Journal of Experimental Psychology:
Learning, Memory, and Cognition
2002, Vol. 28, No. 1, 221–232

0278-7393/02/$5.00 DOI: 10.1037//0278-7393.28.1.221

Processing Picture–Word Stimuli: The Contingent Nature of Picture and of Word Superiority

Yoav Arieh
John B. Pierce Laboratory and Yale University

Daniel Algom
Tel Aviv University

Participants named (Experiments 1–2) or categorized (Experiments 3–4) the picture or the word of picture–word compounds that varied in the relative saliency of the 2 components and in the correlation between them over the experimental trials. Picture–word interference (PWI) was gauged through Stroop and Garner effects. PWI was found to be malleable; its magnitude and very presence depending lawfully on the contextual variations introduced. The contingent nature of PWI is a fact to be reckoned with by theorists of picture–word processing.

When people are asked to name the word or the picture of picture–word compounds (see Figure 1), (a) they name the words faster than the pictures and (b) their performance with the pictures suffers interference from the irrelevant words. When people are asked to categorize the same stimuli, the outcome is the mirror image of the pattern obtained with naming. People categorize the pictures faster than the words, and their performance with the words suffers interference from the irrelevant pictures. In our study, we show word superiority in naming and picture superiority in categorization to be contingent phenomena depending on a few variables of context. Trifling changes in stimuli and design suffice to eliminate both types of picture–word interference (PWI; Lupker & Katz, 1981, 1982; Smith & Magee, 1980). Our means to show the pliability of this species of the Stroop phenomenon (Stroop, 1935) was performing a full Garnerian analysis (Garner, 1962, 1974; Garner & Felfoldy, 1970; Pomerantz, 1983) of PWI.

The Garner Paradigm Applied to Naming and Categorizing Picture–Word Compounds

A single stimulus—a word embedded within a picture—was shown on a trial. In the filtering condition, the participants were timed as they classified values on one dimension (e.g., whether the word was *table* or *apple*) while ignoring trial-to-trial variation on the second, irrelevant dimension (e.g., whether the picture was that of a table or an apple). In the baseline condition, participants again classified values on the criterial dimension of word but values on the irrelevant dimension were held constant (e.g., the irrelevant picture was always that of a table). In the correlated dimensions conditions, the task again was to classify the stimuli on the criterial dimension. However, the components varied in a perfectly correlated manner throughout the sequence of presentations. Thus, all stimuli either matched (*table* embedded in a table and *apple* in an apple) or mismatched, respectively, for positively and negatively correlated sets of stimuli.

The ability to attend selectively was measured by comparing performance in the baseline condition (in which the irrelevant dimension was held constant) with performance in the filtering condition (in which the two dimensions varied in an orthogonal manner). If performance in the filtering condition equaled that in the baseline condition, then selective attention was good. The parity implied that the participants were able to focus on the criterial dimension without suffering distractions from irrelevant variation. Alternatively, if performance in the filtering condition was worse than that in the baseline condition, then selective attention failed. The difference in performance between the filtering and baseline tasks defines Garner interference (Pomerantz, Pristach, & Carson, 1989). Pairs of dimensions that produce substantial Garner interference are called *interacting dimensions*, whereas pairs of dimensions that do not produce Garner interference are called *separable dimensions*.

We also derived the corresponding Stroop effects by calculating the difference in performance between congruent trials (in which the word named the picture) and incongruent trials (in which the word and picture conflicted). Three measures of Stroop congruity commensurate with this definition ensue. In the filtering task, the difference in performance between congruent and incongruent stimuli yielded one measure of the Stroop effect. Another within-condition measure of the Stroop effect was derived in the same way from the baseline task. Finally, the difference in performance between the positively correlated dimensions task (in which all of the stimuli were congruent) and the negatively correlated dimensions task (in which all of the stimuli were incongruent) yielded a third, between-conditions measure of the Stroop effect. For categorization, dimensions and values were defined in terms of category, not individually. Thus, in the baseline task it was the category of the irrelevant dimension that was held constant (i.e., stimuli could vary individually as long as they belonged in the

Yoav Arieh, John B. Pierce Laboratory, Yale University; Daniel Algom; Department of Psychology, Tel Aviv University, Tel Aviv, Israel.

This article is based on a doctoral dissertation carried out by Yoav Arieh and supervised by Daniel Algom at the Department of Psychology, Bar-Ilan University, Ramat Gan, Israel. This study was supported by Israeli Trustees Foundation Grant DC-9/18.

Correspondence concerning this article should be addressed to Yoav Arieh, John B. Pierce Laboratory, Yale University, 290 Congress Avenue, New Haven, Connecticut 06519. E-mail: yarieh@jbpierce.org

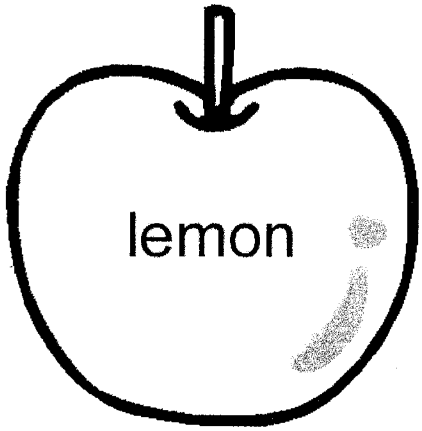

Figure 1. An example of the picture–word compounds used in the study. Note that the stimulus depicted is incongruent (for naming) because the word (*lemon*) does not name the picture (apple). Categorywise, the compound is congruent because the picture and the word are drawn from the same semantic category (fruits). In the study the words were printed in Hebrew, the participants' native language.

same invariant category). Similarly, congruent and incongruent stimuli were defined in terms of shared or conflicting category of the picture and the word. The point to note is that for the first time we derived Stroop and Garner effects for categorization of picture–word stimuli.

We demonstrated the malleability of both superiority effects by manipulating context. Dimensional discriminability (Algom, Dekel, & Pansky, 1996; Melara & Mounts, 1993) and dimensional correlation (Arieh & Algom, 1997; Dishon-Berkovits & Algom, 2000) were the factors of context considered in our study.

Dimensional Discriminability and Dimensional Correlation—Neglected Factors in Processing of Picture–Word Stimuli

Dimensional Discriminability

Dimensional discriminability specifies the size of the psychological differences separating stimulus values along a dimension. It is measured by the speed and accuracy needed to identify stimuli along a dimension as they alternate randomly from trial to trial. Discriminability is matched when the time and accuracy needed to discriminate between the pictures of a table and a chair (with word held constant) are equal to the time and accuracy needed to discriminate between the words *table* and *chair* (with picture held constant). Note that discriminability is an intradimensional index, measured separately for each of the tested dimensions.

Our reading of the literature showed that as a rule words were named faster than pictures (Fraisse, 1969; Glaser, 1992; Glaser & Dungelhoff, 1984; Rosinsky, Golinkoff, & Kukish, 1975; Smith & Magee, 1980; Theios & Amrhein, 1989) but pictures were categorized faster than words (Glaser & Glaser, 1989; Smith & Magee, 1980; Vinson & Theios, 1997). This asymmetry alone suffices to produce the respective patterns of PWI: The more discriminable dimension in each task (words in naming, pictures in categorization) intrudes on performance with the less discriminable dimension, but not vice versa. In this view, word superiority in naming and picture superiority in categorizing depend on relative dimensional discriminability—not necessarily on differential access to lexical and semantic information by words and pictures (this is the lexical hypothesis; Glaser, 1992).

