5A). The first principal cluster in the picture condition consisted of the spoon and the lollipop. The second included the seashell, pinecone, and apple, with the pinecone and apple subsumed under the seashell-pinecone-apple subcluster. The third cluster in the picture condition contained the glove, tape, ruler, and hammer with subclusters of tape-ruler-hammer and ruler-hammer (Figure 5B). For the word condition, the first main cluster contained the seashell and pinecone, the second contained the lollipop and apple, the third contained the spoon and glove, and the fourth contained the tape, ruler, and hammer with the ruler-hammer forming a subcluster (Figure 5C).

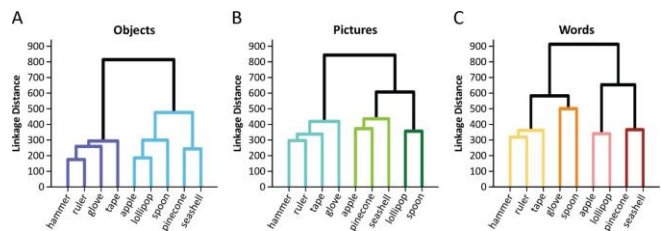


Figure 5: Results from hierarchical agglomerative cluster analysis (linkage method, Ward’s minimized variance; colour threshold: 60% maximum linkage). Vertical axis: linkage distance; horizontal axis: stimuli. (A) The Object Condition – two main clusters: hammer, ruler, glove, tape (dark blue); apple, lollipop, spoon, pinecone, seashell (light blue). (B) The Picture Condition – three main clusters: hammer, ruler, tape, glove (turquoise); apple, pinecone, seashell (light green); and lollipop, spoon (dark green). (C) The Word Condition – four main clusters: hammer, ruler, tape (yellow); glove, spoon (orange); apple, lollipop (pink); and pinecone, seashell (burgundy).

Discussion

Our findings in this experiment comparing similarity judgments made on objects, pictures, and words are consistent with the growing consensus that real objects affect us and are treated by us differently from representations of

those objects (Snow et al. 2011; 2014; Gomez et al. 2018; Romero et al. 2017). In the current case, we compared dissimilarity matrices constructed in our three conditions using both Procrustes and Mantel tests, employing a bootstrapping procedure to assess statistical significance.

In both constructed distributions, the Object-Picture and Object-Word comparisons fell significantly into the tail ends of the distributions. This indicates that the experimental manipulation using pictures of objects versus real objects created significant differences between the similarity judgments of those stimuli, and likewise with objects and words. The Picture-Word comparison falling into the middle of the distribution indicates that the stimuli being a picture rather than a word did not significantly affect the similarity judgements of those stimuli.

From this and previous results (for a review, see Snow & Culham, 2021) it is clear that researchers should exercise great caution when generalizing to the case of real objects when experimental results are obtained using object representations. This is the case even when considering experimental settings with different delivery apparatus. For instance, Holler & Snow (2020) found convergent results in an experimental setting where the real-world objects were handled physically whereas the pictures were manipulated on a computer screen. Our complimentary results were obtained in an experimental setting in which both real objects and pictures are physical manipulated. The combination of both results suggests that the difference between real world objects and object representations is robust in the case of diverse ways of delivering object representations.

We also found support for the hypothesis that the observed differences between real objects and representations of them are partly explained by differences in affordances. Taking the Word condition first, there it seems clear that the two main clusters map neatly onto a distinction between tools and non-tools, suggesting a simple conceptual bifurcation in the classification scheme. (We will return to some of the lower-level details after reviewing the other cluster analyses.)

One initial cue to what is driving the similarity judgments in the Picture condition is the low-level adjacency of spoon and lollipop: overall shape. The impression that in the Picture condition similarity judgments are shape driven is reinforced by the other adjacent sub-cluster: apple, pinecone, and seashell are all roundish, roughly fist-sized objects.

