Title
Semantic Contamination of Visual Similarity Judgments

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 38(0)

Authors
Unger, Layla
Fisher, Anna

Publication Date
2016

Semantic Contamination of Visual Similarity Judgments

Layla Unger (LUnger@andrew.cmu.edu)
Carnegie Mellon University, Department of Psychology, 5000 Forbes Avenue
Pittsburgh, PA 15214 USA

Anna V. Fisher (Fisher49@andrew.cmu.edu)
Carnegie Mellon University, Department of Psychology, 5000 Forbes Avenue
Pittsburgh, PA 15214 USA

Abstract

The roles of semantic and perceptual information in cognition are of widespread interest to many researchers. However, disentangling their contributions is complicated by their overlap in real-world categories. For instance, attempts to calibrate visual similarity based on participant judgments are undermined by the possibility that semantic knowledge contaminates these judgments. This study investigated whether inverting stimuli attenuates semantic contamination of visual similarity judgments in adults and children. Participants viewed upright and inverted triads of familiar animals, and judged which of two test items looked most like the target. One test item belonged to the same category as the target, and one belonged to a different category. Test items' visual similarity to the target either corresponded or conflicted with category membership. Across age groups, conflicting category membership reduced accuracy and slowed reaction times to a greater extent in upright than inverted triads. Therefore, inversion attenuates semantic contamination of visual similarity judgments.

Keywords: semantic knowledge; visual similarity

Introduction

We perceive different things as related to each other in a variety of potentially overlapping ways. For instance, entities may be perceptually similar due to their shared perceptual features, or belong to the same semantic, "taxonomic" category. Many research endeavors and theoretical debates have focused on questions surrounding the influence of these relations on various facets of cognition, such categorization, inductive inference, memory encoding, and visual search, as well as the development and neural underpinnings of these processes (e.g., Deák & Bauer, 1996; Gelman & Markman, 1986; Konkle, Brady, Alvarez, & Oliva, 2010). However, perceptual and semantic relations commonly overlap amongst real-world entities. Moreover, these relations may interact during both learning and online processing of perceptual input. This interplay between semantic and perceptual relations severely complicates the study of questions about their respective contribution to cognition.

The complication of interest in the present paper is the fact that attempts to control for and manipulate perceptual similarity of real-world items independent of the semantic relations between them may be undermined by the contamination of similarity judgments by semantic category knowledge. Below we discuss this issue, and then present a study designed to provide a possible solution to this problem.

Measuring Perceptual Similarity

Researchers have used multiple approaches to measure and calibrate the perceptual similarity between stimuli, including using their own intuition (e.g., Fisher, 2011), collecting similarity judgments from adults (e.g., Deák & Bauer, 1996; Gelman & Markman, 1986), and, in developmental studies, collecting similarity judgments from children (e.g., Long, Lu, Zhang, Li, & Deák, 2012; Sloutsky & Fisher, 2004). The aim of these approaches is to assess visual similarity independent of semantic relatedness. For example, to calibrate stimuli in a match-to-sample task with triads consisting of a target, a perceptual match, and a semantic match, researchers may ask a separate sample of participants to judge the visual similarity of the target to each match item on a likert scale (e.g., Deák & Bauer, 1996; Gelman & Markman, 1986). Alternatively, researchers may calibrate such triads by asking participants to choose which match item looks most like the target in order to obtain ratios of the similarity of the perceptual match to the target versus the semantic category match to the target (e.g., Long et al., 2012; Sloutsky & Fisher, 2004).

The calibration of visual similarity based on participant similarity judgments is common to studies in many areas, such as memory, semantic knowledge, and semantic development (e.g., Blaye, Bernard-Peyron, Paour, & Bonthoux, 2006; Deák & Bauer, 1996; Gelman & Markman, 1986; Konkle et al., 2010). Intrinsic to this approach is the assumption that people can judge visual similarity without being influenced by the semantic knowledge.