Dimensional Correlation

Dimensional correlation refers to the way in which the words and the pictures are combined to create the picture–word compounds. Following a truly random allocation of values, the conditional probability of a word, given a picture, is the same for all words. In a truly random design, the correlation established over the experimental trials between the picture and word dimensions is zero. Surprisingly, this simple stipulation for creating the stimulus ensemble (i.e., randomly combining the two components) has not been satisfied in the great bulk of PWI studies (Glaser & Dungelhoff, 1984; Lupker & Katz, 1982; Smith & Magee, 1980; Vinson & Theios, 1997) or indeed in the vast literature on the Stroop task itself (Algom et al., 1996; Dishon-Berkovits & Algom, 2000).

A correlation between picture and word built into the experimental design is fatal for selective attention to any one component. Under such conditions, the participants attend to the irrelevant dimension and open themselves up to PWI. Presented with a picture–word compound, the participant has a better-than-chance probability of identifying the word once noticing the picture. Similarly, noticing the word provides the participant with information about the picture. Humans are notorious at capturing covariation lurking in their environment (Alloy & Tabachnik, 1984; Kareev, 1995). Establishing a correlation (or correlations) over trials between picture and word, the experimenter dictates the diversion of attention toward the irrelevant dimension. PWI ensues. Portions of PWI reported in the literature might have resulted from the correlation between picture and word built into the experimental designs—not necessarily from the disparate processing of words and pictures (portrayed by the lexical hypothesis).

The Present Study

We monitored and manipulated (a) the relative discriminability of the word and the picture dimensions, and (b) the correlation over the experimental trials between the words and the pictures. For naming, the dimensions were mismatched in favor of words in Experiment 1 and the typical superiority of words emerged. When we matched relative dimensional discriminability in Experiment 2, PWI was largely eliminated. For categorization, the dimensions were mismatched in favor of picture in Experiment 3 and the typical superiority of pictures emerged. Matching dimensional discriminability in Experiment 4 resulted again in the elimination of the interference. Dimensional correlation was manipulated within each experiment (zero in the filtering task; perfect in the correlated dimensions tasks). Commensurate with our contextual approach, PWI was largely confined to the correlated dimensions condition.

Experiment 1

Method

Participants. Seventeen young men and women, undergraduates recruited from the Bar-Ilan University community, were paid to participate.

All participants had normal or corrected-to-normal vision. Their ages ranged between 20 and 25 years.

Stimuli and apparatus. The 10 stimulus items belonged in two semantic categories: fruit (*apple*, *banana*, *cherry*, *lemon*, and *avocado*) and furniture (*table*, *piano*, *vase*, *ashtray*, and *stool*). Each item appeared as both picture (a line drawing) and word. Pairing the words and the pictures into picture–word compounds created the experimental set of 100 stimuli. From these stimuli, we created eight experimental tasks (four involving word naming and four involving picture naming). For picture naming, the participants performed in a baseline task, two correlated dimensions tasks (with picture and word correlated either positively or negatively) and a filtering task. The participants also performed in four complementary tasks, naming words. A brief description of each task follows.

In the baseline task the participants named the pictures of the picture–word compounds while the word remained invariant from trial-to-trial (a different constant word for each participant). There were 10 pictures combined with the constant word; each combination presented 20 times. The 200 trials of the baseline task were presented in a random order for each participant. The complementary task of baseline word naming was similarly constructed, except that now the picture remained invariant from trial-to-trial.

In the filtering task the participant was presented with all 100 stimulus combinations of picture and word. Again, the task was to name the words and ignore the pictures; however, the irrelevant pictures also varied from trial-to-trial. Each stimulus combination was presented twice. The 200 trials of the filtering task were presented in a random order for each participant.

In the positively and negatively correlated dimensions tasks the participant was presented again with all 100 stimulus combinations. This basic set was divided into subsets of 10 and 90 stimuli. The former contained the congruent stimuli in which the word named the picture. The latter contained the incongruent stimuli in which the picture and the word conflicted. Each combination was presented twice, making for 20 trials in the positively correlated dimensions condition and 180 trials in the negatively correlated dimensions condition. Again, the task for the participants was to name the pictures and ignore the words (or, for word as the relevant dimension, to name the words and ignore the pictures).

We calculated the various measures of selective attention as follows. Garner interference was computed as the difference in speed and accuracy between the filtering task and the baseline task. *Redundancy gain* (or loss) was computed as the difference in speed and accuracy between the positively (or negatively) correlated dimensions task and the baseline task. *Stroop congruity* was computed as the difference in speed and accuracy between congruent and incongruent stimuli in each task. The three measures of Stroop congruity were derived from the correlation tasks, the filtering task, and the baseline task.

For additional Stroop analyses, we distinguished between two levels of incongruency. *Partially incongruent stimuli* were those in which the word did not name the picture but in which the picture and the word belonged in the same semantic category. *Fully incongruent stimuli* were those in which the words neither named the picture nor shared its semantic category. A difference in performance for partially incongruent stimuli and fully incongruent stimuli taps the effect of category membership. On the basis of Lupker's (1979) findings, we expected naming performance to be better for fully incongruent stimuli than for partially incongruent stimuli. Such results supply evidence that the semantic processing of the irrelevant dimension extended to category membership.

On a trial, a word embedded in a picture appeared at about the center of the computer screen. The words and the pictures appeared in black on a gray background. A professional artist prepared the 10 line drawings, modeling them after prototypical exemplars taken from children books. Ninety students in an undergraduate class recognized the pictures without making mistakes. The picture subtended 8.1° of visual angle in width and 8.1° in height. The respective visual angles for the word were 1.22° and 0.4°. The words were presented in Hebrew font, Levenim. A Power Mac computer with a 14-in. (35.56-cm) super-VGA color monitor generated the stimuli. Our specifically prepared program also controlled reaction time (RT) measurement and data collection. To avoid adaptation, we introduced a trial-to-trial spatial uncertainty of 10 pixels around the center location of the stimuli. The participant was seated at a distance of approximately 70 cm from the screen.

Procedure. The participants were tested individually in a dimly lit room. They were instructed to attend to the relevant dimension and ignore irrelevant variation. They were also encouraged to respond quickly but accurately. Participants performed the four word-naming tasks and the four picture-naming tasks together as a set, with a random 9 of the participants performing word naming first and 8 performing picture naming first. Within each set, the order of testing was random. Prior to performing a particular block, the participants performed 20 trials of that block as practice. Trials were presented randomly within each task. Naming was made orally by speaking the word or the name of the picture into a microphone. The participant's vocal response interrupted the software timer that measured latency in milliseconds. Subsequently, the experimenter recorded accuracy and next trial was initiated after 1 s. An entire experimental session consisting of 160 practice trials and 1,200 experimental trials lasted about 50 min.

