How does language change perception: a cautionary note

Nola Klemfuss, William Prinzmetal and Richard B. Ivry*

Department of Psychology, University of California Berkeley, Berkeley, CA, USA

The relationship of language, perception, and action has been the focus of recent studies exploring the representation of conceptual knowledge. A substantial literature has emerged, providing ample demonstrations of the intimate relationship between language and perception. The appropriate characterization of these interactions remains an important challenge. Recent evidence involving visual search tasks has led to the hypothesis that top-down input from linguistic representations may sharpen visual feature detectors, suggesting a direct influence of language on early visual perception. We present two experiments to explore this hypothesis. Experiment 1 demonstrates that the benefits of linguistic priming in visual search may arise from a reduction in the demands on working memory. Experiment 2 presents a situation in which visual search performance is disrupted by the automatic activation of irrelevant linguistic representations, a result consistent with the idea that linguistic and sensory representations interact at a late, response-selection stage of processing. These results raise a cautionary note: While language can influence performance on a visual search, the influence need not arise from a change in perception *per se*.

Keywords: language, perception, embodied cognition, working memory, visual search

INTRODUCTION

Language provides a medium for describing the contents of our conscious experience. We use it to share our perceptual experiences, thoughts, and intentions with other individuals. The idea that language guides our cognition was clearly articulated by Whorf (1956) who proposed that an individual's conceptual knowledge was shaped by his or her language. There is clear evidence demonstrating that language directs thought (Ervin-Tripp, 1967), influences concepts of time and space (e.g., Boroditsky, 2001), and affects memory (e.g., Loftus and Palmer, 1974).
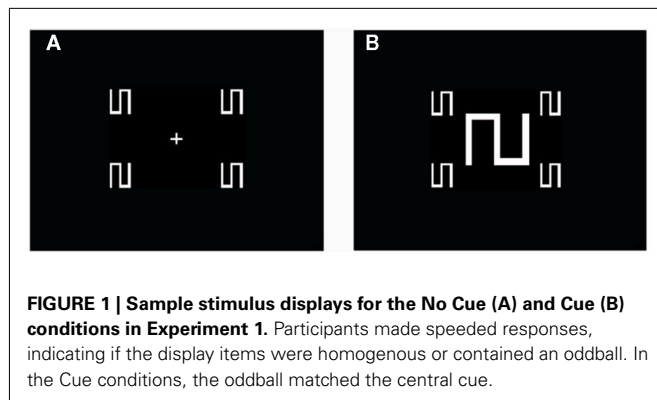
More controversial has been the claim that language has a direct effect on perceptual experience. In a seminal study, Kay and Kempton (1984) found that linguistic labels influence decisions in a color categorization task. In the same spirit, a flurry of studies over the past decade has provided ample demonstrations of how perceptual performance is influenced by language. For example, Meteyard et al. (2007) assessed motion discrimination at threshold for displays of moving dots while participants passively listened to verbs that referred to either motion-related or static actions. Performance on the motion detection task was influenced by the words, with poorer performance observed on the perceptual task when the direction of motion implied by the words was incongruent with the direction of the dot display (see also, Lupyan and Spivey, 2010). Results such as these suggest a close integration of perceptual and conceptual systems (see Goldstone and Barsalou, 1998), an idea captured by the theoretical frameworks of grounded cognition (Barsalou, 2008) and embodied cognition (see Feldman, 2006; Borghi and Pecher, 2011).

There are limitations with tasks based on verbal reports or ones in which the emphasis is on accuracy. In such tasks, language may affect decision and memory processes, as well as perception (see Rosch, 1973). For example, in the Kay and Kempton (1984) study, participants were asked to select the two colored chips that go together best. Even though the stimuli are always visible, a comparison of this sort may engage top-down strategic processes (Pinker, 1997) as well as tax working memory processes as the participant shifts their attentional focus between the stimuli.

To reduce the contribution of memory and decision processes, researchers have turned to simple visual search tasks to explore the influence of language on perception. Consider a visual search study by Lupyan and Spivey (2008). Participants were shown an array of shapes and made speeded responses, indicating if the display was homogeneous or contained an oddball (**Figure 1A**). The shapes were the letters "2" and "5," rotated by 90°. In one condition, the stimuli were described by their linguistic labels. In the other condition, the stimuli were referred to as abstract geometric shapes. RTs were faster for the participants who had been given the linguistic labels or spontaneously noticed that the shapes were rotated letters. Lupyan and Spivey concluded that ". . . visual perception depends not only on what something looks like, but also on what it means" (p. 412).