However, this assumption may be unwarranted. Semantic knowledge may instead influence perceptual similarity judgments through any of multiple routes. Knowledge of semantic relationships between items may influence judgments of their similarity: 1) After perceptual similarity has been independently evaluated (Pylyshyn, 1999), 2) By feeding back into perceptual similarity evaluations (Lupyan, Thompson-Schill, & Swingley, 2010), or 3) By influencing the similarity of items' perceptual representations during prior learning (Goldstone, 1998; Goldstone, Lippa, & Shiffrin, 2001; O'Reilly, Wyatte, Herd, Mingus, & Jilk, 2013).

The degree to which any of these routes truly characterizes cognition is the subject of active research and debate (Chen & Proctor, 2012; Lupyan et al., 2010). Although an in-depth evaluation of this issue is beyond the scope of this paper, the brief overview presented here highlights the many ways in which semantic knowledge may contaminate perceptual similarity judgments. The purpose of the present study is therefore to assess the contamination of perceptual similarity judgments by semantic category knowledge, and test whether this contamination is attenuated by inversion.

The Present Study

The choice to test whether inversion attenuates the influence of semantic category knowledge on perceptual similarity judgments was motivated by numerous findings that rotation away from a canonical orientation impedes the identification of the category to which a familiar item belongs (e.g., Jolicoeur & Milliken, 1989; Lawson & Jolicoeur, 2003). Specifically, increasing misorientation both slows and increases errors for identifying an item's category label. This effect has been attributed to a process in which a perceived item must be mentally normalized to its canonical orientation before its category membership can be retrieved from memory (Lawson & Jolicoeur, 2003). Consequently, inversion may interfere with access to semantic category knowledge, and therefore attenuate the influence of such knowledge on perceptual similarity judgments.

In the present study, participants performed a match-to-sample perceptual similarity judgment task in which they were asked to choose which of two test items "looks most like" a target item for triads of items that were presented in both upright and inverted orientations on different trials. All items were pictures of familiar animals, such as "dog" and "pig". One test item belonged to the same category as the target (e.g., both were pigs), whereas the other test item belonged to a different category (e.g., dog). The visual similarity of the test items to the target was manipulated such that it either corresponded or conflicted with the category membership of the target (see Fig. 1).

We predicted that if semantic category knowledge influences perceptual similarity judgements, participants should choose the visual similarity match less accurately and more slowly when category membership and visual similarity were in conflict. Moreover, if inversion attenuates the influence of category knowledge on perceptual similarity judgments, less accurate and slower responses on conflict versus no-conflict trials should manifest to a greater extent when triads are upright than when they are inverted. To test whether this predicted pattern manifests across ages to whom the categories are familiar, we conducted this study with both a kindergarten-age and an adult sample.

Method

Participants

The total sample of 42 participants included 24 participants in each of two age groups: Kindergarten (M_{age} = 5.45 years, *SD* =0.43 years), and Adults. Kindergarten participants were recruited from schools in a middle-class, metropolitan area in a Northeastern US city, and Adults were recruited via Amazon Mechanical Turk. Adults were compensated at a rate of $5/hour for their participation.

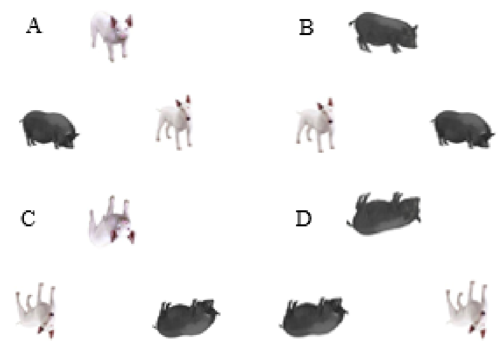

Figure 1. Examples of triads in each condition. A) Conflict, Upright; B) No Conflict, Upright; C) Conflict, Inverted; D) No Conflict, Inverted.