Data analysis. To test for Garner or task effects, in each experiment we performed an analysis of variance (ANOVA) with dimension (word, picture) and task (baseline, filtering, positive correlation, and negative correlation) as within-subject variables. To test for Stroop effects, we performed another ANOVA with dimension, task (baseline, filtering, and correlation), and stimulus type (congruent, incongruent) as variables. Note that the factor of task is redefined in the latter ANOVA to include a single correlation task (the Stroop effect is derived as the difference between positive and negative correlation). Trials in which a participant's RT was either below or above two standard deviations of the mean for that participant in the pertinent condition were removed from the analysis. The values of correlation between speed and accuracy for pictures and words, respectively, were .38 ($p > .1$) and .28 ($p > .1$) in Experiment 1, .14 ($p > .1$) and .24 ($p > .1$) in Experiment 2, .48 ($p = .04$) and .13 ($p > .1$) in Experiment 3, and −.034 and .01 in Experiment 4.

Results

Mean RTs and proportions of accuracy for naming words and pictures appear in Table 1. Overall, observers named words faster than they named pictures, $M = 566$ and 794 ms, respectively, $F(1, 16) = 179.01$, $MSE = 1{,}770{,}391$, $p < .01$. In the baseline task, in particular, average RTs were 562 ms for words and 792 ms for pictures, a large asymmetry of 230 ms, $t(16) = 14.19$, $p < .01$. The participants also were more accurate in the former (99.6%) than in the latter (97.7%) task, $t(16) = 6.6$, $p < .01$. Garner interference and redundancy gains and losses were confined to picture naming, $F(3, 16) = 9.2$, $MSE = 25{,}594.1$, $p < .01$, for the interaction of criterial dimension and task. For picture, the difference in performance between the baseline and the filtering tasks amounted to a Garner interference of 36 ms, $t(16) = 4.2$, $p < .01$. Performance at baseline was slower by 55 ms than in the positively correlated dimensions task, $t(16) = 2.22$, $p < .05$, but faster by 28 ms than in the negatively correlated dimensions task, $t(16) = 3.4$, $p < .05$. For word, in sharp contrast, a Garner interference of 3 ms was negligible and neither task of correlated dimensions differed from baseline, $t(16) = 1.08$, $p > .1$, and $t(16) = 0.7$, $p > .1$, for positively and negatively correlated dimensions, respectively.

Table 1
Mean Reaction Times (RTs; in ms) and Proportions of Accuracy for Word Naming and Picture Naming in Baseline, Filtering, and Correlated Dimension Tasks (Experiment 1)

	Picture naming				Word naming					
	RT		Accuracy		RT		Accuracy		Overall	
Task	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	RT	%
Baseline	792	82	97.7	1.2	562	48	99.6	0.6	677	98.6
Filtering	828	88	97.7	1.6	559	48	99.7	0.4	694	98.7
Positive correlation	737	122	99.7	1.2	587	103	100.0	0.0	662	99.8
Negative correlation	820	92	97.6	1.7	556	58	99.5	0.6	688	98.5
Overall	794	86	98.1	1.6	566	52	99.7	0.5		

The results regarding Stroop congruity are presented in Table 2. Stroop effects plagued only picture naming, $F(2, 32) = 6.5$, $MSE = 13{,}980$, $p < .01$, for the interaction of criterial dimension, task, and congruity, reinforced by the lack of the two-way interaction ($F < 1$) of task and congruity. In the filtering task, observers named pictures of congruent compounds in 781 ms on average but named pictures of incongruent compounds in 833 ms on average, $t(16) = 6.49$, $p < .001$. This Stroop effect of 52 ms contrasted with a larger effect of 83 ms, $F(1, 16) = 4.73$, $MSE = 5{,}905.2$, $p < .05$, obtained in the correlation task (the difference in performance between the positively and negatively correlated conditions, $M =$ 737 and 820 ms, respectively, $t[16] = 3.6$, $p < .01$). A negligible effect of 5 ms was recorded in the baseline task. Accuracy amounted to 98.8% and 97.5% for congruent and incongruent compounds in the filtering task, defining a Stroop congruity of 1.3%, $t(16) = 3.7$, $p < .05$. In the correlated tasks, Stroop congruity was 2.1%, $t(16) = 3.9$, $p < .05$.

In Table 3, we partitioned the incongruent stimuli from the filtering task into partially incongruent (the word and the picture, though different, sharing category) and fully incongruent (word and picture belonging in separate categories) stimuli. Participants named the pictures of partially incongruent stimuli in 842 ms on average but named the pictures of fully incongruent stimuli in 825 ms on average. The difference, 17 ms, $t(16) = 2.3$, $p < .05$, shows that the participants processed the irrelevant words to the level of category membership. For word naming, by contrast, category membership of the irrelevant pictures did not affect performance, $t(16) = 4.2$, $p > .05$.

Discussion

Deploying the full Garnerian regimen, we replicated the common finding of word superiority in naming picture–word stimuli. Naming pictures, our participants could ignore neither trial-to-trial variation in irrelevant word (thus suffering Garner interference) nor the meaning of the words (thus suffering Stroop interference). Apart from noticing the meaning of the individual words, the participants also extracted their semantic category (conflicting words that shared the category of the target pictures intruded more on naming of the pictures than did those belonging in a different category; cf. Lupker, 1979; Lupker & Katz, 1981, 1982). The naming of words, by contrast, was free of Stroop and Garner interference.

Monitoring context, we recorded a glaring mismatch in discriminability favoring the word dimension. The asymmetry rendered the word dimension hard to ignore but rendered the picture dimension relatively easy to ignore. For dimensional correlation, PWI was smaller under the zero correlation of the filtering task than

Table 2
Stroop Congruity: Mean Reaction Times (RTs; in ms) and Proportions of Accuracy for Congruent and Incongruent Stimuli in Naming of Pictures and Words at Baseline, Filtering, and Correlated Tasks (Experiment 1)

	Congruent				Incongruent				Congruity score	
	RT		Accuracy		RT		Accuracy			
Task	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	RT	%
					Picture naming					
Baseline	787	125	97.6	4.3	792	79	97.7	1.3	5	−0.1
Filtering	781	97	98.8	2.1	833	87	97.5	1.5	52	1.3
Correlated tasks	737	122	99.7	1.2	820	92	97.6	1.7	83	2.1
					Word naming					
Baseline	557	59	100.0	0	562	48	99.6	0.6	5	0.4
Filtering	566	58	100.0	0	558	48	99.6	0.4	−8	0.4
Correlated tasks	587	103	100.0	0	556	58	99.5	0.6	−31	0.5

Table 3
Mean Reaction Times (RTs; in ms) for Picture–Word Compounds in the Filtering Task as a Function of Semantic Relation (Experiment 1)

Task	Partially incongruent		Fully incongruent	
	M	*SD*	*M*	*SD*
Picture naming	842	88	825	89
Word naming	556	48	560	48

under the full contingency of the correlation task. In the latter task, the participants noticed the predictive relation established between word and picture, compromising selectivity to picture to a great extent.

In Experiment 2, we attempted to redress the dimensional imbalance of Experiment 1 favoring words. A concerted effort preceded the experiment in an attempt to eliminate the mismatch in discriminability. We asked, will word superiority vanish for (roughly) matched dimensions?

Experiment 2

Method

Participants. Sixteen young men and women, undergraduates recruited from the Bar-Ilan University community, were paid to participate. All participants had normal or corrected-to-normal vision. Their ages ranged between 20 and 29. None had participated in Experiment 1.

Stimuli and apparatus. To achieve the desired match between dimensions, we attempted to reduce the discriminability of words but enhance the discriminability of pictures. For words, the manipulation was twofold. We reduced the contrast between the print and the background. We also added an *x* as a prefix and as a suffix to each word (for example, the word *chair* appeared as *xchairx*). For pictures, we reduced size to achieve the optimal visual angle for best performance (Theios & Amrhein, 1989). The pictures thus subtended 5.5° of visual angle in width and 6.5° in height. The apparatus, design, and conditions of stimulus presentation were those used in Experiment 1.