Visual search has been widely employed as a model task for understanding early perceptual processing (Treisman and Gelade, 1980; Wolfe, 1992). Indeed, we have used visual search to show that the influence of linguistic categories in a detection task is amplified for stimuli presented in the right visual field (Gilbert et al., 2006, 2008). While our results provide compelling evidence that language can influence performance on elementary perceptual tasks, the mechanisms underlying this interaction remain unclear. Lupyan and Spivey (2008; Lupyan, 2008) suggest that the influence of language on perception reflects a dynamic interaction in which linguistic representations sharpen visual feature detectors.

FIGURE 1 | Sample stimulus displays for the No Cue (A) and Cue (B) conditions in Experiment 1. Participants made speeded responses, indicating if the display items were homogenous or contained an oddball. In the Cue conditions, the oddball matched the central cue.

By this view, feedback connections from linguistic or conceptual representations provide a mechanism to bias or amplify activity in perceptual detectors associated with those representations (Lupyan and Spivey, 2010), similar to how attentional cues may alter sensory processing (e.g., Luck et al., 1997; Mazer and Gallant, 2003).

While there is considerable appeal to this dynamic perspective, it is also important to consider alternative hypotheses that may explain how such interactions could arise at higher stages of processing (Wang et al., 1994; Mitterer et al., 2009; see also, Lupyan et al., 2010). Consider the Lupyan and Spivey task from the participants' point of view. The RT data indicate that the displays are searched in a serial fashion (Treisman and Gelade, 1980). When targets are familiar, participants compare each display item to an image stored in long-term memory, terminating the visual search when the target is found. With unfamiliar stimuli, the task is much more challenging (Wang et al., 1994). The participant must form a mental representation of the first shape and maintain this representation while comparing it to each display item. It is reasonable to assume that familiar shapes, ones that can be efficiently coded with a verbal label, would be easier to retain in working memory for subsequent use in making perceptual decisions (Paivio, 1971; Bartlett et al., 1980). In contrast, since unfamiliar stimuli lack a verbal representation in long-term memory, the first item would have to be encoded anew on each trial. We test the memory hypothesis in the following experiment, introducing a condition in which the demands on working memory are reduced.

EXPERIMENT 1

For two groups, the task was similar to that used by Lupyan and Spivey (2008): participants made speeded responses to indicate if a display contained a homogenous set of items or contained one oddball. For two other groups, a cue was present in the center of the display, indicating the target for that trial. Within each display type, one group was given linguistic primes by being told that the displays contained rotated 2's and 5's. The other group was told that the stimuli were abstract forms.

The inclusion of a cue was adopted to minimize the demands on working memory. By pairing the search items with a cue of the target, the task is changed from one requiring an implicit matching process in which each item is compared to a stored representation to one requiring an explicit matching process in which

each item is compared to the cue. If language influences perception by priming visual feature detectors, we would expect that participants who were given the linguistic labels would exhibit a similar advantage with both types of displays. In contrast, if the verbal labels reduce the demands on an implicit matching process (e.g., because the verbal labels provide for dual coding in working memory, see Paivio, 1971), then we would expect this advantage to be eliminated or attenuated when the displays contain an explicit cue.

MATERIALS AND METHODS
Participants
Fifty-three participants from the UC Berkeley Research Participation pool were tested. They received class credit for their participation. The research protocol was conducted in accordance with the procedures of the University's Institutional Review Board.

Stimuli
The visual search arrays consisted of 4, 6, or 10 white characters, presented on a black background. The characters were arranged in a circle. The characters were either a "5" or "2," rotated 90°clockwise. The characters fit inside a rectangle that measured 9 cm × 9 cm, and participants sat approximately 56 cm from the computer monitor. For the no cue (NC) conditions, a fixation cross was presented at the center of the display. For the Cue groups, the fixation cross was replaced by a cue.

Procedure
The participants were randomly assigned to one of four groups. The two NC groups provided a replication of Lupyan and Spivey (2008). They were presented with stimulus arrays (**Figure 1A**) and instructed to identify whether the display was composed of a homogenous set of characters, or whether the display included one character that was different than the others. One of the NC groups was told that the display contained 2's and 5's whereas the other NC group was told that the displays contained abstract forms. For the two Cue groups, the fixation point was replaced with a visual cue (**Figure 1B**). For these participants, the task was to determine if an array item matched the cue. As with the NC conditions, one of the Cue groups was told that the display consisted of 2's and 5's and the other Cue group was told that the display contained abstract forms.