Materials

Both kindergarten and adult participants completed a Visual Similarity Judgment task, described below. Kindergarten participants viewed this presentation on a laptop computer, and made responses using a Cedrus RB-530 response box. To help kindergarten participants distinguish between the buttons they were instructed to use in the study (see Procedure), these buttons were given different-colored plastic covers. Adult participants viewed the presentation on Qualtrics, an online survey platform, and made responses using their personal keyboards. The Qualtrics version of the presentation was designed to record keyboard response times.

Visual Similarity Judgment Task This task consisted of triads of animal pictures presented on a computer screen that consisted of a Target item (e.g., a white pig) on the top, a Same Category test item (e.g., a black pig) on the bottom to one side, and a Different Category test item (e.g., a white dog) on the bottom to the other side. The pictures were photo-realistic images manipulated in graphics editing software to create two Semantic Conflict conditions: 1) No Conflict, in which the Same Category test item was visually similar to the target and the Different Category test item was visually dissimilar to the target, and 2) Conflict, in which the correspondence between category membership and visual similarity was reversed (see Fig. 1). Visual similarity was manipulated across several characteristics of the stimuli, including color, shape, and internal features.

Both No Conflict and Conflict triads were presented in two orientation conditions: Upright, and Inverted. The position of the two types of test items on the left or the right of the bottom of the screen was counterbalanced. Both, the Conflict condition and the Orientation condition were manipulated within participants.

The presentation included four practice trials, and 32 experiment trials. The practice trials consisted of a Conflict and a No Conflict triad presented in upright and inverted orientations, and the experiment trials consisted of eight Conflict and eight No Conflict triads presented in upright and inverted orientations. Each triad was presented once in upright and inverted orientations. The order of the experiment trials was pseudo-randomized such that different versions of a given triad did not appear consecutively, and such that no more than two triads in the same combination of visual similarity and orientation conditions appeared consecutively.

Procedure

Children Participants were tested individually in a quiet space. To begin, participants were seated in front of the laptop and button response box, and the first practice trial was displayed. Participants were told that they were going to play a game in which they decide which of two animals on the bottom of the screen looks like the animal on the top. They were further asked to use the buttons on the response box to indicate which animal they chose. Participants were then allowed to proceed through the practice and experiment trials at their own pace. The button instructions were repeated on subsequent trials if participants either failed to make a response for several seconds, or started to press the buttons quickly and randomly.

Adults Participants completed the task via Qualtrics. The version adults completed was identical to the version children completed, with the exception that participants were instructed to use "z" and "m" keyboard keys rather than the left and right buttons of a response box.

Results and Discussion

First, to ensure that participants understood that the purpose of the task was to identify the similarity rather than the category match, the accuracy with which participants in each age group in each condition chose the similarity match was compared to chance (i.e., .5). All contrasts revealed significantly above chance performance (*p*s $< .0001$).

To test the prediction that conflict between semantic and visual similarity would decrease accuracy and slow response times (RTs) for Conflict versus No Conflict trials in the Upright and not the Inverted condition, we analyzed the effects of the Semantic Conflict and Orientation factors on accuracy and RT using repeated measures ANOVAs. Specifically, for each outcome measure, we calculated each participant's mean score for the four combinations of conditions produced by our Semantic Conflict and Orientation factors, and submitted these mean scores to separate repeated measures ANOVAs for each age group.

Accuracy Analysis

For adult participants, this analysis revealed a main effect of Semantic Conflict ($F(1,23)=10.122$, $p=.004$, $\eta^2=.306$), and a main effect of Orientation ($F(1,23)=6.457$, $p=.018$, $\eta^2=.219$). More importantly, both main effects were qualified by a significant interaction ($F(1,23)=15.826$, $p=.001$, $\eta^2=.408$). To explore this interaction, we conducted t-tests comparing Upright versus Inverted RTs separately for the Conflict and No Conflict conditions. In the Conflict condition, adults were more accurate on Inverted ($M_{accuracy}$=89.96%) than on Upright ($M_{accuracy}$=82.29%) trials ($t(23)=3.680$, $p=.001$, Cohen's $d=1.54$), whereas in the No Conflict condition, adult accuracy did not significantly differ on Inverted and Upright trials ($t(23)=1$, $p=.328$) (see Fig. 2).