Procedure. The procedure was that of Experiment 1, with half of the participants performing word naming first and half performing picture naming first.

Results

Mean RTs and proportions of accuracy in the various tasks appear in Table 4. Overall, responding was faster by 37 ms to words than to pictures, $M = 659$ and 696 ms, respectively, $F(1, 15) = 6.57$, $MSE = 43{,}618.7$, $p < .05$. At baseline, the advantage of words was 55 ms, $t(15) = 3.6$, $p < .01$, although accuracy was comparable for words (99.1%) and pictures (98.5%), $t(15) = 1.6$, $p > .1$. Garner interference was present for neither picture (0 ms) nor word (9 ms), $t(15) = 2.1$, $p > .05$; its absence was reinforced by the lack of effects for task, $F(2, 30) = 2.68$, $MSE = 792.31$, $p > .05$, and for the interaction of task and dimension ($F < 1$). The results of the positive correlation condition, inconsequential for the derivation of Garner interference, were excluded from these analyses. Garner interference was absent for accuracy too, amounting to 98.5% (baseline) and 98.1% (filtering) for picture and to 99.1% (both tasks) for word. The single intertask effect in the data of Table 4 was the redundancy gain of 56 ms observed for picture, $M = 709$ and $M = 653$ ms, respectively, in the baseline and the positive correlation tasks, $F(3, 45) = 5.2$, $MSE = 4{,}086.7$, $p < .05$.

Mean RTs and accuracy rates for Stroop congruity are presented in Table 5. Larger values of Stroop congruity were present for picture than for word, $F(2, 30) = 4.03$, $MSE = 5{,}161.5$, $p < .05$, yet this contrast derived from a single large value obtained in the correlation task, congruity score of 60 ms, $t(15) = 4.4$, $p < .01$. Apart from this effect, none of the other conditions entailed a significant effect of Stroop. Performing an ANOVA without the correlation task confirmed the absence of the Stroop effect from the baseline and the filtering tasks for both dimensions ($F < 1$, for the three-way interaction of criterial dimension, task, and congruity). For picture, the Stroop effect in the correlation condition (the only significant effect in the entire data of Table 5) was 5 times the value of that in the filtering condition, congruity scores of 60 and 13 ms, respectively, $t(15) = 3.18$, $p < .01$. For word, the respective values were 21 and 9 ms.

In Table 6 we list the mean RTs for incongruent stimuli whose components shared (partially incongruent stimuli) and did not share (fully incongruent stimuli) semantic category. There was no effect of semantic category, $F(1, 15) = 3.81$, $MSE = 993.4$, $p > .05$, for either word or picture ($F < 1$).

Table 4
Mean Reaction Times (RTs; in ms) and Proportions of Accuracy for Word Naming and Picture Naming in Baseline, Filtering, and Correlated Dimension Tasks (Experiment 2)

Task	Picture naming				Word naming				Overall	
	RT		Accuracy		RT		Accuracy			
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	RT	%
Baseline	709	58	98.5	1.9	654	63	99.1	1.0	682	98.8
Filtering	709	63	98.1	1.4	663	66	99.1	1.0	686	98.5
Positive correlation	653	56	98.7	2.8	649	58	99.0	2.0	651	98.9
Negative correlation	713	59	97.7	1.2	670	73	98.9	1.1	692	98.4
Overall	696		98.3		659		99.0			

Table 5
Stroop Congruity: Mean Reaction Times (RTs; in ms) and Proportions of Accuracy for Congruent and Incongruent Stimuli in Naming of Pictures and Words at Baseline, Filtering, and Correlated Tasks (Experiment 2)

	Congruent				Incongruent				Congruity score	
	RT		Accuracy		RT		Accuracy			
Task	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	RT	%
Picture naming										
Baseline	727	86	98.7	2.2	707	57	98.5	2.1	−20	−0.2
Filtering	695	70	99.0	2.0	708	56	97.9	1.6	13	1.1
Correlated tasks	653	56	98.7	2.8	713	59	97.7	1.2	60	1.0
Word naming										
Baseline	643	64	99.3	1.7	656	64	99.0	1.0	13	0.3
Filtering	655	70	99.0	2.0	664	66	99.0	0.9	9	0.0
Correlated tasks	649	58	99.0	2.0	670	73	98.9	1.1	21	0.1

Discussion

Reducing dimensional imbalance (from 228 ms in Experiment 1 to 37 ms in Experiment 2—a six-fold drop) largely eliminated PWI. First, our participants could ignore trial-to-trial variation in picture when naming the words and they could ignore trial-to-trial variation in word when naming the pictures (hence they suffered no Garner interference). Second, the meaning of the irrelevant component intruded only minimally on performance with the relevant component whether for word or picture (hence, the absence of Stroop interference). Third, the semantic category of the irrelevant component did not affect performance with either word or picture. Therefore, the contextual factor of relative dimensional discriminability exerted a profound influence on the size of PWI. When the constituent dimensions were equally salient in Experiment 2, neither intruded on the other and performance was largely free of PWI.

Dimensional correlation, another variable of context, also exerted an appreciable influence on performance. A statistically significant effect of Stroop obtained only in the correlation tasks. Numerically too, the effects were larger in the correlation task than in the filtering task. The predictive context established in the correlation task dictated the allocation of attention to the nominally irrelevant component, the cost to performance expressed as a large Stroop effect.

In Experiments 1–2 we used naming for response. In Experiments 3–4 we used categorization rather than naming to further examine the effects of context on PWI. Because we introduced no special precautions in Experiment 3, we expected the usual advantage of pictures over words in categorization to emerge.

Table 6
Mean Reaction Times (RTs; in ms) for Picture–Word Compounds in the Filtering Task as a Function of Semantic Relation (Experiment 2)

	Partially incongruent		Fully incongruent	
Task	*M*	*SD*	*M*	*SD*
Picture naming	706	62	716	68
Word naming	668	69	674	69

Experiment 3

Method

Participants. Eighteen young men and women, undergraduates recruited from the Bar-Ilan University community, were paid to participate. All participants had normal or corrected-to-normal vision. Their ages ranged between 21 and 30 years. None had participated in the previous experiments.

Stimuli and apparatus. The same 100 stimuli from Experiment 1 were used. Because the task was one of categorization, there was but a pair of responses per dimension (the semantic categories of fruit and furniture). The change in task (from naming to categorizing) also entailed modifications in the four experimental conditions.

The defining feature of the Garnerian baseline task is that the irrelevant dimension is held constant. In Experiment 3, the category of the irrelevant dimension (word or picture) was held constant at fruit or furniture but we allowed the individual items within that category to vary randomly from trial to trial. Each task of baseline (two tasks per dimension) contained 50 stimuli that were presented twice.

The filtering task was similar to those of Experiments 1 and 2, except that in Experiment 3 participants categorized the components of the picture–word compounds rather than naming them.

In the positively correlated dimensions condition the participant was presented with 50 congruent stimuli in which the picture and the word shared category. In the negatively correlated dimensions condition the participant was presented with the remaining 50 incongruent stimuli in which the categories of the picture and the word conflicted.