Each trial started with the onset of either a fixation cross (NC groups) or cue (CUE groups). The search array was added to the display after a 300-ms delay. Participants responded on one of two keys, indicating if the display contained one item that was different than the other display items. Following the response, an accuracy feedback screen was presented on the monitor for 1000 ms. The screen was then blanked for a 500-ms inter-trial interval. Average RT and accuracy were displayed at the end of each block.

The experiment consisted of a practice block of 12 trials and four test blocks of 60 trials each. At the beginning of each block, participants in both the NC and Cue groups were informed which character would be the target for that block of trials, similar to the procedure used by Lupyan and Spivey (2008). Each character served as the oddball for two of the blocks. The oddball was present on 50% of trials, positioned on the right and left side of the screen with equal frequency.

At the end of the experiment, the participants completed a short questionnaire to assess their strategy in performing the task. We were particularly interested in identifying participants in the abstract groups who had generated verbal labels for the rotated 2's and 5's given that such strategies produced a similar pattern of results as the Cue group in the Lupyan and Spivey (2008) study. Three participants in the NC group and two participants in the Cue reported using verbal labels, either spontaneously recognizing that the symbols were tilted 2's and 5's, or creating idiosyncratic labels (one subject reported labeling the items "valleys" and "mountains"). These participants were replaced, yielding a total of 12 participants in each of the four groups for the analyses reported below.

RESULTS

Overall, participants were correct on 89% of the trials and there was no indication of a speed–accuracy trade-off. Excluding incorrect trials, we analyzed the RT data (**Figure 2**) in a three-way ANOVA with two between-subject factors (1) task description (linguistic vs. abstract) and (2) task set (NC vs. Cue), and one within-subject factor, (3) set size (4, 6, or 10 items). The effect of set size was highly reliable, consistent with a serial search process, $F(2, 88) = 289.35$, $p < 0.0001$. Importantly, the two-way interaction of task description and task set was reliable, $F(1, 44) = 4.96$, $p < 0.05$, and there was also a significant three-way interaction, $F(2, 88) = 6.23$, $p < 0.005$, reflecting the fact that the linguistic advantage was greatest for the largest set size, but only for the NC group.

To explore these higher-order interactions, we performed separate analyses on the NC and Cue groups. For the NC groups, the data replicate the results reported in Lupyan and Spivey (2008). Participants who were instructed to view the characters as rotated numbers (linguistic description) responded much faster compared to participants for whom the characters were described as abstract symbols. Overall, the RT advantage was 303 ms, $F(1, 22) = 10.12$, $p < 0.001$.

We used linear regression to calculate the slope of the search functions, restricting this analysis to the target present data. The

FIGURE 2 | Reaction time data for Experiment 1, combined over target present and target absent trials. Confidence intervals in the figure were calculated using the three-way interaction (Loftus and Masson, 1994).

mean slopes for the linguistic and symbol groups were 112 and 143 ms, respectively. This difference was not reliable, ($p = 0.10$). However, there was one participant in the symbol group with a negative slope (-2 ms/item), whereas the smallest value for all of the other participants in this group was at least 93 ms/item. When the analysis was repeated without this participant, the mean slope for the symbol group rose to 155 ms/item, a value that was significantly higher than for the linguistic group ($p = 0.03$). In summary, consistent with Lupyan and Spivey (2008), the linguistic cues not only led to faster RTs overall, but also yielded a more efficient visual search process.

A very different pattern of results was observed in the analysis of the data from the two Cue groups. Here, the linguistic advantage was completely abolished. In fact, mean RTs were slower by 46 ms for participants who were instructed to view the characters as rotated numbers, although this difference was not reliable $F(1, 22) = 0.072$, ns. Similarly, there was no difference in the efficiency of visual search, with mean slopes of 126 and 105 ms/item for the linguistic and symbol conditions, respectively. Thus, when the demands on working memory were reduced by the inclusion of a cue, we observed no linguistic benefit.

The results of Experiment 1 challenge the hypothesis that linguistic labels provide a top-down priming input to perceptual feature detectors. If this were so, then we would expect to observe a linguistic advantage regardless of whether the task involved a standard visual search (oddball detection) or our modified, matching task. *A priori*, we would expect that with either display, the linguistic description of the characters should provide a similar priming signal.

