Journal of Experimental Psychology: General
2009, Vol. 138, No. 3, 400–415

0096-3445/09/$12.00 DOI: 10.1037/a0015835

Distance-Dependent Processing of Pictures and Words

Elinor Amit
New York University

Daniel Algom
Tel-Aviv University

Yaacov Trope
New York University

A series of 8 experiments investigated the association between pictorial and verbal representations and the psychological distance of the referent objects from the observer. The results showed that people better process pictures that represent proximal objects and words that represent distal objects than pictures that represent distal objects and words that represent proximal objects. These results were obtained with various psychological distance dimensions (spatial, temporal, and social), different tasks (classification and categorization), and different measures (speed of processing and selective attention). The authors argue that differences in the processing of pictures and words emanate from the physical similarity of pictures, but not words, to the referents. Consequently, perceptual analysis is commonly applied to pictures but not to words. Pictures thus impart a sense of closeness to the referent objects and are preferably used to represent such objects, whereas words do not convey proximity and are preferably used to represent distal objects in space, time, and social perspective.

Keywords: pictures, words, distance

Pictures and words have served humans to represent their environment and to convey event or object information since times immemorial. In antiquity, essential differences in the nature of the two tools of representation were recognized and accentuated. The biblical prohibition against pictorial representation of God is well known, although portrayals in words were endorsed if not actively sought (Halbertal & Margalit, 1992). In other times and contexts, the surrender of a name was a horrendous sin because names (but not pictures) were thought to reveal the essential (secret) nature of their referents (cf. Putnam, 1981). In modern psychology, the emphasis has shifted to explicating differences in processing between pictures and words. Amidst the many findings, the original role of pictures and words as means of representation has been somewhat underreported. In the present study, we focus again on the referential aspects of words and pictures.

Pictures and Words as Dedicated Means of Representation

Virtually all objects can be represented by words or by pictures, but the type of representation available or used first in individual instances is not random or incidental. We argue that each type of representation is best suited to represent the referent object under selective conditions. Optimal representation enables fast and less error prone processing (e.g., identification, categorization). What determines optimal representation? One clue comes from the aforementioned example of the contrast in representations of God. In monotheistic religions, God is transcendental. God is immeasurably remote and secluded from humans and their pursuits. Is pictorial representation forbidden because it violates this infinitely great distance? If so, does the picture-induced violation—forming a looming God—result from the concrete, contextual way in which pictures construe their referents? Under this hypothesis, words are well suited (hence used) to represent God because they convey information in a more generic, de-contextualized, even categorical fashion. Indeed, entire chapters in the Bible entail verbal depiction of God.[1] It is this idea of a distance-related difference between words and pictures that we elucidate and test in the present study.

Consequences of the Picture–Word Partition of Representation

The renewed stress on the quality of pictures and words as communicative media carries substantial ramifications for existing picture–word research. An immediate corollary is the importance

Elinor Amit and Yaacov Trope, Department of Psychology, New York University; Daniel Algom, Department of Psychology, Tel-Aviv University, Tel-Aviv, Israel.

We thank Yaniv Mama and Assi Schupak for help in running the experiments and for many helpful ideas.

Correspondence concerning this article should be addressed to Elinor Amit (c/o Yaacov Trope), Department of Psychology, New York University, 6 Washington Place, New York, NY 10003. E-mail: amitelin@gmail.com

[1] The Jewish prohibition against pronouncing the name of God does not violate the generic rule of permitting, and indeed encouraging, verbal descriptions of God. Note that the distinction we are drawing is between types of *construal*—a pictorial versus a verbal construal of deity—and not between single-item, particular prohibitions. All kinds of pictorial construal are prohibited, whereas virtually all kinds of verbal ones are permitted. Note also that the name of God is a proper noun, which concretizes God to the extent that a picture does. This might well be the root cause of the prohibition to pronounce the name. If so, this single-item verbal prohibition fits well within the generic scheme of distance-governed representations.

of the identity of the referent object and, consequently, of the particular representation used. In our approach it makes a noticeable difference whether a soccer ball (a culturally close object for a European observer) or the Tower of Pisa (a remote object for an Asian observer) comes as a picture or as a word. The variable of distance critically depends on the identity of the referent object. This distance, in turn, determines the optimal representation. The serviceability of pictures and words is distance determined.

Another corollary to reckon with is the nature of the respective mental processes. Pictures bear physical similarity to the referent objects, and as a result, pictures are subject to (largely) the same perceptual processes that are applied to the referent objects. Because pictures are such high-fidelity icons, they are also as specific, concrete, and particular as objects in nature are. Words do not bear physical similarity to the referent objects and are not subjected to perceptual analysis. A word denotes a larger set of referent objects than does a picture. Each word is actually a category, whereas a picture can often form a singular representation.

A final corollary refers to the ways in which concepts are organized in the cognitive system. According to one particular principle, concepts are organized in a hierarchical structure of categories. For example, the concepts of "bulldog," "dog," and "mammal" form a sequence of progressively larger categories in which each includes the previous ones (see Murphy, 2002, for a full discussion). Our approach poses an important limit on this method of organization: It applies to objects denoted by words but not to objects denoted by pictures. We return to discuss these implications of our approach.

To recap, picture- and word-symbols differ (a) in whether they share physical features with the referent object, (b) in the presence of perceptual analysis, (c) in the set-size of the represented referents, and (d) in the way that the represented stimuli are cognitively organized. Pictures and words thus carry extra-representation *qualia* or added psychological value above and beyond the representation of their referent objects. We suggest that these qualia are associated with the variable of psychological distance.[2]

Psychological Distance and Its Social–Cognitive Prospects

Psychological distance specifies the subjective closeness or remoteness of the experienced stimulus. Psychological distance is egocentric; that is, its reference point is the self here and now. An object is psychologically distant from the self to the extent that it is remote in time (future or past; temporal distance), place (spatial distance), or is experienced by others (unfamiliar events and stimuli or strangers; social–cultural distance). These (and other) dimensions of psychological distance correspond to the ways in which objects or events can be removed from the self (see Liberman & Trope, 2008, for a recent discussion and review). Because the various embodiments of psychological distance map onto the same fundamental variable, all are interrelated (i.e., can be congruent or incongruent with one another; Bar-Anan, Liberman, Trope, & Algom, 2007).

A recent theory of social cognition, construal level theory (Liberman, Trope, & Stephan, 2007; Trope & Liberman, 2003), links psychological distance from objects to the mental construal of those objects. According to the theory, objects or events can be mentally represented at multiple levels. At the one end, a high-level construal is an abstract representation that extracts the gist of event information. Such representations are general, goal-oriented, de-contextualized, and coherent. At the other end, a low-level construal is concrete and context bound. Such representations are goal-indifferent and include incidental features of the referent object or event. A major factor that determines the level of construal is the psychological distance—temporal, spatial, social, or cultural—of the referent from the self (Bar-Anan et al., 2007). Objects that are more distant along any of the dimensions will be represented at a higher construal level, because this abstract level captures those features of objects that remain invariant over distance-induced transformations. Research conducted within the framework of construal level theory (e.g., Henderson, Fujita, Trope, & Liberman, 2006; Nussbaum, Liberman, & Trope, 2006; Trope & Liberman, 2003; Wakslak, Trope, Liberman, & Alony, 2006) has indeed shown that the greater the psychological distance, the more likely events are to be represented at a higher level of construal.

Pictures Versus Words: The Psychological Distance Hypothesis

We propose that words are examples of high-level construal, whereas pictures are examples of low-level construal. As we recounted, each word is actually a fairly wide category. Naming a particular object (i.e., assigning a word to the object) is tantamount to placing that object in a semantic category. The verbal representation of an object always involves the bestowal of meaning. By their nature, such representations tend to be generic (carrying the gist of the object or event), de-contextualized, goal-relevant, and often abstract. By contrast, pictures are particular, contextual, and concrete representations (if they are not singular, they certainly denote smaller categories than do words). Pictures are also goal irrelevant in the sense that vital and marginal information are provided equal prominence in pictures. In general, the main defining features of high- and low-level construal apply well to words and pictures as specific cases (see Amit, Algom, Trope, & Liberman, 2009, for a full discussion of the parallelism between high- and low-level construal, on the one hand, and words and pictures, on the other hand).

The conviction that pictures are subject to perceptual analysis akin to that performed on the objects themselves is widely held (Glaser, 1992; Stenberg, 2006; see also DeLoache, Pierroutsakos, Uttal, Rosengren, & Gottlieb, 1998, and DeLoache, Pierroutsakos, & Uttal, 2003). However, the corollary of closeness to the referent is not widely recognized. Perception occurs only in the presence of the perceived object. This trivial stipulation is indispensable, or the associated stimulation (photons, sound pressure, or airborne taste/smell molecules) does not impinge on the sensory surface. Perception thus presupposes proximity of the referent object. Because pictures are perceived, they too convey a feeling of proximity. This sensation of closeness to the object or event (to the point at times of intimacy) is missing from verbal representations. Words often instantiate a certain

[2] We note that psychological distance is only one of many dimensions on which words and pictures differ. For instance, Huettig and McQueen (2007) have shown recently that language-mediated attention entails matches at perceptual (feature), phonological, and semantic levels of processing and that the type of mapping depends on the type of representation available (picture or word). Nevertheless, psychological distance is a major factor when generic effects of the type of representation are sought. Note also that psychological distance itself depends on many factors including the purpose and function of the representation.

distance between the object or event and its representation. Historically, the level of literacy was coextensive with the radius of control exercised in a given society (Coulmas, 2003), illustrating the function of words to convey information at a distance. Pictures instantiate the "here" and the "now," whereas words transcend perceptual bounds to convey information over space, time, and culture.

Arguably, pictures and words serve different cognitive functions. Words preserve the essential properties of a stimulus across momentary changes in appearance and through changes in space and time, functioning somewhat like perceptual constancies do (Rock, 1983). Words are capable of fulfilling this function due to their inclusive semantic nature. Because words are emancipated from the particularity of the situation, they are able to represent it across (temporal, spatial, social) distance. Pictures, by contrast, preserve the event in minute detail (relevant and irrelevant parts are given equal attention) to serve for immediate or early use.