Garner and Stroop effects were calculated following the procedures of Experiment 1–2. In addition, categorization allows for the partition of the congruent stimuli into two subsets. *Partially congruent* stimuli were those in which the word and the picture were drawn from the same category but in which the two did not refer to the same object. *Fully congruent* stimuli were those in which the word named the picture.

Apart from these changes, the apparatus, design, and conditions of stimulus presentation were those used in Experiment 1.

Procedure. The procedure was that of Experiment 1 with two notable exceptions. First, participants were asked to categorize rather than name the components of the picture–word compounds presented. Second, the

Table 7
Mean Reaction Times (RTs; in ms) and Proportions of Accuracy for Word Categorization and Picture Categorization in Baseline, Filtering, and Correlated Dimension Tasks (Experiment 3)

	Picture categorization				Word categorization					
	RT		Accuracy		RT		Accuracy		Overall	
Task	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	RT	%
Baseline	564	42	96.7	0.4	651	47	95.6	0.6	607	96.1
Filtering	561	43	96.4	0.5	681	50	94.8	0.9	621	95.6
Positive correlation	570	44	97.2	0.6	598	42	97.7	0.3	584	97.4
Negative correlation	587	48	97.0	0.4	703	54	94.9	0.8	645	95.9
Overall	570		96.8		658		95.7			

participants responded manually rather than orally by pressing the appropriate key on the computer keyboard. The two categories were assigned to the letters *D* and *K* on the keyboard. Key assignment was counterbalanced across participants. Participants were instructed to press the key that corresponded to the category of the relevant component as fast and as accurately as possible. The response interrupted the software timer that measured latency in milliseconds; the same software also measured accuracy. The next stimuli followed in 1 s. An entire experimental session, consisting of 160 practice trials and 1,200 experimental trials, lasted about 50 min.

Results

Mean RTs and proportions of accuracy for categorization of words and pictures appear in Table 7. Overall, the participants responded faster to pictures than to words, $M = 570$ and 658 ms, respectively, $F(1, 17) = 56.4$, $MSE = 278{,}520$, $p < .01$. At baseline, average RTs were 564 ms for pictures and 651 ms for words, a difference of 87 ms, $t(17) = 64.5$, $p < .01$. Participants were also more accurate categorizing pictures (96.7%) than words (95.6%), $t(17) = 5.8$, $p < .05$. Garner interference and virtually all effects of redundancy were confined to words, $F(3, 17) = 15.3$, $MSE = 16{,}020.3$, $p < .01$, for RT, and $F(3, 17) = 6.2$, $MSE = 11.9$, $p < .01$, for accuracy. For RT, the difference in performance between the baseline ($M = 651$ ms) and the filtering ($M = 681$ ms) tasks defined a Garner interference of 30 ms, $t(17) = 7.5$, $p < .01$. For accuracy, the respective values of 95.6% and 94.8% produced a Garner interference of 0.8%, $t(17) = 3.05$, $p = .08$. Categorization at baseline was slower by 53 ms than in the positive correlation task, $t(17) = 23.7$, $p < .01$, but faster by 52 ms than in the negative correlation task, $t(17) = 23.6$, $p < .01$. For accuracy, we recorded a redundancy gain of 2.1%, $t(17) = 20.9$, $p < .01$. For picture, in sharp contrast, a Garner interference of 3 ms was negligible and the single intertask effect recorded was a redundancy loss of 23 ms, $t(17) = 4.2$, $p = .04$.

The results concerning Stroop congruity are presented in Table 8. Stroop effects affected only categorization with the less discriminable dimension of word, $F(2, 34) = 10.69$, $MSE = 9{,}463.3$, $p < .01$. The effect was 33 ms in the filtering task, $t(17) = 3.5$, $p < .01$, but amounted to an appreciable 105 ms, $t(17) = 94.6$, $p < .01$, in the correlation task—a threefold increase, $F(1, 17) = 21.66$, $MSE = 24{,}282.4$, $p < .01$. At baseline, a difference of 11 ms between congruent and incongruent compounds was not reliable, $t(17) = 1.5$, $p > .1$, as were all the results with picture (congruity scores of 11, 5, and 17 ms, respectively, at the baseline, filtering, and correlation tasks). In the filtering task, the participants were more accurate to categorize words of congruent compounds than words of incongruent compounds by 4.9%, $t(17) = 3.4$, $p < .01$. The Stroop effect was 2.8% in the correlation task, $t(17) = 37.7$, $p < .01$, and it was reliable at baseline,

Table 8
Stroop Congruity: Mean Reaction Times (RTs; in ms) and Proportions of Accuracy to Congruent and Incongruent Stimuli in Categorization of Pictures and Words at Baseline, Filtering, and Correlated Tasks (Experiment 3)

	Congruent				Incongruent				Congruity score	
	RT		Accuracy		RT		Accuracy			
Task	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	RT	%
Picture categorization										
Baseline	559	47	97.1	2.3	570	47	96.3	2.0	11	0.8
Filtering	560	37	96.3	2.9	565	44	96.5	2.8	5	−0.2
Correlated tasks	570	44	97.2	0.6	587	48	97.0	0.4	17	0.2
Word categorization										
Baseline	646	52	96.5	2.6	657	50	94.8	3.5	11	1.7
Filtering	664	45	97.3	2.3	697	50	92.4	6.5	33	4.9
Correlated tasks	598	42	97.7	0.3	703	54	94.9	0.8	105	2.8

too, 1.7%, $t(17) = 2.7$, $p < .05$. None of the minuscule effects of Stroop obtained with pictures was reliable.

In Table 9, we partitioned the congruent stimuli into fully congruent stimuli (with the picture and the word components sharing name and category) and partially congruent stimuli (with the picture and word components sharing only category). Neither the main effect of component identity, $F(1, 17) = 3.58$, $MSE = 4{,}122.5$, $p > .05$, nor the interaction of identity and dimension ($F < 1$) was reliable, showing that the identity of the individual components did not affect categorization for either picture or word.

Discussion

We reproduced the often-reported superiority of pictures in categorization. Categorizing words, the participants could ignore neither trial-to-trial variation in category of the irrelevant pictures (thus suffering Garner interference) nor the semantic category of those pictures (thus suffering Stroop interference). The semantic category of the irrelevant pictures was noticed over and above their individual identity. Note that it was not mere variation of the individual pictures that hindered the categorization of words. Irrelevant pictures varied from trial-to-trial in both the baseline (within a category) and the filtering tasks (across categories). It was the variation in category that incurred cost to performance (Garner interference). In a similar vein, stimuli in which the word named the picture were not categorized faster than other category-congruent stimuli with different individual components. What counted for performance was semantic category, not individual exemplars.

The results of Experiment 3 form a mirror image of those obtained in Experiment 1. Naming picture–word compounds in Experiment 1 resulted in word superiority; categorizing the same stimuli resulted in picture superiority. We implicate the contextual factors of relative dimensional discriminability and dimensional correlation. For the former, we recorded a large mismatch in discriminability favoring picture. The advantage rendered the pictures hard to ignore, but rendered the words easy to ignore. For the latter, PWI was larger in the correlation task than in the filtering task. Target selectivity was compromised to a larger extent under the predictive relationship established in the latter condition. Commensurate with the current contextual approach, we asked, would PWI in categorization evaporate under matched discriminability?