Kindergarten participants exhibited similar patterns of accuracy. The repeated measures ANOVA for this age group also revealed a main effect of Semantic Conflict ($F(1,23)=6.4$, $p=.019$, $\eta^2=.218$), though unlike adults, the main effect of Orientation did not reach significance ($F(1,23)=1.15$, $p=.295$). Moreover, like adults, the analysis with this age group revealed a significant interaction between Semantic Conflict and Orientation ($F(1,23)=5.522$, $p=.028$, $\eta^2=.194$). T-tests comparing Upright and Inverted trials for each Semantic Conflict condition revealed that Kindergarten participants were marginally more accurate on Inverted $M_{accuracy}$=84.90%) than on Upright ($M_{accuracy}$=77.08%) trials in the Conflict condition ($t(23)=2.01$, $p=.057$, Cohen's $d=.45$), and numerically more accurate on Upright than Inverted trials in the No Conflict condition ($t(23)=1.772$, $p=.09$) (see Fig. 2).

RT Analysis

Prior to calculating each participant's mean RT score, we filtered RTs to remove inaccurate trials and trials on which the participant responded either faster than 250 msec, or more than three standard deviations more slowly than the average RT for their age group.

For adults, this analysis revealed a main effect of Semantic Conflict ($F(1,23)=15.133$,$p=.001$, $\eta^2=.397$), and no main effect of Orientation ($F(1,23)=1.116$, $p=.302$). Critically, the main effect of Semantic Conflict was qualified by a significant interaction between this factor and Orientation ($F(1,23)=26.14$,$p=.000$, $\eta^2=.532$). In the Conflict condition, inverted trials yielded faster RTs than upright trials ($M_{Upright}$=1601ms, $M_{Inverted}$=1199ms, $t(23)=4.08$, $p<.0001$), whereas in the No Conflict condition, inverted trials yielded slower RTs ($M_{Upright}$=912ms, $M_{Inverted}$=1195ms, $t(23)$=-

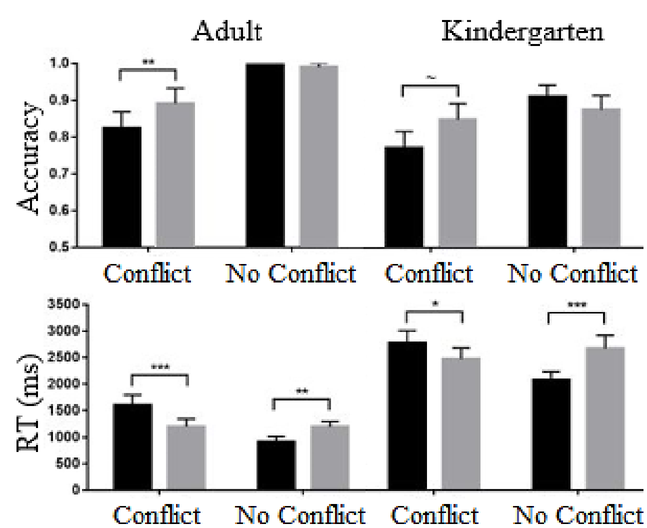

Figure 2. Accuracy and RT outcomes. Upright trials are depicted in black, and Inverted trials are depicted in gray. Error bars represent standard errors of the mean.

3.753,p=.001). These results are consistent with our prediction that semantic category knowledge influences perceptual similarity judgments such that participants are slower to judge perceptual similarity when it is in conflict with category membership, and that this influence is attenuated by inversion.