An important prediction follows from this analysis. Given that (a) psychological distance determines the level of construal and that (b) pictures and words are species of low- and high-level construal, it follows that pictures best function when they represent near objects or events and that words best function when representing distal objects or events. Moreover, we expect that distance-congruent settings facilitate and distance-incongruent ones hinder performance. These predictions are tested in the present study for various instantiations of psychological distance.

The Present Research

The experiments focused on the task of speeded classification. The two variables manipulated were distance of the referent object from the observer (large, small) and mode or medium of presentation (word, picture). According to the present hypothesis, each mode of presentation conveys a value of implicit distance (word = large, picture = small). Consequently, each presented stimulus entailed two values of distance, objective and implied. The two values could correspond (both large or both small) or conflict. Specifically, congruent stimuli included pictures that represented proximal objects or words that denoted remote objects. The incongruent stimuli included pictures that represented distal objects and words that denoted proximal ones.

Experiments 1–6

The stimuli in the first six experiments included pairs of objects, one distal and the other proximal. Each object could appear as a word or as a picture. A single stimulus, a word or a picture, was presented on each trial. The participant's task was to decide which of the two possible objects appeared—regardless of mode of presentation. We predicted better classification performance when the medium of presentation was distance-appropriate than when it was distance-inappropriate.

We manipulated three species of distance: spatial, temporal, and social. Spatial distance was created through scenes rich in depth cues. This enabled us to place the stimulus at different perceived distances from the observer. Temporal distance was conveyed through history. The target object could be modern (e.g., a car, an object that is temporally proximal to the observer) or ancient (e.g., a carriage, an object that is temporally distant from the observer). Social distance was conveyed through everyday experience. The target object could belong in the culture of the observer (e.g., local currency) or to a different culture (e.g., foreign currency).

Experiments 7–8

The stimuli in Experiments 7 and 8 were picture–word compounds. One set consisted entirely of distal objects (i.e., both the picture and the word referred to distal objects) and another set consisted of proximal objects (again, both the picture and the word components denoted proximal objects). The task was speeded classification of the picture component in one condition (separately for distal and proximal objects) and speeded classification of the word component in another condition (again, separately for the two sets of objects).

Note that the picture–word compounds create other types of congruity. The word can be the name of the picture (congruent compound), belong in the category of the picture (category-congruent compound), or mismatch the picture in name and/or category (incongruent compound). The difference in classification performance between congruent and incongruent compounds is actually the Stroop effect (Stroop, 1935) with picture–word stimuli (MacLeod, 1991; Melara & Algom, 2003). The corresponding difference in categorization performance yields another species of the Stroop effect (Arieh & Algom, 2002). These Stroop effects were measured in the sets of proximal and distal objects with the word and with the picture component as the relevant dimension.

The prototypical finding in the Stroop literature is that performance with picture–word compounds depends on the task (Arieh & Algom, 2002; Glaser, 1992; M. C. Smith & Magee, 1980). For naming or classification, responses are faster to the words than to the pictures. Moreover, words intrude on performance with the pictures (the latter are named/classified faster with congruent than with incongruent stimuli) but not vice versa. These findings replicate the original Stroop effect by which selective attention is good for words but fails for pictures. However, when the task changes to one of categorization, the results are a mirror image of the former pattern. Pictures are categorized faster than words and they intrude on the categorization of the words. For categorization, selective attention is good for pictures, but it fails for words. According to the most common account (Glaser, 1992), words (but not pictures) can be named (read) without engaging the semantic system, affording words both a speed advantage and immunity from interference in tasks of naming, identification, or classification. Because pictures are closely associated with the semantic system (a picture cannot be named without accessing its meaning) and because categorization is a semantic task, pictures enjoy the same double advantage in tasks of categorization.[3]

We predicted that the manipulations of distance and medium in our study would modify these well-documented patterns. For classification (Experiment 7), we expected that the pictures would withstand interference from the words better when the pictures represented proximal items than when the pictures represented distal items. For categorization (Experiment 8), we expected that

[3] These findings from cognitive psychology refer to the early perceptual processing of word and picture stimuli. Note that the variables tested in this study exert their influence only after the words and the pictures have acquired full meaning so that they can function as representations for the referent objects. It is at this point that the main variable of psychological distance is consequential.

the words would sustain less interference from the pictures when the words denoted far objects than when the words denoted near objects. In general, the patterns of Stroop effects to emerge in the different sets (distal, proximal), for each of the two components as the relevant dimension for responding (picture, word), and in each of the two tasks (classification, categorization), served as a singularly potent test to support or to reject the present hypotheses.

Experiment 1

Suppose that a teenager from England classifies balls into either soccer balls or footballs (an American football, that is).[4] The balls appear either as pictures or as words. Does her or his classification performance differ with the two modes of representations? We claim that it does. We predicted that the teenager would respond faster to the picture of a soccer ball than to the word, SOCCERBALL, given the cultural proximity of this stimulus object. However, the reverse pattern was expected for the football: The teenager would respond faster to the word, FOOTBALL, than to the picture of a football because this object is more remote from the observer's cultural milieu. In our hypothesis, pictures are associated with closeness whereas words are associated with distance. We tested this prediction in Experiment 1.

Method

Participants. Fourteen Tel-Aviv University undergraduates participated in the experiment in partial fulfillment of course credit. All of the participants were native speakers of Hebrew.

Stimuli and apparatus. The stimulus objects were selected on the basis of social distance from the observer. There were three proximal objects (Israeli shekel, soccer ball, and the Israeli parliament building, the Knesset) and three distal objects from corresponding domains (American dollar, football, and the Tower of Pisa). On a trial, the observer was presented with a single stimulus, either a picture or a word. The words in this and all subsequent experiments were in Hebrew. The pictures were colored photographs recreated from standard computer files. The stimuli (the picture or the word) were presented at the center of the computer screen. The observer was seated approximately 60 cm from the screen. The width of the pictures was 7 cm (maximum) and their height varied between 3.5–4.5 cm. The words appeared in Hebrew font Miriam, size 26. In this and all subsequent experiments, the stimuli were generated by an IBM-compatible microcomputer (PC 486) and presented on a VGA color monitor set at a resolution of 600×800 pixels.

Design. There were three blocks of trials, separated by short breaks. Two stimulus items from the same domain, one proximal and one distal (e.g., an Israeli shekel and a U.S. dollar), appeared in a block. These objects appeared as a picture or as a word. Each of the four items thus produced (a picture of an Israeli shekel, a picture of a U.S. dollar, the word SHEKEL, and the word DOLLAR) were presented six times in random order, making for 24 stimuli per block. For the words, length was (very nearly) matched (both words were of four letters in the first block, of six letters in the second, and of three and four letters in the third).

On a trial, a single stimulus appeared (a picture or a word) and the participant's task was to decide, while timed, whether the object was a shekel or a dollar—regardless of mode of presentation (picture or word). The same speeded classification task was used in the other two blocks with the other pairs of near–far objects (soccer ball–football and Knesset–Tower of Pisa).

The participant responded by pressing one of a pair of lateralized keys corresponding to the respective stimuli. Stimuli were response terminated. The next stimulus appeared 500 ms after the response.

Procedure. The participants were tested individually in a dimly lit room. They were instructed to classify the presented item as one of two possible objects (e.g., football vs. soccer ball) and to ignore the medium of presentation (picture or word). They were encouraged to respond quickly but accurately. Prior to performing a particular block, the participant performed four trials of that block as practice. A block consisted of 24 trials, and the entire experimental session consisted of 72 trials. All the participants conducted the three blocks in the same order: shekel–dollar, soccer ball–football, and Knesset–Tower of Pisa. The order of trials within each block was randomized across participants.

Data analysis. Trials with a reaction time (RT) shorter than 150 ms or longer than 1,500 ms were discarded. Errors were negligible, amounting to less than 0.7% of all responses. The following analyses therefore focus on RTs for correct responses.

Results and Discussion

Figure 1 presents the mean latencies for the names and the pictures of the socially near and distant objects in Experiment 1. A Medium $\times$ Social Distance analysis of variance (ANOVA) (within-subjects analysis) yielded the predicted interaction between these two factors, $F(1, 13) = 7.63$, $MSE = 1{,}095.7$, $p < .05$, partial $\eta^2 = .41$. Post hoc comparisons using Fisher's least significant difference (LSD) test indicated that when the objects were represented verbally, the responses were faster to remote objects than to objects that belonged in the observer's cultural sphere (Ms = 598 and 634 ms, respectively, $p = .02$). When the objects were represented pictorially, responses to the objects belonging to the participant's culture were faster than responses to the objects that did not belong to the participant's culture, although this difference was not statistically dependable (Ms = 588 and 605 ms, respectively).

Notably, the pattern of results observed with the group data held in the data of virtually all of the individual participants. Of the 14 individual comparisons for words between distal and proximal objects, the data of 13 participants exhibited the group pattern of better performance in the distal condition. The chance probability for this outcome is less than .0001. Of the parallel comparisons for pictures, again, the data of only a single participant deviated from the predicted pattern of faster responding to proximal objects. The evidence for the hypothesized association of medium and psychological distance is solid at the level of individual performance as well.

In an auxiliary experiment performed on an independent group of 16 participants, we tested the present stimuli for familiarity. Quite expectedly, the proximal stimuli were judged more familiar

[4] Recall that American football is virtually nonexistent in the United Kingdom as well as in Israel and that, in any case, it is much lesser known than in the United States where it is a popular major sport.