Table 9
Mean Reaction Times (RTs; in ms) for Picture–Word Compounds in the Filtering Task as a Function of Semantic Relation (Experiment 3)

Task	Partially congruent		Fully congruent	
	M	*SD*	*M*	*SD*
Picture naming	562	86	552	80
Word naming	669	79	649	91

Experiment 4

Method

Participants. Seventeen young men and women, undergraduates recruited from the Bar-Ilan University community, were paid to participate. All participants had normal or corrected-to-normal vision. Their ages ranged between 22 and 31 years. None had participated in the previous experiments.

Stimuli and apparatus. To attain the desired match between dimensions, we replaced one of the categories from Experiments 1–3 with a new category such that the variation among the drawings belonging to the two categories was small and difficult to detect. The change rendered the categorization of pictures more difficult while leaving the categorization of words unaffected. In Experiment 3, classifying pictures was easy because members of the furniture category usually come by angular features whereas members of the fruit category usually come by circular features. In Experiment 4, we replaced furniture with vegetables that, like fruits, also are depicted by circular features (see, Snodgrass & McCullough, 1986). The five items of the vegetable category selected were *carrot*, *eggplant*, *mushroom*, *cucumber*, and *pumpkin*. Again, a professional artist prepared the five new line drawings, modeling them after prototypical exemplars taken from children's books. In a pilot study, 90 students correctly categorized each picture as that of the intended vegetable. The apparatus, design, and conditions of stimulus presentation were those used in Experiment 3.

Procedure. The procedure was that of Experiment 3. A random 9 of the participants performed word categorization first and 8 performed picture categorization first. Participants were instructed to categorize the word or the picture as fast and accurately as possible.

Results

Mean RTs and proportions of accuracy for categorizing words and pictures appear in Table 10. Overall, categorization did not differ for picture and word ($M = 689$ and 694 ms, respectively, $F < 1$). At baseline too, the small difference of 9 ms was not reliable, $t(17) = 0.4$, $p > .1$, as was the minuscule difference of 0.3% in accuracy, $t(16) = 0.25$, $p > .1$. Garner interference (as well as all effects of redundancy) was absent from the data for both dimensions ($F < 1$, for task and for the interaction of task and criterial dimension). Accuracy mimicked RT in exhibiting symmetry across picture and word and freedom of Garner interference, $F(1, 17) = 2.4$, $MSE = 26.9$, $p > .1$, $F(1, 17) = 1.49$, $MSE = 4.8$, $p > .1$, and $F < 1$, respectively, for dimension, task, and their interaction.

The results regarding Stroop congruity are presented in Table 11. Categorization of neither pictures nor words was plagued by Stroop interference, $F(1, 16) = 1.32$, $MSE = 2{,}670.45$, $p > .1$, for congruity, and $F < 1$ for all the other effects. The results of the ANOVA notwithstanding, separate tests performed on the two largest scores (Stroop values of 16 ms, $t[16] = 1.8$, $p > .05$, and 13 ms, $t[16] = 2.04$, $p > .05$, respectively, for filtering with pictures and words) further underscored the lack of Stroop congruity in the data. For accuracy, an ANOVA did not yield significant effects for dimension, task, congruity, or any of the interactions ($F < 1$, for all of the effects).

Mean RTs for fully congruent stimuli and for partially congruent stimuli in the filtering task are presented in Table 12. We did not find effects of the identity of the individual component for either picture, $t(16) = 0.9$, $p > .1$, or word, $t(16) = 0.9$, $p > .1$. Components of stimuli in which the word named the picture (fully

Table 10
Mean Reaction Times (RTs; in ms) and Proportions of Accuracy for Word Categorization and Picture Categorization in Baseline, Filtering, and Correlated Dimension Tasks (Experiment 4)

	Picture categorization				Word categorization					
	RT		Accuracy		RT		Accuracy		Overall	
Task	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	RT	%
Baseline	681	86	95.0	5.2	690	42	94.7	5.2	685	94.8
Filtering	683	96	95.2	3.5	681	41	94.0	3.6	682	94.6
Positive correlation	695	81	95.7	4.4	703	64	94.5	3.7	699	95.1
Negative correlation	697	78	94.7	4.7	703	49	93.8	4.5	700	94.2
Overall	689		95.1		694		94.2			

congruent stimuli) were not categorized faster than components of stimuli in which the picture and word merely came from the same category (partially congruent stimuli).

Discussion

Matching the discriminability of the picture and word dimensions, pictures no longer held an advantage over words in categorization. The participants could ignore trial-to-trial variation in the category of the irrelevant pictures (or words) when categorizing words (or pictures). Hence, classification performance on either dimension was free of Garner interference. The absence of PWI extended to the Stroop effects: Performance was on par for category-congruent and category-incongruent stimuli. Hence, categorization of words and pictures alike was free from Stroop interference. The complete elimination of PWI is the signature of our study's results.

The results of a study by Snodgrass and McCullough (1986) underscore the need to match the discriminability of the picture and word components in studies of classification. These authors showed that pictures drawn from visually dissimilar categories were classified faster than were the pictures' names, documenting picture superiority in categorization. However, the same authors also showed that pictures drawn from visually similar categories were classified slower than were their names, thereby showing word superiority in categorization. In the Snodgrass and McCullough study, the pictures and the words were presented separately; hence, they did not compete for the participant's attention. In our study, by contrast, the participants had to abstract the relevant pictures or words from the picture–word compounds. PWI (or lack thereof) critically depended on relative dimensional discriminability.

The results of Experiment 4 show that PWI is a contingent feature—not a mandatory feature—of the processing of picture–word stimuli. The results of two auxiliary experiments further reinforce this conclusion. We tested a new group of 12 participants in the filtering task of Experiment 4 and again found comparable performance for pictures and words. A striking feature of the data was the absence of Stroop effects from performance with both criterial dimensions. We attribute the elimination of Stroop interference to the conditions of matched discriminability and zero correlation that prevailed throughout that experiment. In a second experiment, another group of 17 participants categorized orally (rather than manually) the stimuli of Experiment 4. An important finding of the research within the Garnerian tradition (Flowers & Stoup, 1977; Flowers, Warner, & Polansky, 1979; Melara & Mounts, 1993; Sabri, Melara, & Algom, 2001) is the dependency of dimensional discriminability on the response mode used. Equally discriminable stimuli for a given mode of responding

Table 11
Stroop Congruity: Mean Reaction Times (RTs; in ms) and Proportions of Accuracy to Congruent and Incongruent Stimuli in Categorization of Pictures and Words at Baseline, Filtering, and Correlated Tasks (Experiment 4)

	Congruent				Incongruent				Congruity score	
	RT		Accuracy		RT		Accuracy			
Task	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	RT	%
					Picture categorization					
Baseline	680	87	95.4	4.9	681	91	94.6	5.8	1	0.8
Filtering	675	102	95.8	4.6	691	92	94.5	3.5	16	1.3
Correlated tasks	695	81	95.7	4.4	697	78	94.7	4.7	2	1.0
					Word categorization					
Baseline	685	38	95.2	3.3	695	56	93.9	2.8	10	1.3
Filtering	675	34	95.1	4.2	688	51	93.0	3.7	13	2.1
Correlated tasks	703	64	94.5	3.7	703	49	93.8	4.5	0	0.7

Table 12
Mean Reaction Times (RTs; in ms) for Picture–Word Compounds in the Filtering Task as a Function of Semantic Relation (Experiment 4)

	Partially congruent		Fully congruent	
Task	*M*	*SD*	*M*	*SD*
Picture naming	678	111	665	78
Word naming	678	59	663	39

(manual) may or may not be equally discriminable when another mode of responding (oral) is used. We found that oral responding slowed down RTs by 63 ms on average. However, the slowdown affected pictures and words similarly, and the dimensions remained matched for discriminability. Stroop and Garner interference was eliminated for oral responding too. We conclude that matched discriminability is conducive to the collapse of PWI in processing of picture–word stimuli.