The pattern of results for kindergarten participants was similar to the pattern observed for adults. The repeated measures ANOVA revealed a main effect of Semantic Conflict ($F(1,23)=6.096$, $p=.021$, $\eta^2=.210$), and no main effect of Orientation ($F(1,23)=2.072$, $p=.164$). As in adults, the main effect of Semantic Conflict was qualified by a significant interaction between this factor and Orientation ($F(1,23)=23.105$, $p<.0001$, $\eta^2=.501$). Children's responses were also faster for inverted trials in the Conflict condition (M_{Upright}=2788ms, M_{Inverted}=2487ms, $t(23)=2.153$, $p=.042$), and slower for inverted trials in the No Conflict condition (M_{Upright}=2092ms, M_{Inverted}=2685ms, $t(23)=-4.376$, $p<.0001$).

The pattern of results in this experiment is broadly consistent with the prediction that decrements in accuracy and RT of visual similarity judgments due to conflict between category membership and visual similarity is attenuated by inversion. Both adults and children were less accurate and slower to respond on Upright than on Inverted trials in which category membership and visual similarity conflicted (though the effects on accuracy were marginally significant in children). In contrast, inversion did not improve performance in the No Conflict condition. Participants were instead similarly accurate on both upright and inverted No Conflict trials, and in fact slower on inverted than upright trials.

The observation of slower RTs in the inverted versus upright trials in the No Conflict condition was not specifically predicted in our semantic contamination hypothesis. This finding indicates that, in the absence of semantic conflict, inversion generally slows down responses in even non-semantic tasks such as the visual similarity judgment task used here. This possibility underscores the importance of the finding that inversion *speeds up* responses in the presence of semantic conflict.

General Discussion

Investigating the contributions of visual similarity and semantic relatedness is the focus of a wide range of research endeavors (Deák & Bauer, 1996; Gelman & Markman, 1986; Konkle et al., 2010; Lupyan et al., 2010; Sloutsky & Fisher, 2004). A critical component of this research is calibrating the degree to which the stimuli used in experiments are visually similar or semantically related. With respect to calibrating visual similarity, approaches taken to date in which researchers intuit or ask participants to judge visual similarity are undermined by the possibility that semantic knowledge contaminates visual similarity judgments. The purpose of this study was to examine this possibility and test whether semantic contamination is attenuated by inverting stimuli.

Our findings show that both adults and children show an effect of semantic conflict on visual similarity judgments that lowers accuracy and slows response times, and that is attenuated by inversion. Conversely, in the absence of semantic conflict, inversion slowed response times. Taken together, these findings suggest that semantic knowledge contaminates visual similarity judgments, and that inversion attenuates semantic contamination. Therefore, inverting stimuli for which visual similarity judgments are elicited provides a viable approach to calibrating stimuli for research investigating the role of visual similarity and semantic relatedness in various facets of cognition.

For example, recent cognitive neuroscience studies have investigated the degree to which knowledge about semantic relations is encoded as similar patterns of activity in brain regions involved in visual processing versus only in regions involved in subsequent amodal processing (Bruffaerts et al., 2013; Weber, Thompson-Schill, Osherson, Haxby, & Parsons, 2009). However, overlap between semantic relatedness and perceptual similarity renders it difficult to determine which of these sources is responsible for observed brain activity pattern similarities. For instance, a study conducted by Weber et al. (2009) found that the similarity of brain activity patterns in visual cortex evoked by viewing a set of animals correlated with behavioral judgments of their semantic similarity, but the fact that behavioral judgments of semantic similarity were in turn highly correlated with judgments of visual similarity renders it difficult to determine whether brain pattern similarity was related to semantic or visual similarity.

The approach introduced here to collecting visual similarity judgments that are relatively uncontaminated by semantic knowledge provides a route for attenuating such confounds. For example, the paradigm introduced in this study could be used to calibrate both items that are semantically but not visually similar, and items that are visually but not semantically similar, for studies that aim to disentangle the contributions of these forms of similarity to various facets of cognition and/or brain activity. The finding that this approach is similarly effective in both adult and child samples renders it viable for studies of both adult and developmental cognition. Therefore, the validation of the inversion approach demonstrated by the present study has the potential to support progress in a variety of lines of research.