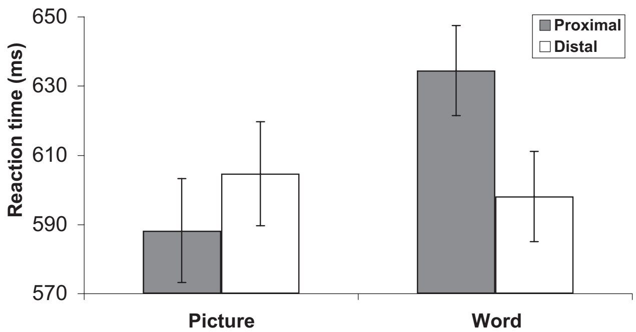

Figure 1. Mean reaction time (in milliseconds) for socially proximal and socially distal objects represented as pictures and as words, in Experiment 1. Error bars represent standard errors of the mean.

than the distal stimuli (*M*s = 5.75 and 6.6 on a 1 [*familiar*] to 7 [*unfamiliar*] scale), $F(1, 15) = 9.0$, $p < .01$. This result validates our choice of the two classes of stimuli. Notably, the pictures and the words within each class were judged equally familiar, but, significantly, there was not an interaction between medium (word, picture) and social distance for familiarity ($F < 1$). It is the identification of the referent via the different types of representation that is subject to the modifying effect of social distance, not other variables such as familiarity.

In sum, the presence of a congruity between distance and medium facilitated performance, and incongruity between medium and distance took a toll on performance. Note that the distance–medium contingency overrode any effect of familiarity on performance. Socially proximal pictures and words may well be more familiar to the participant than socially distal pictures and words, yet the decisive influence was that of the particular distance–medium combination. The results support our claim that pictures and words function in an optimal fashion at different values of psychological distance. As a result, they are not fully interchangeable means of representation.

Experiment 2

The logic of Experiment 1 is not limited to stimuli distinguished by social distance from the perceiver. In this experiment, we applied the design of Experiment 1 to stimuli that differ in their temporal distance from the perceiver. Half of the stimuli were modern (e.g., a car) and the other half were ancient (e.g., a carriage). We predicted that participants would respond faster to the pictures of temporally proximal objects than to the pictures of temporally distant objects, but that the opposite would hold for the names of the same objects at the respective time frames.

Method

Participants. The participants were 14 Tel-Aviv University undergraduate students who took part in the experiment in partial fulfillment of course credit. All of the participants were native speakers of Hebrew.

Stimuli and apparatus. The stimulus objects were selected on the basis of temporal distance from the observer. There were three modern objects (car, ball-point pen,[5] and electric lamp) and three parallel but more antique objects (carriage, quill pen, and oil lamp). Stimulus preparation and mode of presentation followed those of Experiment 1. The observers were seated approximately 60 cm from the screen. The width of the pictures varied between 3–16 cm, and their height varied between 7.5–10.5 cm. The words appeared in Hebrew font Miriam, size 26.

Design. The design was the same as that of Experiment 1. In each of three blocks, two objects from the same domain, one proximal and one distal (e.g., a car and a carriage), appeared in a random sequence as a picture or as a word. In this experiment, too, word length was (very nearly) matched (the words were of five and four letters in the first block, of six and five letters in the second, and both words were of five letters in the third). On each trial, a single item (a picture or a word) appeared and the participant's task was to decide, while timed, which of the two objects (e.g., car or carriage) she or he was watching. The participant responded by pressing one of a pair of lateralized keys standing for the stimuli.

Procedure. The procedure was the same as in Experiment 1. Again, the participants were instructed to classify the stimulus item by its meaning and to ignore the medium of presentation (picture or word). They were encouraged to respond quickly but accurately. Prior to performing a particular block, the participant performed four trials of that block as practice. All the participants performed the three blocks in the same order: car and carriage, quill pen and Pilot pen, and electric lamp and oil lamp. The order of trials within each block was random.

Data analysis. Again, only correct RTs between 150 and 1,500 ms were included in our analysis. Errors were negligible, amounting to less than 5% of all responses. The following analyses therefore focus on RTs for correct responses.

Results and Discussion

Figure 2 presents the mean latencies for the names and the pictures of the modern and ancient objects in Experiment 2. A Medium × Temporal Distance ANOVA (within-subjects analysis) yielded the predicted interaction between these two factors, $F(1, 13) = 6.13$, $MSE = 854.94$, $p < .05$, partial $\eta^2 = .32$. Responses were faster to pictures of modern objects than to those of ancient ones (*M*s = 556 and 583 ms, respectively, $p = .02$). The differences between the responses to words denoting antique objects and those denoting modern ones were in the opposite direction, although this difference was not dependable (*M*s = 576 and 588 ms, respectively).

The individual data were, again, fully consistent with the group data. For the pictures, all of the 14 individual comparisons between distal and proximal objects yielded the expected advantage for the latter. For the words, 11 out of the 14 individual comparisons showed the expected advantage for distal objects ($p < .001$).

As in Experiment 1, we checked the stimuli for familiarity in this experiment, too. The results for the present temporally defined items duplicated those obtained for the culturally defined items of Experiment 1: a difference in familiarity favoring modern over ancient items (*M*s = 4.85 and 6.8 on the same scale), $F(1, 15) =$

[5] We used the Hebrew term Pilot-Pen in Experiment 2 because in Hebrew this word has been overgeneralized from a brand name to a general name for a (modern) pen.

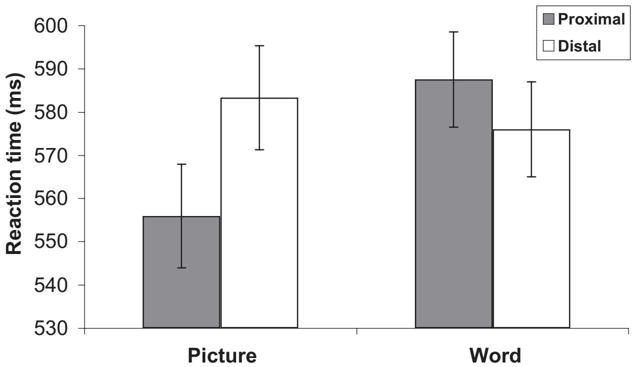

Figure 2. Mean reaction time (in milliseconds) for temporally proximal and for temporally distal objects, represented as pictures or as words, in Experiment 2. Error bars represent standard errors of the mean.

19.9, $p < .05$, but a lack of a difference across words and pictures in the same distance category as well as a lack of a Medium × Distance interaction for familiarity ($F < 1$ for both terms). Clearly, the present check validated the choice of items for Experiment 2 and showed that familiarity did not in any way constrain the generality of the results with respect to representation quality.

The results of Experiment 2 demonstrated that the medium–distance association governed the pattern of responses to objects varying in their temporal separation from the observer. For the simple effects, pictures best represent temporally proximal objects, whereas words best represent temporally distal objects. In the next experiment, we tested the medium–distance hypothesis with stimuli defined by the most natural embodiment of distance, spatial layout. Spatial distance has another advantage. The same stimulus can appear at close and large distance from the observer. When testing the effects of social and temporal distance, one must use different stimuli at the various values of distance. Spatial distance, by contrast, enables examination of the same stimulus presented at various distances from the observer.

Experiment 3

In Experiment 3, we tested responses to stimuli (pictures and words) presented in spatially near or far-off positions. As noted earlier, spatial distance is the most natural embodiment of distance with humans. We used scenes rich in depth cues to place the targets (a dog or a chair). The objects appeared close or distant as a picture or as a word. In order to enhance the illusion of depth while keeping the stimuli perceptually the same, the stimulus in the distal position was a bit smaller than the same stimulus in the proximal position. This adjustment, needed to preserve subjective size, is an important feature of perception (it was recently shown via functional magnetic resonance imaging that the retinotopic map in the primary visual cortex, V1, is already involved in the process; Murray, Boyaci, & Kersten, 2006). We predicted that responses would be faster for near pictures than for distal pictures and faster for distal words than for near words.

Method

Participants. The participants were 18 students from Tel-Aviv University who took part in the experiment in partial fulfillment of course credit. All of the participants were native Hebrew speakers.

Stimuli and apparatus. The background scenes were eight color photographs of outdoor scenery, rich in depth cues, entailing mountains, valleys, and rivers. Superimposed on this scenery was a large vertical arrow in green, placed at a near or at a far position in the scene (see Bar-Anan et al., 2007).

The stimuli were photographs of a dog and of a chair and the words DOG and CHAIR, each placed inside a green arrow. For proximal pictures, the width of the dog was 2.5 cm and its height was 1.8 cm, and the width of the chair was 2.5 cm and its height was 3.1 cm. The size of the distal pictures was 80% of the size of the proximal pictures. For words, the font size was 26 for the proximal words and 22 for the distal words. Their length was three and four letters. The observers were seated approximately 60 cm from the screen.

Design. There was a single block of trials with a dog and chair as the stimuli. Each stimulus appeared as a picture and as a word. There were a total of eight stimulus combinations: (a) a picture of a dog in the distance, (b) a picture of a dog near the observer, (c) the word DOG in the distance, (d) the word DOG near the observer, (e) a picture of a chair in the distance, (f) a picture of a chair near the observer, (g) the word CHAIR in the distance, and (h) the word CHAIR near the observer. Each of the eight combinations was presented 10 times with each of the eight background scenes, making for a total of 640 trials. On each trial, a single stimulus appeared on the background of a scenic picture, and the participant's task was to decide, while timed, if it was a dog or a chair—regardless of mode of presentation (picture, word) and spatial location (near, far).

The participant responded by pressing one of a pair of lateralized keys, one standing for dog and the other standing for chair. Stimuli were response terminated. The next stimulus appeared 500 ms after the response.

Procedure. The participants were tested individually in a dimly lit room. They were instructed to classify the stimulus item and to ignore the medium of presentation (and distance). The participants were encouraged to respond quickly but accurately. The procedure was the same as in Experiments 1 and 2. Prior to performing the task, the participant performed eight trials as practice. Classification was made by pressing one of two keys (*L* for a dog and *A* for a chair). The order of trials within the block was randomized across participants. As before, our analysis was based on correct RTs between 150 and 1,500 ms. Errors were negligible, amounting to less than 2% of all responses. The following analyses therefore focus on RTs for correct responses.

Data analysis. Again, trials with RTs shorter than 150 ms or longer than 1,500 ms were discarded. Errors were negligible, amounting to less than 0.5% of all responses.