General Discussion

The results of our study demonstrate that mutual interference in processing picture–word compounds is malleable to the point of elimination. PWI has been thought to be a robust feature of such processing and it forms the cornerstone of dual-channel theories such as those developed by Glaser (1992), Seymour (1973), and Theios and Amrhein (1989). According to our approach, by contrast, PWI is not entirely fixed by differential access to lexical and semantic information by words and pictures. The perceiver is sensitive to context, including covariation of the components (dimensional correlation) and their relative saliency (dimensional discriminability). Often, performance and interference depend on these factors of context and not on whether the target stimuli are pictures or words.

Our results pose a challenge to traditional explanations of the Stroop effect. The Stroop effect is often considered to be the outcome of automatic processing of the meaning of words. The absence of Stroop interference in several conditions of our study implies that people can ignore the words while processing the pictures. The contextual factors of dimensional discriminability and correlation moderated the activation of meaning considerably. Besner, Stolz, and their colleagues have recently concluded that “Contrary to the view expressed repeatedly over the past 60 years in the Stroop literature, at least some aspects of word recognition are not automatic” (Besner, Stolz, & Boutilier, 1997, p. 224) and that “the Stroop effect is reduced in magnitude or eliminated, depending on details of context” (Besner & Stolz, 1999, p. 99; see also Bauer & Besner, 1997; Besner, 2001; Stolz & Besner, 1999). Our findings and conclusions are consistent with those of Besner, Stolz, and their colleagues in challenging the alleged robustness of the Stroop effect.

Concerning the present factors of context, discriminability is not the trivial manipulation by which one affects the perception of the words and the pictures. Manipulations of discriminability leave the perception, identification, recognition, and psychophysical properties of the stimuli intact. Performance was virtually errorless even with the less discriminable dimension. Discriminability affects attention, not sensation, by making one or the other dimension more salient. In the vast literature on the Stroop task, the color words used were invariably more discriminable than the colors used—a mismatch that easily produced the classic Stroop asymmetry by which words intrude on colors but not vice versa. However, when the colors were made more discriminable than the words, the reverse of the classic Stroop effect appeared, and strikingly; when the dimensions were matched in discriminability the Stroop effect vanished (Melara & Mounts, 1993; see also, Algom et al., 1996; Pansky & Algom, 1999; Shalev & Algom, 2000).

Concerning correlation, the insidious creeping into the experimental design of a hidden correlation between word and picture (or between word and color) started with Stroop (1935) himself. He only presented incongruent stimuli (none of the words appeared in its matching print color in Stroop’s classic study), thus establishing a negative correlation between word and color. Modern studies of the Stroop effect have typically applied a stimulus arrangement in which congruent and incongruent stimuli are presented in equal number. The implicit assumption held by many investigators is that by equating the frequency of congruent and incongruent stimuli the word is not predictive of the picture and vice versa. However, the assumption is incorrect. Because there are fewer congruent combinations, each congruent stimulus is typically presented more often than each incongruent stimulus. A positive correlation between dimensions ensues. Note that any deviation from a truly random allocation of the dimensional values is bound to establish an informative relationship between word and picture over the experimental trials. This information in turn undermines selective attention to either the picture or the word (Dishon-Berkovits & Algom, 2000).

In our study we recorded the profound effects of these factors of context on PWI. In Figure 2 we summarize the results obtained with the two tasks that carried the most theoretical interest in the study: naming pictures and categorizing words. In the past these tasks have uniformly yielded to the detrimental effects of Stroop and Garner interference. In Figure 2 we illustrate how the interference can be brought under experimental control by the judicious manipulation of context.

Consider first PWI in picture naming, depicted in Panel A of Figure 2. When the dimensions were grossly mismatched in favor of word (Experiment 1), considerable effects of Stroop and Garner plagued performance with the pictures. When the dimensions were more balanced (Experiment 2), Stroop and Garner interference almost vanished (rendering the naming of pictures almost as easy as that of the words). The Stroop effects obtained were also larger under the correlation contexts than under orthogonal contexts. The naming of pictures was governed by the contextual factors of discriminability, $F(1, 31) = 10.6$, $MSE = 26{,}967.9$, $p < .01$, and correlation, $F(2, 62) = 9.34$, $MSE = 25{,}407.0$, $p < .01$. The outcome of a planned test, $t(32) = 2.7$, $p < .01$, reinforced further the conclusion that Stroop interference was larger under correlation contexts than under random contexts.

In Panel B of Figure 2, we show the parallel results for categorization of words. As relative dimensional discriminability varied from one favoring picture (Experiment 3) to equality (Experiment