Conclusions

This study demonstrated that conflict between visual similarity and semantic category membership slows and reduces accuracy of visual similarity judgments of upright, but not inverted stimuli for both adults and children. Therefore, this study suggests that semantic knowledge contaminates visual similarity judgments in adults and children, and that inversion attenuates this contamination.

Acknowledgements

This work was supported by a Graduate Training Grant awarded to Carnegie Mellon University by the Department of Education, Institute of Education Sciences (R305B040063) and by the James S. McDonnell Foundation 21st Century Science Initiative in Understanding Human Cognition–

Scholar Award (220020401) to the second author. We thank children, parents, teachers, and participating institutions for making this project possible: the Children's School at Carnegie Mellon University, Sacred Heart Elementary School, and Phipps Conservatory. Finally, we thank Anna Vande Velde and Manon Sohn for their help collecting data.

References

Blaye, A., Bernard-Peyron, V., Paour, J.-L., & Bonthoux, F. (2006). Categorical flexibility in children: Distinguishing response flexibility from conceptual flexibility. *European Journal of Developmental Psychology, 3*, 163-188.

Bruffaerts, R., Dupont, P., De Grauwe, S., Peeters, R., De Deyne, S., Storms, G., & Vandenberghe, R. (2013). Right fusiform response patterns reflect visual object identity rather than semantic similarity. *NeuroImage, 83*, 87-97.

Chen, J., & Proctor, R. W. (2012). Influence of category identity on letter matching: Conceptual penetration of visual processing or response competition? *Attention, Perception, & Psychophysics, 74*, 716-729.

Deák, G., & Bauer, P. (1996). The dynamics of preschoolers' categorization choices. *Child development, 67*, 740-767.

Fisher, A. V. (2011). Processing of perceptual information is more robust than processing of conceptual information in preschool-age children: Evidence from costs of switching. *Cognition, 119*, 253-264.

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition, 23*, 183-209.

Goldstone, R. L. (1998). Perceptual learning. *Annual review of psychology, 49*, 585-612.

Goldstone, R. L., Lippa, Y., & Shiffrin, R. M. (2001). Altering object representations through category learning. *Cognition, 78*, 27-43.

Jolicoeur, P., & Milliken, B. (1989). Identification of disoriented objects: effects of context of prior presentation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 200-210.

Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010). Conceptual distinctiveness supports detailed visual long-term memory for real-world objects. *Journal of Experimental Psychology: General*, *139*, 558-578.

Lawson, R., & Jolicoeur, P. (2003). Recognition thresholds for plane-rotated pictures of familiar objects. *Acta Psychologica, 112*, 17-41.

Long, C., Lu, X., Zhang, L., Li, H., & Deák, G. O. (2012). Category label effects on Chinese children's inductive inferences: Modulation by perceptual detail and category specificity. *Journal of experimental child psychology, 111*, 230-245.

Lupyan, G., Thompson-Schill, S. L., & Swingley, D. (2010). Conceptual penetration of visual processing. *Psychological Science*682-691.

O'Reilly, R. C., Wyatte, D., Herd, S., Mingus, B., & Jilk, D. J. (2013). Recurrent processing during object recognition. *Frontiers in psychology, 4*, 1-14.

Pylyshyn, Z. (1999). Is vision continuous with cognition?: The case for cognitive impenetrability of visual perception. *Behavioral and brain sciences*, *22*, 341-365.

Sloutsky, V. M., & Fisher, A. V. (2004). Induction and categorization in young children: a similarity-based model. *Journal of Experimental Psychology: General, 133*, 166-187.

Weber, M., Thompson-Schill, S. L., Osherson, D., Haxby, J., & Parsons, L. (2009). Predicting judged similarity of natural categories from their neural representations. *Neuropsychologia*, *47*, 859-868.