Results and Discussion

Figure 3 gives the results of Experiment 3. An overall ANOVA revealed a significant interaction of spatial distance and medium, $F(1, 17) = 5.4$, $MSE = 981.45$, $p < .05$, partial $\eta^2 = .25$. For pictures, the responses were faster to near objects than to far objects (*M*s = 685 and 713 ms, respectively, $p = .02$). For words, the responses were faster to far objects than to near objects, although this difference was not statistically significant (*M*s = 768 and 775, respectively). In addition, we found a main effect for medium, such that participants responded faster to pictures than to

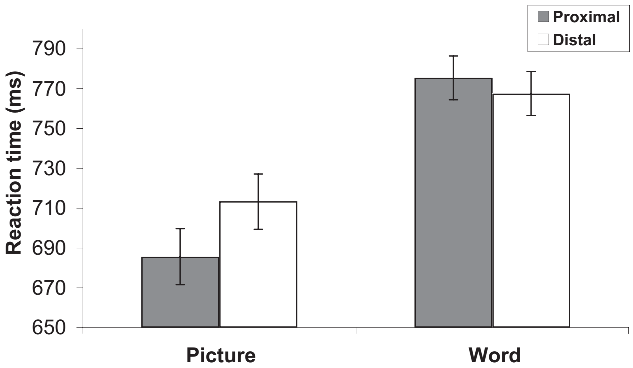

Figure 3. Mean reaction time (in milliseconds) for spatially proximal and spatially distal objects, represented as pictures (colored photographs) or as words, in Experiment 3. Error bars represent standard errors of the mean.

words (*M*s of 699 and 771 ms, respectively), $F(1, 17) = 22.71$, $MSE = 3{,}890$, $p < .001$.

The individual data were in full agreement with the data pooled over the participants. For words, the data of only 4 of the 18 participants deviated from the overall pattern of an advantage for distal objects ($p < .0001$). For pictures, the data of 5 participants showed a different outcome than the main pattern of advantage for proximal objects ($p < .001$).

The responses to pictures were faster than those to words. In tasks of naming, the responses to pictures are typically slower than those to words (MacLeod, 1991). However, the current task was manual classification, not verbal naming. This task requires semantic processing, a mode that gives rise to an advantage of pictures over words.

Earlier we recounted an important feature of the present procedure, conferring a unique advantage for hypothesis testing: The same stimuli appear at close and distal range. Now, it is important to realize that the sameness in question refers to subjective appearance. The stimuli must look the same at various apparent distances. In order to realize this goal, it is necessary to adjust their size; otherwise, the stimuli do not look the same. It was for this reason that we presented the distal stimuli at 80% their proximal size. Did the adjusted proximal and distal stimuli look alike? The results of an auxiliary experiment on a separate sample of 16 participants showed that they did. The participants rated the apparent size (Teghtsoonian, 1965) of the adjusted stimuli roughly the same (when presented singly on a 1–7 scale; *M*s = 4.16 and 4.28), $t(15) = 0.47$, $p > .05$. This result reinforces the power of our present manipulation.[6]

In sum, a main effect for medium was modified by distance such that the interaction of medium and spatial distance governed performance. This medium–distance interaction cannot be accounted for by mode of presentation (picture, word) or by stimulus variability (the same stimuli were presented close and far off).[7]

Experiment 4

In Experiment 4, we made use of the Ponzo illusion to create the feeling of distance. In the previous experiment, we manipulated psychological distance through realistic scenery that was rich in depth information. The minimalist device employed in the current experiment was complemented by equally minimalist outline drawings of the objects (rather than the photographs used in Experiments 1–3). This setup provides for a cross-validation of the observed medium–distance association by eliminating all sorts of extraneous variables.

Method

Participants. The participants were 16 students from Tel-Aviv University who took part in the experiment in partial fulfillment of course credit. All of the participants were native Hebrew speakers.

Stimuli and apparatus. A pair of vertically oriented converging straight lines was used to create the illusion of depth. These lines served to place the target stimuli (pictures and words) in a near (bottom) or a distal (top) position.

The objects were a bird and a pear (Block 1) and an ice cream and a lamp (Block 2). The pictures were outline drawings recreated from the International Picture Naming Project (IPNP; Szekely et al., 2004). For proximal pictures, height varied between 5–6 cm, and width varied between 2.5–6 cm. The size of distal pictures was 70% of the size of the proximal pictures. For words, font size was 26 for the proximal words and 22 for the distal words. In the first block they both were of five letters, whereas in the second block they were of five and three letters. A single word stood for each object. The observers were seated approximately 60 cm from the screen.

Design. There were two blocks of trials separated by a short break. Two objects appeared in each block. In the first block, the objects were a bird and a pear. In the second block, the objects were an ice cream and a lamp. Therefore, the eight stimulus combinations in the first block were (a) a picture of a bird in the distance, (b) a picture of a bird near the observer, (c) the word BIRD in the distance, (d) the word BIRD near the observer, (e) a picture of a pear in the distance, (f) a picture of a pear near the observer, (g) the word PEAR in the distance, and (h) the word PEAR near the observer. The eight combinations in the second block were (a) a picture of an ice cream in the distance, (b) a picture of an ice cream near the observer, (c) the words ICE CREAM in the distance, (d) the words ICE CREAM near the observer, (e) a picture of a lamp in the distance, (f) a picture of a lamp near the observer, (g) the word LAMP in the distance, and (h) the word LAMP near the observer. Each object was presented 10 times in a block, making for a total of 80 trials. On each trial, a single stimulus appeared and the participant's task was to decide,

[6] As is well known from elementary chapters on perception, a given object does not look the same at various implied distances created by depth cues. If no adjustments are made, the objects look much bigger at long distances (this trivial result was reproduced by our current participants when we presented them the same stimuli at 100% size at both values of distance). Therefore, adjustments must be introduced in order to preserve constant (perceived) size. This is the basis of the well-known perceptual constancies, in particular that of size constancy. In everyday life the adjustment is effected in a natural fashion by the diminishing visual angle of a receding object.

[7] In order to get an idea of the subjective scaling of the apparent size of words, we performed another preliminary experiment on a fresh sample of 16 participants. Our results showed that the participants judged the size-adjusted words at distal and proximal locations to be roughly equal (means of 3.81 and 4.09), $t(15) = 1.4$, $p > .05$.

while timed, if it was a bird or a pear (Block 1) or an ice cream or a lamp (Block 2), regardless of mode of presentation (picture, word) and spatial location (near, far). Participants responded by pressing one of a pair of lateralized keys, one standing for bird (ice cream) the other for pear (lamp). Stimuli were response terminated. The next stimuli appeared 500 ms after the response.

Procedure. The procedure was the same as in Experiments 1–3. The participants were instructed to classify the stimulus item and to ignore the medium of presentation (and the object's distance). Prior to performing the task, the participant performed four trials as practice. Classification was made by pressing a key (*L* for the bird in the first block and for the ice cream in the second block, and *A* for the pear in the first block and for the lamp in the second block). The entire experimental session consisted of 160 trials. All participants performed the two blocks in the same order: bird–pear first and then ice cream–lamp. The order of trials within each block was randomized across participants. As before, our analysis was based on correct RTs between 150 and 1,500 ms. Errors were negligible, amounting to less than 2% of all responses.

Results and Discussion

Figure 4 presents the results of Experiment 4. A Medium × Social distance ANOVA (a within-subjects analysis) yielded the predicted interaction between these two factors, $F(1, 15) = 6.3$, $MSE = 1{,}591.5$, $p < .05$, partial $\eta^2 = .28$. For words, the responses were faster in a distal position than in a near position (*M*s = 635 ms for the distal position and 668 ms for the proximal position, $p = .02$). For pictures, the responses were faster in a near position than in a distal position, although this difference was not statistically dependable (*M*s = 619 and 635 ms, respectively). In addition, we found a main effect for medium, $F(1, 15) = 7.2$, $MSE = 1{,}386.1$, $p < .05$. The participants responded faster to pictures than to words (*M*s = 626 and 650 ms, respectively), although the asymmetry was not as large as in Experiment 3.

Again, the features of the pooled data were exhibited in the data of virtually all of the individual participants. For words, the data of 13 out of the 16 participants showed the advantage of distal over proximal presentation. For pictures, there were merely two deviant patterns that did not yield the advantage for proximal presentation. The chance likelihood for these outcomes is negligibly small.

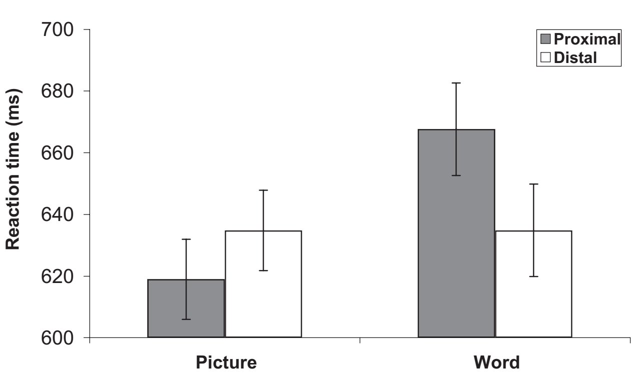

Figure 4. Mean reaction time (in milliseconds) for spatially proximal and spatially distal objects, represented as pictures (outline drawings) or as words, in Experiment 4. Error bars represent standard errors of the mean.

Clearly, the pattern imposed by the association of medium and psychological distance is a real constraint that affects the efficacy of the two tools of representation.

The pictures and words used in this experiment all looked approximately the same from the "near" and "far" positions (*M*s = 3.8 and 4.28), $t(15) = 1.9$, $p > .05$.

In sum, the current experiment indicates that even the most impoverished setup of implied distance suffices to engender the medium-dependent effect of distance. For the simple effects, the responses to pictures and to words were affected differently by their spatial location. Words benefited from spatial distance. There was an opposite trend for pictures, although this trend was not statistically dependable.

In all of the experiments reported so far, the task for the participant was speeded identification. The next experiment sought to extend the generality of our findings by showing that the association among pictures, words, and distance applies to speeded categorization too.