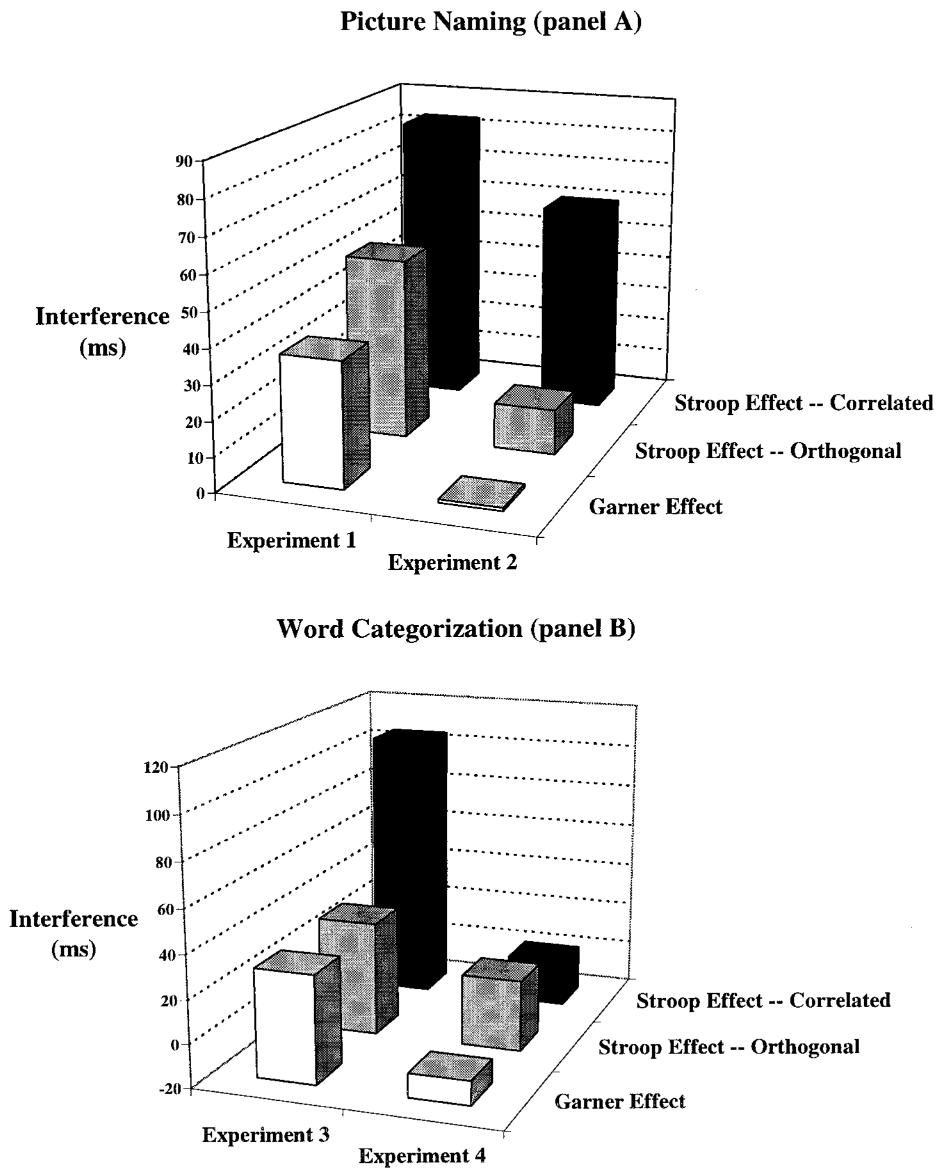

Figure 2. Stroop and Garner effects as a function of dimensional discriminability and correlation. Panel A depicts Stroop and Garner effects for naming of pictures under mismatched discriminability (Experiment 1) and under more balanced discriminability (Experiment 2). Panel B depicts Stroop and Garner effects for categorization of words under mismatched discriminability (Experiment 3) and under balanced discriminability (Experiment 4). Dimensional correlation was manipulated within each experiment.

4), so did Garner and Stroop interference diminish (along with picture superiority). Again the factors of discriminability, $F(1, 33) = 32.2$, $MSE = 79{,}185.6$, $p < .01$, and correlation, $F(2, 66) = 6.4$, $MSE = 17{,}052.7$, $p < .01$, affected performance (recall that in Experiment 4 none of the interference effect was significant statistically). A planned test confirmed the difference in Stroop interference between the correlation and the filtering tasks, $t(34) = 2.07$, $p < .05$. We conclude that designs entailing equally salient words and pictures in a correlation-free arrangement are the paradigms of choice for testing processing with picture–word stimuli.

References

Algom, D., Dekel, A., & Pansky, A. (1996). The perception of number from the separability of the stimulus: The Stroop effect revisited. *Memory & Cognition, 24,* 557–572.

Alloy, L. G., & Tabachnik, N. (1984). Assessment of covariation by humans and animals: The joint influence of prior expectations and current situational information. *Psychological Review, 91,* 112–149.

Arieh, Y., & Algom, D. (1997). Reading words in Stroop-like tasks: The effect of contingency between semantic and visual components. *Psychologia: Israel Journal of Psychology, 6,* 7–19.

Bauer, B., & Besner, D. (1997). Mental set as a determinant of processing in the Stroop task. *Canadian Journal of Experimental Psychology, 51,* 61–68.

Besner, D. (2001). The myth of ballistic processing: Evidence from the Stroop paradigm. *Psychonomic Bulletin & Review, 8,* 324–330.

Besner, D., & Stolz, J. A. (1999). What kind of attention modulates the Stroop effect? *Pyschonomic Bulletin & Review, 6,* 99–104.

Besner, D., Stolz, J. A., & Boutilier, C. (1997). The Stroop effect and the myth of automaticity *Psychonomic Bulletin & Review, 4,* 221–225.

Dishon-Berkovits, M., & Algom, D. (2000). The Stroop effect: It is not the robust phenomenon that you have thought it to be. *Memory & Cognition, 28,* 1437–1449.

Flowers, J. H., & Stoup, C. M. (1977). Selective attention between words, shapes, and colors in speeded classification and vocalization tasks. *Memory & Cognition, 5,* 299–307.

Flowers, J. H., Warner, J. L., & Polansky, M. L. (1979). Response and encoding factors in ignoring irrelevant information. *Memory & Cognition, 7,* 86–94.

Fraisse, P. (1969). Why is naming longer than reading? *Acta Psychologica, 30,* 96–103.

Garner, W. R. (1962). *Uncertainty and structure as psychological concepts.* New York: Wiley.

Garner, W. R. (1974). *The processing of information and structure.* Potomac, MD: Erlbaum.

Garner, W. R., & Felfoldy, G. L. (1970). Integrality of stimulus dimensions in various types of information processing. *Cognitive Psychology, 1,* 225–241.

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 word and picture processing. *Journal of Experimental Psychology: General, 118,* 13–42.

Kareev, Y. (1995). Positive bias in the perception of covariation. *Psychological Review, 102,* 490–502.

Lupker, S. J. (1979). The semantic nature of response competition in the picture–word interference task. *Memory & Cognition, 7,* 485–495.

Lupker, S. J., & Katz, A. N. (1981). Input, decision, and response factors in picture–word interference. *Journal of Experimental Psychology: Human Learning and Memory, 7,* 269–282.

Lupker, S. J., & Katz, A. N. (1982). Can automatic picture processing influence word judgments? *Journal of Experimental Psychology: Human Learning and Memory, 8,* 418–434.

Melara, R. D., & Mounts, J. R. W. (1993). Selective attention to Stroop dimensions: Effects of baseline discriminability, response mode, and practice. *Memory & Cognition, 21,* 627–645.

Pansky, A., & Algom, D. (1999). Stroop and Garner effects in comparative judgment of numerals: The role of attention. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 39–58.

Pomerantz, J. R. (1983). Global and local precedence: Selective attention in form and motion perception. *Journal of Experimental Psychology: General, 112,* 515–540.

Pomerantz, J. R., Pristach, E. A., & Carson, C. E. (1989). Attention and object perception. In B. E. Shepp & S. Ballesteros (Eds.), *Object perception: Structure and process* (pp. 53–89). Hillsdale, NJ: Erlbaum.

Rosinsky, R. R., Golinkoff, R. M., & Kukish, K. S. (1975). Automatic semantic processing in a picture–word interference task. *Child Development, 46,* 247–253.

Sabri, M., Melara, R. D., & Algom, D. (2001). A confluence of contexts: Asymmetric versus global failures of selective attention to Stroop dimensions. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 515–537.

Seymour, P. H. K. (1973). A model for reading, naming, and comparison. *British Journal of Psychology, 64,* 35–49.

Shalev, L., & Algom D. (2000). Stroop and Garner effects in and out of Posner's beam: Reconciling two conceptions of selective attention. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 997–1017.

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

Snodgrass, J. G., & McCullough, B. (1986). The role of visual similarity in picture categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 147–157.

Stolz, J. A., & Besner, D. (1999). On the myth of automatic semantic activation in reading. *Current Directions in Psychological Science, 8,* 61–65.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643–662.

Theios, J., & Amrhein, P. C. (1989). Theoretical analysis of the cognitive processing of lexical and pictorial stimuli: Reading, naming, and visual and conceptual comparisons. *Psychological Review, 96,* 5–24.

Vinson, D. P., & Theios, J. (November, 1997). *Asymmetric interference between incongruent pictures and words in semantic categorization.* Poster session presented at the 38th Annual Meeting of the Psychonomic Society, Philadelphia.

Received September 8, 2000
Revision received June 27, 2001
Accepted June 27, 2001 ■