The most telling feature of the Object cluster are the tuples apple-lollipop-spoon and hammer-ruler-glove. The first suggests an eating-related cluster, with a specific “edible” category, and the second a hand-related cluster with a specific “employable” category.

There is also evidence here that many different factors influenced similarity judgments, and it must be acknowledged that we have not in the current set of experiments ruled out plausible alternatives to the affordance hypothesis. Hammer, ruler glove, tape, for instance, are closely related in all cases, indicating sensitivity to a tool category across conditions. In fact, hammer-ruler are adjacent in all cases. But these items are related in multiple

ways: the way they might be wielded (affordances); being tools (category); and being relatively elongated (shape). Similarly, there is evidence in these clusters for the influence of natural vs. artificial (e.g., apple, pinecone, seashell being closely related). This is why it is crucial to attend not just to local adjacency, but also to larger neighborhoods when discerning the dominant factor driving similarity judgment in a given condition. Hence, when the adjacent pair apple-lollipop are clustered with spoon, that suggests “edible” is a dominant consideration, whereas when apple-lollipop are clustered with pinecone-seashell this suggests that a “non-tools” category is dominant, despite the fact that edibility is clearly part of the overall judgment. Similar reasoning applies to the hammer-ruler dyad when it appears in the company of glove vs. tape; the former indicates hand-related affordances, the latter a more general tool category.

Cluster analyses are always exploratory, and we understand this is just one plausible interpretation among others of the relevant factors guiding the categorization of the different stimuli. These results therefore point to several possibilities for future work. Employing the same behavioral paradigm used here, further evidence for the influence of affordances on similarity judgments might be found using objects with different highly specific and salient affordances, such as specialized tools, pitting these against one another and/or against the sorts of more “everyday” objects that we used in the current case. Taking this notion further, one could attempt to manipulate the salience of the affordances by altering the context within which they are encountered, e.g., a generic lab vs. a workshop vs. a kitchen. Alternately, one could manipulate the state of the participants themselves, e.g., hungry vs. satiated. Evidence that the effect of affordances can be modulated by circumstance would further underscore the important role played by affordance perception in the first place (one wouldn’t expect such manipulations to modulate shape or category perception, for instance).

Moving into the domain of neuroimaging, it would be instructive to repeat the study done by Mur et al. (2013) using real objects as a contrast condition. They found, using a representational similarity analysis (RSA), that the similarity of medial temporal (MT) cortex activation in response to different objects closely matched the judged similarity of those objects, as measured by the paradigm employed here. But they used only pictures as stimuli. Given the observed difference in judged similarity using objects vs. pictures, a question arises: will MT activation look the same when using real objects, or different? If the same, what drives the different behavioral result? If different, what modulates MT activity so it differs when presented with real objects vs. pictures? Either outcome promises new insight into the relationship between brain and behavior.

Generally, we believe the examination of the differences in the experimental effects of real objects and of representations of objects is interesting in its own right, and also of practical relevance for a field largely dominated by the use of representations such as pictures as stimuli.

Acknowledgements

We are thankful to Krystyna B. Wiczerzak for her help in improving the figures of the paper. We are also thankful to Tony Chemero and The EMRG Lab (emrglab.org) for their helpful comments on different versions of this manuscript. This research was supported in part by a Canada Research Chair award to MLA, grant number SSHRC 950-231929, and by a Mitacs Research Training Award to REW, award number IT19642.