Experiment 5

In this experiment, the task for the observer was speeded categorization. The objects belonged in one of two categories: animals or pieces of clothing. Each object was presented as a picture or as a word. As in the previous experiment, two converging lines served to place the target stimuli near (bottom) or far away (top) from the observer. In each trial, the participant categorized the object as an animal or as a piece of clothing. We predicted that categorization would be faster for near pictures than for distal pictures and that it would be faster for distal words than for near words.

Testing the currently proposed dissociation between pictures and words with categorization is important for two reasons. First, we wished to test our hypothesis using another task beyond that of classification or identification. Second, the task of categorization is a semantic task with both pictures and words (tasks of naming and classification are semantic with pictures, but can be nonsemantic with words). Applying the task of categorization, we thus attempted to rule out a possible confound in the form of a Medium × Task interaction (recall that the Medium × Distance interaction is the cornerstone of the present study).

Method

Participants. The participants were 10 students from Tel-Aviv University who took part in the experiment in partial fulfillment of course credit. The participants were native Hebrew speakers.

Stimuli and apparatus. Two straight lines that converged toward the top were used to create the illusion of depth. They served to place the target stimuli (a picture or a word) in a near or distal position.

The objects belonged in one of two groups: animals (camel, bird, lion, and elephant) or pieces of clothing (tie, dress, pants, and jacket). Each of these objects appeared as a picture and as a word. The pictures were outline drawings taken from the IPNP database (Szekely et al., 2004). For proximal pictures, height varied between 5.5–6.5 cm and width varied between 4–7.5 cm. The size of the distal pictures was 70% of the size of the proximal pictures. For words, font size was 26 when proximal and 16 when distal. Their length was between three and six letters. The observers were seated approximately 60 cm from the screen.

Design. There was a single block of trials with an animal or a piece of clothing as the object. The task for the observer was to indicate whether the object was an animal or a piece of clothing. Because each object appeared as a picture and as a word in a proximal and in a distal position, there were 32 stimuli in all. Each stimulus appeared four times, making for 128 trials in all.

On a trial, a single stimulus appeared and the participant's task was to decide, while timed, whether it was an animal or a piece of clothing—regardless of mode of presentation (picture, word) and spatial location (near, far).

The participant responded by pressing one of a pair of lateralized keys, one standing for animals and the other standing for clothes. Stimuli were response terminated. The next stimulus appeared 500 ms after the response.

Procedure. The procedure was the same as in Experiments 1–4. The participants were instructed to categorize the objects and to ignore the medium of presentation and stimulus distance. Prior to performing the task, the participant performed four trials as practice. Categorization was made by pressing a key (*1* for an animals and *Z* for a piece of clothing). The order of trials within the block was random and different for each participant. As before, our analysis was based on correct RTs between 150 and 1,500 ms. Errors occurred at a low rate and consisted of less than 10% of all responses.

Results and Discussion

Figure 5 gives the results of Experiment 5. The interaction of medium and distance, $F(1, 9) = 5.5$, $MSE = 844.83$, $p < .05$, partial $\eta^2 = .38$, confirmed that distance had opposite effects on pictures and on words. For pictures, the responses were faster to near objects than to far objects (*M*s = 618 and 648 ms, respectively, $p = .04$). For words, the responses to distal objects were faster than those to proximal objects, although this difference was not significant (*M*s = 732 and 746 ms, respectively). In addition, we found a main effect for medium. Responses were faster to pictures than to words (*M*s = 633 and 739 ms, respectively), $F(1, 9) = 59.24$, $MSE = 1{,}889.64$, $p < .001$.

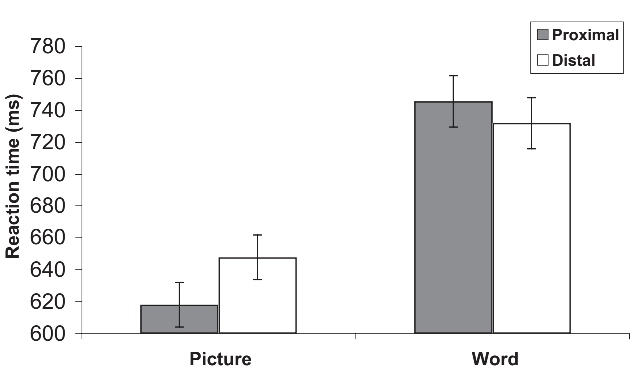

Figure 5. Mean reaction time (in milliseconds) for categorization of spatially proximal and spatially distal objects, represented as pictures (outline drawings) or as words, in Experiment 5. Error bars represent standard errors of the mean.

When examining the data of each of the individual participants, one observes largely the same features documented in the pooled data. For words, the data of only 2 individuals deviated from the general pattern (advantage of proximal over distal presentation; $p < .001$). For pictures, the individual data were equivocal, with 4 participants deviating from the expected advantage of proximal presentation.

The results of this experiment suggest that the medium–distance association holds for the task of categorization. Categorization, unlike classification, entails extensive semantic processes, namely, the placing of discriminably different objects under a common rubric. The current experiment shows that when people are asked to perform such semantic tasks, their performance is still affected by the way that the to-be-categorized objects are presented.

Experiment 6

Experiment 6 was a control experiment designed to rule out alternative artifactual explanations for the results of Experiments 4–5. In these experiments on spatial distance, the same object appeared in a near and in a distal position, but its size differed. We presented somewhat smaller stimuli in distal locations to increase the illusion of depth. Could this difference in size determine the results? In other words, is it possible that mere difference in size, rather than perceived distance, produced the observed pattern?

In order to address this concern, Experiment 6 manipulated the sizes of the objects of Experiment 4, without creating any illusion of depth or distance. The stimuli of Experiment 4 were presented singly without any context, notably without the frame of the Ponzo lines. The participant's task was to classify the presented object (and ignore medium of presentation and size).

Method

Participants. The participants were 10 Tel-Aviv University students who took part in the experiment in partial fulfillment of course credit. All were native Hebrew speakers.

Stimuli and apparatus. The stimuli and apparatus were the same as in Experiment 4, with a single notable exception: The stimuli were presented at the center of the screen *without* the Ponzo-like converging lines. The observers were seated approximately 60 cm from the screen.

The stimulus objects were a bird and a pear (Block 1) and an ice cream and a lamp (Block 2). For big pictures, height varied between 5–6 cm, and width varied between 2.5–6 cm. The size of the small pictures was 70% of the size of the big pictures. For words, font size was 26 for the large words and 22 for the small words. The objects, a picture or a word, appeared at the center of a computer screen.

Design. The design was the same as in Experiment 4. There were two blocks of trials separated by a short break. Two objects appeared in each block. In the first block, the objects were a bird and a pear. In the second block, the objects were an ice cream and a lamp. Each item was presented 10 times in a block, making for 80 trials in all. On a trial, a single stimulus appeared and the participant's task was to decide, while timed, if it was a bird (ice cream) or a pear (lamp), regardless of mode of presentation (picture, word).

The participant responded by pressing one of a pair of lateralized keys, one standing for the bird (ice cream) the other for the pear (lamp). Stimuli were response terminated. The next stimuli appeared 500 ms after the response.

Procedure. The procedure was the same as in Experiment 4. The participants were instructed to classify the stimulus item and to ignore the medium of presentation and the size of the object. Prior to performing the task, the participant performed four trials as practice. Classification was made by pressing a key (*L* for the bird in the first block and for the ice cream in the second block, and *A* for the pear in the first block and for the lamp in the second block). The experimental session consisted of 80 trials. All the participants performed the two blocks in the same order: bird–pear first and then ice cream–lamp. The order of trials within each block was randomized across participants. Our data analysis was based on correct RTs between 150 and 1,500 ms.

Results and Discussion

Figure 6 gives the results of Experiment 6. RTs were somewhat faster to large objects than to small objects (*M*s = 453 and 464 ms, respectively), $F(1, 9) = 3.7$, $MSE = 263.88$, $p = .09$, partial $\eta^2 = .00001$. This was true for pictures (*M*s = 452 and 463 ms for large and small size, respectively) as well as words (*M*s = 454 and 464 ms for large and small size, respectively).

The results of the individual participants were uneventful. All of the individual data reproduced precisely the pattern observed with the data pooled over the participants.

The fact that size had the same effect on pictures and words indicates that it is psychological distance rather than mere size that was responsible for the interaction with medium observed in all the experiments of this research.

Experiments 7 and 8

In Experiments 7–8, the stimuli were picture–word compounds. When people are asked to name (identify or classify) the word or the picture of a picture–word compound, they can attend selectively to the word but cannot attend selectively to the picture (Arieh & Algom, 2002; Glaser & Dungelhoff, 1984; Glaser & Glaser, 1989; MacLeod, 1991; M. C. Smith & Magee, 1980; see also Damian & Bowers, 2003, for a recent review). People do not suffer in their reading performance from the irrelevant pictures: They name a word superimposed on a conflicting picture as quickly and accurately as they do one superimposed on a matching picture. However, when people name the pictures, selective attention fails. They name a picture with a conflicting word more slowly than they do one in which the word is the name of the picture (M. C. Smith & Magee, 1980). Consequently, there is a Stroop effect for naming pictures, but none for naming words.

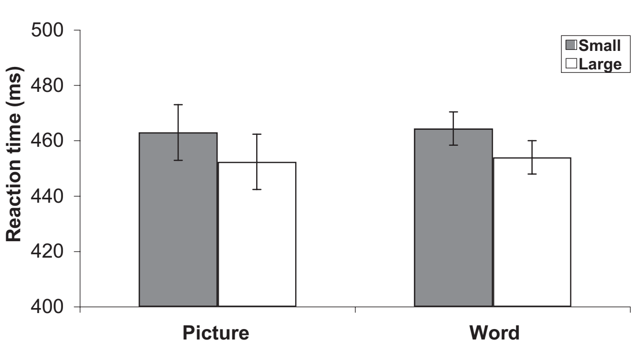

Figure 6. Mean reaction time (in milliseconds) for small and large objects, represented as pictures (outline drawings) or as words, in Experiment 6. Error bars represent standard errors of the mean.

When the task changes to that of categorization, the outcome is reversed. People do not suffer in their categorization performance with pictures with the irrelevant words: They categorize the pictures at the same speed regardless of whether the word belongs to the category as the picture or to a different category. For words, matching pictures facilitate and conflicting pictures impede categorization (Arieh & Algom, 2002; M. C. Smith & Magee, 1980). Therefore, there is a Stroop effect for categorizing words, but none for categorizing pictures.

Does psychological distance modify these patterns? We suggest that it does. The final pair of experiments, ordered by the type of the distance dimension used, tested this prediction. In Experiments 7 and 8 we tested how these patterns are affected by social (Experiment 7) or spatial (Experiment 8) distance.

In Experiment 7, we investigated the effect of social distance, creating two different sets of picture–word compounds. In one set, the words and the pictures belonged in the participant's culture. In the other set, the words and the pictures belonged in another culture. The participant's task was to categorize the words in one session and to categorize the pictures in another session. On the basis of the categorization results mentioned, we predicted that social distance would modify this standard pattern. We expected the Stroop effects for words to be smaller in the distal than in the proximal set for two interrelated reasons. Because words are processed more efficiently when distal, they command good selective attention in that position and are more immune to interference. In a complementary fashion, because pictures are processed less efficiently when distal, their ability to interfere with words is diminished. The same two factors conspire to enhance the interference for the categorization of words in the proximal set.

Experiment 7

The participants were asked to categorize pictures (in one condition) or words (in another condition) of picture–word compounds into religious or nonreligious items. In one set, the compound (i.e., all components) belonged in the Jewish–Israeli milieu (the participant's culture), whereas in the other set the items belonged in the Christian–American culture (a foreign culture). The compounds within each set were either congruent (the word and picture belonged in the same category) or incongruent (the word and the picture belonged in different categories). The difference in responding between the congruent and incongruent compounds gauged the quality of selective attention to the task-relevant component. Because the task was categorization, this Stroop effect is expected only in performance with the words. In the present experiment, we made a more fine-tuned prediction: The Stroop effect for words would be larger in the proximal than in the distal set.

Method

Participants. The participants were 20 students from Tel-Aviv University who took part in the experiment in partial fulfillment of course credit. All of the participants were native Hebrew speakers.

Stimuli and apparatus. The stimuli were selected on the basis of social distance from the observer. There were four proximal stimuli and four distal ones. Within each value of social distance, the stimuli were divided into religious and nonreligious items. For proximal stimuli, the religious items were a yarmulke and a menorah, and the nonreligious items were a soccer ball and an Israeli shekel. For the distal stimuli, the religious items were a Christmas tree and a cross and the nonreligious items were an American dollar and a football. Within each set, the four words and the four pictures were combined in a factorial fashion, making 16 compounds. Of these compounds, eight were congruent (the picture and the word belonged in the same category) and eight were incongruent (the picture and the word belonged in different categories). The words appeared on the top of or just below the pictures. Because each compound was presented in both arrangements, there were 32 compounds in all.

The participants were seated 60 cm from the screen. The stimuli were presented at the center of the screen. The pictures were colored photographs recreated from standard computer files. For pictures, height varied between 4.5–9 cm and width between 3.5–6.5 cm. The words were written in Hebrew font Miriam, size 26.

Design and procedure. For each set (proximal, distal), there were two blocks of trials, one for categorization of the pictures and the other for categorization of the words. Within each set, the order of the two blocks was randomized across subjects. In each block, the 32 picture–word compounds were presented twice each, making for 62 trials per block. The order of stimuli within a block was random and different for each participant.

On each trial, a picture–word compound appeared and the participant's task was to decide, while timed, whether the target feature was a religious item. The participants were instructed to attend to the relevant component (picture, word) and to ignore the irrelevant component. They were encouraged to respond quickly but accurately. Responses were made by pressing one of a pair of lateralized keys standing for the respective categories (*L* for religious and *A* for nonreligious). Stimuli were response terminated. The next stimulus appeared 500 ms after the response. Procedure and data analysis were the same as in the previous experiments.

Results and Discussion

Figure 7 presents the Stroop effect (the difference in performance between congruent and incongruent compounds) for proximal and distal pictures and words in Experiment 7. As can be seen in Figure 7, for words, the Stroop effect was larger in the proximal set than in the distal set (*M*s = 95 and 25 ms, respectively). Post hoc comparisons using Fisher's LSD test indicated that the differences between the Stroop effect for proximal words and distal words was reliable ($p < .05$). For pictures, the Stroop effect was larger in the distal set than in the proximal set, although this difference was not statistically reliable (*M*s = 15 and 4 ms, respectively). The difference in the Stroop effect for pictures and words as a function of distance was significant, $F(1, 19) = 4.09$, $MSE = 9{,}979$, $p = .057$.

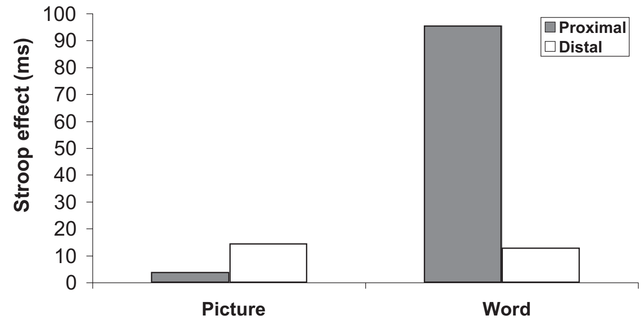

Figure 7. Stroop effects (in milliseconds) for categorization of proximal and distal objects, represented in pictures and in words, in Experiment 7.

For the absolute RTs we found that participants categorized pictures faster than words (*M*s = 598 and 704 ms, respectively), $F(1, 19) = 101.79$, $MSE = 13{,}325.13$, $p < .001$.

For the individual data, the same features documented in the pooled data were present. For words, the data of only 5 out of 20 individuals deviated from the general pattern (larger Stroop effect for proximal than distal presentation; $p < .001$). For pictures, the individual data were more variable, with those for 8 participants violating the expected advantage for proximal presentation.

For this experiment, too, a postexperimental check on familiarity confirmed that proximal items were more familiar than distal items (*M*s = 5.6 and 6.65, respectively), $F(1, 15) = 8.2$, $p < .05$, and that there was not an interaction between relevant mode for responding and distance ($F < 1$).

In sum, selective attention for words was better when they were standing for distal than for proximal stimuli. When categorizing words, the participants were better able to ignore irrelevant pictorial information if the words were representing socially distal objects. Distance helped the words to overcome the interference from pictures. In contrast, distance did not reliably improve selective attention to the pictures.

Experiment 8

In Experiment 8, we manipulated spatial distance. A pair of vertically oriented converging straight lines was used to create the illusion of depth. These lines served to place the target stimuli (picture–word compounds) in a near (bottom) or a distal (top) position. The compounds were either congruent (the word and picture refer to the same concept) or incongruent (the word and the picture refer to different concepts). The difference in response speed between the two types of compounds documents the quality of selective attention to the task-relevant item. The participants were asked to name the picture (in one condition) or read the word (in another condition) of the picture–word compounds. Past research suggests that the Stroop effect (the difference in response speed to congruent and incongruent compounds) is large for pictures but much smaller for words (Arieh & Algom, 2002; see MacLeod, 1991, for a review). This pattern granted, we still expected smaller Stroop effects for proximal pictures than for distal pictures for two reasons. Because pictures are better processed when proximal, they command attention in a powerful fashion in that position. In complementary fashion, because words are processed less efficiently when proximal, their ability to interfere with the pictures is diminished.

Method

Participants. The participants were 22 students from Tel-Aviv University who took part in the experiment in partial fulfillment of course credit. All of the participants were native speakers of Hebrew.

Stimuli and apparatus. Two straight lines that converged toward the top were used to create the illusion of depth in a Ponzo-like setup. They served to place the target stimuli in a bottom (near) or in a top (distal) position. The stimulus objects were picture–word compounds, such that the words were embedded inside the picture (see Figure 8). There were two objects, a vase and a banana. The pictures were outline drawings recreated from the IPNP. For proximal pictures, height varied between 4.5–5 cm and width varied between 4.5–5 cm. The size of distal pictures was 70% of the size of the proximal pictures. For words (font Miriam), size was 20 for the proximal words and 12 for the distal words. The two words were combined in a factorial fashion with the two pictures. Each compound appeared in either a proximal location or a distal location, thus creating eight distinct stimuli. Participants were seated 60 cm from the screen.[8]

Design and procedure. There were two blocks of trials. For half of the participants, the task in the first block was to name the picture and that in the second block was to name the word while ignoring the picture. For the other half of the participants, the order of blocks was reversed.

There were four stimulus items: the word VASE embedded inside a picture of a vase, the word VASE embedded inside a picture of a banana, the word BANANA embedded inside a picture of a banana, and the word BANANA embedded inside a picture of a vase. These four picture–word compounds appeared either in a proximal or in a distal location, creating eight different combinations. Each combination appeared four times in a block, making 32 trials for a block and a total of 64 trials for the experiment. The order of trials within each block was randomized across participants.

In each trial, a word embedded in a picture appeared in either a proximal location or a distal location. The participant's task was to decide, while timed, whether the object depicted by the relevant component (word, picture) was a vase or a banana and to ignore the distractor. The participant responded by pressing one of a pair of lateralized keys, one standing for a vase and the other for a banana. Stimuli were response terminated. The next stimuli appeared 500 ms after the response.

Prior to performing the task, the participant performed four trials as practice. Participants were encouraged to respond quickly and accurately. Again, correct RTs between 150 and 1,500 ms were considered in our analysis. Errors were negligible, amounting to less than 6% of all responses. The following analyses therefore focus on RTs for correct responses.

Figure 8. An example of the picture–word compound stimuli used in Experiment 8. The vase shows the Hebrew word for banana.