References

- Castellini, C., Tommasi, T., Noceti, N., Odone, F., & Caputo, B. (2011). Using object affordances to improve object recognition. *IEEE transactions on autonomous mental development*, 3(3), 207-215.
- Chao, L., Haxby, J. & Martin, A. (1999) Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature Neuroscience* 2, 913–919.
- Chemero, A. (2003). An outline of a theory of affordances. *Ecological Psychology*, 15(2), 181-195.
- Cisek, P. (2007). Cortical mechanisms of action selection: The affordance competition hypothesis. *Philosophical Transactions of the Royal Society B*, 362(1485), 1585–1599.
- Clark, A. (2015). *Surfing uncertainty*. London: Oxford University Press.
- Dotov, D. G., de Wit, M. M., & Nie, L. (2012). Understanding affordances: History and contemporary development of Gibson’s central concept. *AVANT. Pismo Awangardny Filozoficzno-Naukowej*, 28–39.
- Fajen, B. R., Riley, M. A., & Turvey, M. T. (2009). Information, affordances, and the control of action in sport. *International Journal of Sport Psychology*, 40, 79–107.
- Gibson, J. J. (1979/2015). *The ecological approach to visual perception: classic edition*. New York: Psychology Press. (originally published in 1979).
- Glerean, E., Pan, R.K., Salmi, J., Kujala, R., Lahnakoski, J.M., Roine, U., Nummenmaa, L., Leppämäki, S., Nieminen-von Wendt, T., Tani, P., Saramäki, J., Sams, M., Jääskeläinen, I.P. (2016) Reorganization of functionally connected brain subnetworks in high-functioning autism. *Human brain mapping*, 37(3), 1066-1079.
- Gomez, M. A., Skiba, R. M. & Snow, J. C. (2018). Graspable objects grab attention more than images do. *Psychological Science*, 29(2), 206-218.
- Grill-Spector, K. & Malach, R. (2001). fMR-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta psychologica*, 107(1-3), 293-321.
- Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, 293(5539), 2425-2430.
- Heras-Escribano, M. (2019). *The philosophy of affordances*. Cham: Palgrave Macmillan.
- Holler, D. E., Fabbri, S., & Snow, J. C. (2020). object responses are highly malleable, rather than invariant, with changes in object appearance. *Scientific Reports*, 10(1), 1-14.
- Kriegeskorte, N., Mur, M., & Bandettini, P. A. (2008). Representational similarity analysis-connecting the branches of systems neuroscience. *Frontiers in systems neuroscience*, 2, 4.
- Mur, M., Meys, M., Bodurka, J., Goebel, R., Bandettini, P. A., & Kriegeskorte, N. (2013). Human object-similarity judgments reflect and transcend the primate-IT object representation. *Frontiers in Psychology*, 4.
- Pezzulo, G., & Cisek, P. (2016). Navigating the affordance landscape: Feedback control as a process model of behavior and cognition. *Trends in Cognitive Sciences*, 20(6), 414–424.
- Romero, C. A., Compton, M. T., Yang, Y. & Snow, J. C. (2017). The real deal: Willingness-to-pay and satiety expectations are greater for real foods versus their images. *Cortex*, 107, 78-91.
- Snow, J. C., & Culham, J. C. (2021). The treachery of images: how realism influences brain and behavior. *Trends in Cognitive Sciences*, 25(6), 506-519.
- 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(1), 1-10.
- 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.
- Turvey, M. T. (1992). Affordances and prospective control: An outline of the ontology. *Ecological Psychology*, 4, 173–187.
- Valyear, K. F., Culham, J. C., Sharif, N., Westwood, D., & Goodale, M. A. (2006). A double dissociation between sensitivity to changes in object identity and object orientation in the ventral and dorsal visual streams: a human fMRI study. *Neuropsychologia*, 44, 218-228.
- Valyear, K. F., Gallivan, J. P., McLean, A. D., & Culham, J. C. (2012). fMRI repetition suppression for familiar but not arbitrary actions with tools. *Journal of Neuroscience*, 32, 4247-4259.
- Van Weelden, L., Maes, A., Schilperoord, J., & Cozijn, R. (2011). The role of shape in comparing objects: How perceptual similarity may affect visual metaphor processing. *Metaphor and Symbol*, 26(4), 272-298.
- Zhang, S., & Schmitt, B. (1998). Language-dependent classification: The mental representation of classifiers in cognition, memory, and ad evaluations. *Journal of Experimental Psychology: Applied*, 4(4), 375–385.