Results and Discussion

Figure 9 presents the Stroop effect for naming words and pictures in near and far-off positions in Experiment 8. For pictures, the Stroop effect was larger for distal than for proximal stimuli (*M*s = 30 and –9 ms, respectively). Post hoc comparisons using Fisher's LSD test indicated that the difference in interference to proximal and distal pictures was reliable ($p < .05$). For words, the pattern of the Stroop effect was reversed, such that interference was larger for proximal than for distal stimuli, although this difference was not significant (*M*s = 45 and 28 ms, respectively). The main result is the difference in the Stroop effect for pictures and words as a function of distance, $F(1, 21) = 6.1$, $MSE = 1{,}419.66$, $p < .05$.

For the absolute RTs, we found that participants named distal objects faster than they named proximal objects (*M*s = 633 and 666 ms, respectively), $F(1, 21) = 19.68$, $MSE = 2{,}801.67$, $p < .001$. In addition, participants named pictures faster than they named words (*M*s = 598 and 704 ms, respectively), $F(1, 21) = 47.7$, $MSE = 10{,}379.85$, $p < .001$.

The individual data were fully consistent with the group data. For pictures, the data of only 2 individuals deviated from the general pattern (larger Stroop effect for distal than proximal presentation; $p < .001$). For words, the individual data were equivocal, with those of 7 participants deviating from the expected disadvantage of proximal presentation.

Taken together, the results of Experiment 8 show that interference to naming a picture is affected by psychological distance. Selective attention to pictures was better in a proximal location. When naming pictures, the participants were better able to ignore irrelevant verbal information when the pictures represented spatially proximal objects. Proximity helped pictures overcome the interference from words. In contrast, selective attention to words was not better in the proximal location than in the distal location.

An unexpected finding was that words suffered interference from pictures, in both the proximal and distal conditions. This finding can be explained by the fact we asked participants to identify objects manually and not orally. It is possible that manual

[8] The stimuli presented in this and all the other experiments are available upon request from the authors.

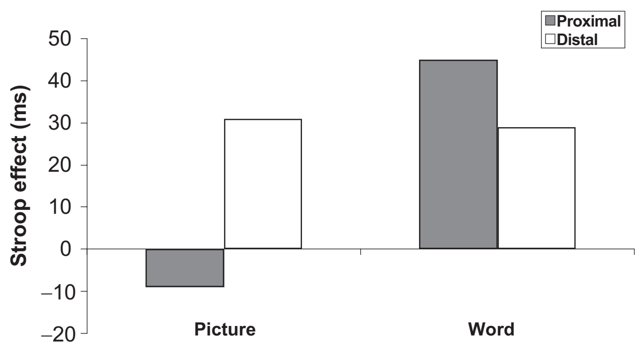

Figure 9. Stroop effects (in milliseconds) for naming proximal and distal objects, represented in pictures and in words, in Experiment 8.

identification eliminates the advantage of words over pictures in the naming task by disabling the direct route between the graphemes seen and the phonemes spoken. Most important, the interference to word reading was affected by perceived distance in a way consistent with our hypothesis.

General Discussion

The surface differences between pictures and words are obvious. Pictures are spatial representations that bear physical similarity to the referent objects, whereas words are strings of signs, often processed sequentially, that represent the object by convention. Nevertheless, the two are alternative ways of conveying object or event information; to wit, "images and verbal processes are . . . alternative coding systems or modes of symbolic representation" (Paivio, 1971, p. 8). In everyday life, you can read about a certain traffic accident in your newspaper or watch it on the evening news on your TV set. In scientific reporting, editors prefer one or the other means of conveying the data, guiding prospective authors to produce one or the other representation. In this study, we demonstrated that each representation is associated with a distinct set of stimulus conditions. We suggest that words best represent stimuli that are removed from the observer in space, time, society, or culture. Pictures, by contrast, best represent proximal objects in all of these respects. In a commensurate manner, pictures convey a sense of proximity to the represented stimuli that words do not. These claims are supported by the converging results of eight experiments.

Cross-validation support for the present conclusions is provided by a meta-analysis of effect size performed following the procedures recommended by Johnson and Eagly (2000). We conducted two separate analyses; one analysis concerned pictures and the other, words. The data of Experiments 1–5 were included in the analysis. Effect sizes (d) were computed as the standardized difference between the mean of the proximal condition and the mean of the distal condition. The results of these analyses are shown in Table 1 (pictures) and Table 2 (words).

For pictures, the meta-analysis yielded a weighted mean (d) of 0.4, with a 95% confidence interval (CI) that did not include zero (95% CI = 0.07, 0.74). Evaluating significance, the unit-normal z value for the weighted mean (d) was 2.4 ($p < .001$). This outcome indicates that, across our studies, the responses to proximal pictures were faster than to distal pictures. The results of the meta-analysis also indicate that effect sizes were homogenous across the experiments, $Q(4) = 0.58$, $p = .96$.

For words, the meta-analysis yielded a weighted mean (d) of 0.34, with a confidence interval that again did not include zero (95% CI = 0.008, 0.67). Evaluating significance, the unit-normal z value for the weighted mean (d) was 2.01 ($p < .05$). This outcome indicates that, throughout the study, the reactions to distal words were faster than to proximal words. The results of the meta-analysis also indicate that effect sizes were homogenous across the experiments, $Q(5) = 1.86$, $p = .75$.

Taken together, the present findings show that performance was better to pictures when they represented proximal objects and to words when they represented distal objects than with the reverse arrangement. Inappropriate or suboptimal combination of medium and distance took a toll on performance.

Implications of the Representation Approach for Picture–Word Research

We mentioned several consequences of the present approach for studying picture–word stimuli. We expand on two issues of theoretical import here.

The importance of the referent. We recounted the importance of the identity of the referent object in our approach. This perspective contrasts with that espoused in the great majority of contemporary picture–word studies in which it little matters what particular pictures and words are presented or recalled (once the categorical membership of the picture and the word have been specified). Whether the words recalled in a particular experiment were TREE or CLOCK or the recalled pictures were those of a house or a car little mattered (as long as they were not "lure" items or some other well-defined targets; e.g., Dodson & Schacter, 2001, 2002). Often, the particular item-specific responses are not reported, only the overall performance is. In a similar vein, whether the word presented in a particular experiment was BANANA and the picture was of a book or whether the reverse arrangement was presented was not of consequence.

This indifference to content characterizes much picture–word research, whether conducted within the picture–word interference paradigm (e.g., Damian & Bowers, 2003), the Stroop task (Arieh & Algom, 2002; Roelofs, 2003; see also Melara & Algom, 2003), cross-modal priming (e.g., Bajo, 1988), or memory (e.g., Dodson & Schacter, 2001; Weldon & Roediger, 1987). In our approach, by

Table 1
Meta-Analysis of Effect Sizes (ds) for Comparison Between Proximal and Distal Objects Represented in Pictures of Experiments 1–5 and Mean Effect Sizes Across All Measures

Experiment	Effect size (d)
Experiment 1	0.27
Experiment 2	0.51
Experiment 3	0.42
Experiment 4	0.27
Experiment 5	0.63
Weighted mean d	0.41
z value of weighted mean d	2.40^{*}

* $p < .05$.

Table 2
Meta-Analysis of Effect Sizes (ds) for Comparison Between Proximal and Distal Objects Represented in Words of Experiments 1–5 and Mean Effect Sizes Across All Measures

Experiment	Effect size (*d*)
Experiment 1	0.74
Experiment 2	0.25
Experiment 3	0.06
Experiment 4	0.49
Experiment 5	0.23
Weighted mean *d*	0.34
z value of weighted mean *d*	2.01*

* $p < .05$.

contrast, the identity of the picture and the word *is* of consequence. Our perspective is best understood by recognizing the different goal of much of current research.

The goal in virtually all contemporary picture–word studies was that of studying processes of speech production and not that of studying pictures and words as means of representation. Commensurate with that goal, words were seldom tested (because words are already lexical entities), and interest focused on features of lexicalization with pictures. Even for the latter, the focus was not on pictorial representations per se, but on the referent objects, with the pictures merely serving as convenient experimental surrogates to the "things" themselves (Jescheniak, Hantsch, & Schriefers, 2005). In contrast again, our interest focuses on the pictures and the words themselves as symbols carrying information. In this respect, we demonstrated that these two modes of construal have crystallized into dedicated means of representation.

Limitations on the hierarchical structure of categories. We also recounted the different ways in which the same objects are represented as pictures and words, noting that a hierarchical organization applies only to words. In this respect, it is often overlooked that Rosch's (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976) influential taxonomy of objects into superordinate, basic, and subordinate categories applies to words and not to pictures. There cannot be a picture of the superordinate category of "furniture." And, it is arguable that there exists a picture to represent the basic level term "chair." The most impoverished outline drawing of a chair is still specific to an extent that compromises the attempt at a pictorial representation of this category. The subordinate level in Rosch's taxonomy is the most concrete, yet it is a fairly wide category, too. Consider the subordinate term SPORTSCAR. Any outline drawing would reveal details (number of doors, size, shape, insignia, or plate) that are not carried by the word SPORTSCAR. Even a subordinate term cannot capture the concreteness of a picture.

In retrospect, the use of pictures in the study by Rosch et al. (1976; to show that objects of basic and subordinate categories had similar shapes) was unfortunate. It served to exacerbate the problem of demarcation between words and pictures as (different) means of representation. Pictures have been used extensively in more recent studies of categorization, mainly to elucidate the perceptual implications of level of specification. Participants were asked to verify, while timed, that a picture was a superordinate-, basic-, or subordinate-level stimulus after listening to the name of the pertinent category (Jolicoeur, Gluck, & Kosslyn, 1984; Murphy & Brownell, 1985; Murphy & Wisniewski, 1989; E. E. Smith, Balzano, & Walker, 1978; see, Murphy, 2002, for a review). One should be extremely careful not to take the pictures presented in these studies to stand for a category (they properly functioned as members of a category). Unlike a word, a picture cannot represent a category; it can merely be an (prototypical) exemplar of a category. As David Hume has famously observed, one can conceive of the category of trees, but one can have an image of only a particular tree.

Pictures and Words in Priming and Memory: New Insights

Due to their abstract nature, words are able to survive surface changes and serve as potent conveyors of meaning. Due to their concrete nature, pictures are more context bound and are more limited means of information transmission. The priming study by Durso and Johnson (1979) nicely demonstrates this contrast in a naming task. The participants named a picture target preceded by the same picture or by a word that was the picture's name. In the complementary condition, the participants named a word preceded by itself or by a pictorial representation of the word. Word primes facilitated the subsequent naming of the target regardless of the modality of the target (picture or word). By contrast, picture primes facilitated the naming of picture targets but not of word targets. Word naming (reading) was not facilitated by prior exposure to a picture. The authors concluded that pictures are like "words in context—they generate . . . specific representations . . . [whereas a] word . . . activates . . . a large set of semantic features" (Durso & Johnson, 1979, p. 457).

The same resistance to change affects memory for pictures but not for words. The study by Mintzer and Snodgrass (1999), which probed memory for concepts presented in different formats, nicely illustrates this contrast. The participants studied a list of objects presented either as pictures or as words. On a subsequent recognition test, the studied stimuli were presented to the participants in the same medium or in the other medium. The results showed that studied words survived changes in their format during test, whereas studied pictures were vulnerable to changes in their appearance.

How does the current conception fare with the one formulation in the cognitive psychology literature that is devoted explicitly to exploring the differences between pictures and words? We conclude by contrasting the current approach with dual coding theory (Paivio, 1986).

Dual Coding Theory

According to Paivio's dual coding theory (Paivio, 1986, 2007), there are two separate representation systems. One is specialized for representing information conveyed by spatial, nonverbal stimuli, and the other is specialized for representing information conveyed verbally. Words are initially represented by the verbal system, and scenes and pictures are initially represented by the nonverbal or imagery system. In subsequent processing, each stimulus can be coded by the other system as well, hence the possibility of dual coding. The theory does not posit mutually exclusive processing of words and pictures: Each stimulus can be encoded and processed by each of the two systems. The major explanatory variable in the theory is the imaginability of the input. Pictures are readily imaginable. Concrete words are also easily imaginable; therefore, such words are likely to be encoded by the image system as well as by the verbal system. In contrast, abstract words are likely to be encoded by the verbal system only. This difference explains the superiority in memory of concrete over abstract words, and, often, of pictures over words.

The concrete–abstract distinction in dual coding theory is well taken. Note however that the concrete–abstract distinction in dual coding theory is orthogonal to the picture–word distinction as means of representation. Words can be as concrete as pictures (although pictures cannot be as abstract as some words). When concreteness (or abstractness) is comparable, processing is comparable according to dual coding theory. In our approach, a word is always more abstract than is a picture and hence differences are always expected.

Concrete words, such as the word TABLE, and pictures, such as an outline drawing of your table, are close relatives in dual coding theory. Both stimuli are dual coded by both systems. Consequently, comparable performance is expected with such stimuli. In the current approach, there is a difference between the word TABLE and the picture of a table, although admittedly both items are highly imaginable. The reason is that the word TABLE encompasses a larger category (innumerable individual tables including your desk) than does the picture of a table. As a result, pictures are categorized better than words (including concrete, readily imaginable words), are less valuable primes, are more vulnerable to incidental changes, and impart a sense of proximity. All of these features are difficult to account for by dual coding theory.

Conclusion

Pictures are concrete representations that bear physical resemblance to their referent objects, whereas words are abstract representations that carry the essence of that object. Pictures should therefore best function when they represent near objects or events and words best function when they represent distal objects or events. Consistent with this argument, the results of the present experiments show that the processing of pictures is facilitated when they are psychologically near and the processing of words is facilitated when they are psychologically distant. These results were obtained with various psychological distance dimensions (spatial, temporal, and social), different tasks (classification and categorization), and different measures (speed of processing and selective attention).

References

Amit, E., Algom, D., Trope, Y., & Liberman, N. (2009). "Thou shalt not make unto thee any graven image": The distance-dependence of representation. In K. D. Markman, W. M. P. Klein, & J. A. Suhr (Eds.), *The handbook of imagination and mental simulation* (pp. 53–68). New York: Psychology Press.

Arieh, Y., & Algom, D. (2002). Processing picture–word stimuli: The contingent nature of pictures and of word superiority. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 221–232.

Bajo, M. T. (1988). Semantic facilitation with pictures and words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 579–589.

Bar-Anan, Y., Liberman, N., Trope, Y., & Algom, D. (2007). Automatic processing of psychological distance: Evidence from a Stroop task. *Journal of Experimental Psychology: General, 136,* 610–622.

Coulmas, F. (2003). *Writing systems.* Cambridge, United Kingdom: Cambridge University Press.

Damian, M. F., & Bowers, J. S. (2003). Locus of semantic interference in picture–word interference tasks. *Psychonomic Bulletin & Review, 10,* 111–117.

DeLoache, J. S., Pierroutsakos, S. L., & Uttal, D. H. (2003). The origins of pictorial competence. *Current Directions in Psychological Science, 12,* 114–118.

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

Dodson, C. S., & Schacter, D. L. (2001). "If I had said it I would have remembered it": Reducing false memories with a distinctiveness heuristic. *Psychonomic Bulletin & Review, 8,* 155–161.

Dodson, C. S., & Schacter, D. L. (2002). When false recognition meets metacognition: The distinctiveness heuristic. *Journal of Memory and Language, 46,* 782–803.

Durso, F. T., & Johnson, M. K. (1979). Facilitation in naming and categorization of repeated pictures and words. *Journal of Experimental Psychology: Human Learning and Memory, 5,* 449–459.

Glaser, W. R. (1992). Picture naming. *Cognition, 42,* 61–105.

Glaser, W. R., & Dungelhoff, F. J. (1984). The time course of picture–word interference. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 640–654.

Glaser, W. R., & Glaser, M. O. (1989). Context effects in Stroop-like and picture processing. *Journal of Experimental Psychology: General, 118,* 13–42.

Halbertal, M., & Margalit, A. (1992). *Idolatry.* Cambridge, MA: Harvard University Press.

Henderson, M. D., Fujita, K., Trope, Y., & Liberman, N. (2006). Transcending the "here": The effect of spatial distance on social judgment. *Journal of Personality and Social Psychology, 91,* 845–856.

Huettig, F., & McQueen, J. M. (2007). The tug of war between phonological, semantic and shape information in language-mediated visual search. *Journal of Memory and Language, 57,* 460–482.

Jescheniak, J. D., Hantsch, A., & Schriefers, H. (2005). Context effects on lexical choice and lexical activation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 905–920.

Johnson, B. T., & Eagly, A. H. (2000). Quantitative synthesis of social psychological research. In C. M. Judd & H. T. Reis (Eds.), *Handbook of research methods in social and personality psychology* (pp. 496–528). London: Cambridge University Press.

Jolicoeur, P., Gluck, M. A., & Kosslyn, S. M. (1984). Pictures and names: Making the connection. *Cognitive Psychology, 16,* 243–275.

Liberman, N., & Trope, Y. (2008, November 21). The psychology of transcending the here and now. *Science, 322,* 1201–1205.

Liberman, N., Trope, Y., & Stephan, E. (2007). Psychological distance. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (Vol. 2, pp. 353–383). New York: Guilford Press.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109,* 163–203.

Melara, R. D., & Algom, D. (2003). Driven by information: A tectonic theory of Stroop effects. *Psychological Review, 110,* 422–471.

Mintzer, M. Z., & Snodgrass, J. G. (1999). The picture superiority effect: Support for the distinctiveness model. *American Journal of Psychology, 112,* 113–146.

Murphy, G. L. (2002). *The big book of concepts.* Cambridge, MA: MIT Press.

Murphy, G. L., & Brownell, H. H. (1985). Category differentiation in object recognition: Typicality constraints on the basic category advantage. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11,* 70–84.

Murphy, G. L., & Wisniewski, E. J. (1989). Categorizing objects in isolation and in scenes: What a superordinate is good for. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 572–586.

Murray, S. O., Boyaci, H., & Kersten, D. (2006). The representation of perceived angular size in human primary visual cortex. *Nature Neuroscience, 9,* 429–434.

Nussbaum, S., Liberman, N., & Trope, Y. (2006). Predicting the near and distant future. *Journal of Experimental Psychology: General, 135,* 152–161.

Paivio, A. (1971). *Imagery and verbal processes.* Oxford, United Kingdom: Holt, Rinehart, & Winston.

Paivio, A. (1986). *Mental representations.* New York: Oxford University Press.

Paivio, A. (2007). *Mind and its evolution: A dual coding theoretical approach.* Mahwah, NJ: Erlbaum.

Putnam, H. (1981). *Reason, truth and history.* Cambridge, United Kingdom: Cambridge University Press.

Rock, I. (1983). *The logic of perception.* Cambridge, MA: MIT Press.

Roelofs, A. (2003). Goal-reference selection of verbal action: Modeling attentional control in the Stroop task. *Psychological Review, 110,* 88–125.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8,* 382–439.

Smith, E. E., Balzano, G. J., & Walker, J. (1978). Nominal, perceptual, and semantic codes in picture categorization. In J. W. Cotton & R. L. Kaltzky (Eds.), *Semantic factors in cognition.* Hillsdale, NJ: Erlbaum.

Smith, M. C., & Magee, L. E. (1980). Tracing the time course of picture–word processing. *Journal of Experimental Psychology: General, 109,* 373–392.

Stenberg, G. (2006). Conceptual and perceptual factors in the picture superiority effect. *European Journal of Cognitive Psychology, 18,* 813–847.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 38–48.

Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., et al. (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory and Language, 51,* 247–250.

Teghtsoonian, M. (1965). The judgment of size. *The American Journal of Psychology, 78,* 392–402.

Trope, Y., & Liberman, N. (2003). Temporal construal. *Psychological Review, 110,* 403–421.

Wakslak, C. J., Trope, Y., Liberman, N., & Alony, R. (2006). Seeing the forest when entry is unlikely: Probability and the mental representation of events. *Journal of Experimental Psychology: General, 135,* 641–653.

Weldon, M. S., & Roediger, H. L. (1987). Altering retrieval demands reverses the picture superiority effect. *Memory & Cognition, 15,* 269–280.

Received January 22, 2008
Revision received January 21, 2009
Accepted January 30, 2009 ■