In contrast, the results are consistent with our working memory account. In particular, we assume that the linguistic advantage in the NC condition arises from the fact that participants must compare items in working memory during serial search, and that this process is more efficient when the display items can be verbally coded. Mean reaction time was faster and search more efficient (e.g., lower slope) when the rotated letters were associated with verbal labels. In this condition, each item can be assessed to determine if it matches the designated target, with the memory of the target facilitated by its verbal label (especially relevant here given that each target was tested in separate blocks). When the rotated letters were perceived as abstract symbols, the comparison process is slower, either because there is no verbal code to supplement the working memory representation of the target, or because participants end up making multiple comparisons between the different items.

The linguistic advantage was abolished when the target was always presented as a visual cue in the display. We can envision two ways in which the cue may have altered performance on the task. First, it would reduce the demands on working memory given that the cue provides a visible prompt. Second, it eliminates the need for comparisons between items in the display since each item can be successively compared to the cue. By either or both of these hypotheses, we would not expect a substantive benefit from verbal labels. RTs increase with display size, but at a similar rate for the linguistic and abstract conditions.

Mean RTs were slower for the Cue group compared to the NC group when the targets were described linguistically. This result
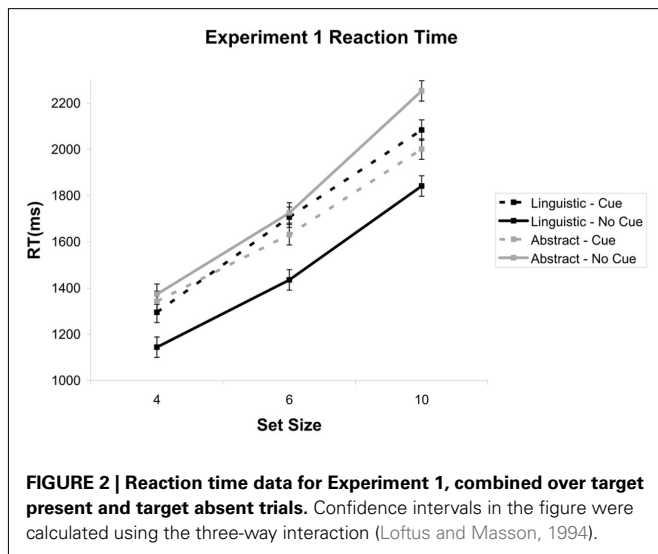
might indicate that the inclusion of the cues introduced some sort of interference with the search process. However, this hypothesis fails to account for why the slower RTs in the Cue condition were only observed in the linguistic group; indeed, mean RT was faster in the Cue condition for the abstract group. One would have to posit a rather complex model in which the inclusion of the cue somehow negated the beneficial priming from verbal labels.

Alternatively, the inclusion of the cue can be viewed as changing the search process in a fundamental way, with the task now more akin to a physical matching task rather than a comparison to a target stored in working memory. *A priori*, we cannot say which process would lead to faster RTs. However, the comparison of the absolute RT values between the Cue and NC conditions is problematic given the differences in the displays. One could imagine that there is some general cost associated with orienting to the visual cue at the onset of the displays for the Cue groups. Nonetheless, if the verbal labels were directly influencing perceptual detectors, we would have expected to see a persistent verbal advantage in the Cue condition, despite the slower RTs. The absence of such an advantage underscores our main point that the performance changes in visual search for the NC condition need not reflect differences in perception *per se*.

EXPERIMENT 2

We take a different approach in Experiment 2, testing the prediction that linguistic labels can disrupt processing when this information is task irrelevant. To this end, we had participants make an oddball judgment based on a physical attribute, line thickness. We presented upright or rotated 2s and 5s, assuming that upright numbers would be encoded as linguistic symbols, while rotated numbers would not. If language enhances perception, performance should be better for the upright displays. Alternatively, the automatic activation of linguistic codes for the upright displays may produce response conflict given that this information is irrelevant to the task.

MATERIALS AND METHODS
Participants
Twelve participants received class credit for completing the study.

Stimuli
Thick and thin versions of each character were created. The thick version was the same as in Experiment 1. For the thin version, the stroke thickness of each character was halved.

Procedure
Each trial began with the onset of a fixation cross for 300 ms. An array of four characters was then added to the display and remained visible for 450 ms (**Figure 3**). Participants were instructed to indicate whether the four characters had the same thickness, or whether one was different. The characters were either displayed in an upright orientation or rotated, with the same orientation used for all four items in a given display. Upright and rotated trials were randomized within a block. Each participant completed four blocks of 80 trials each. All other aspects were identical to Experiment 1.
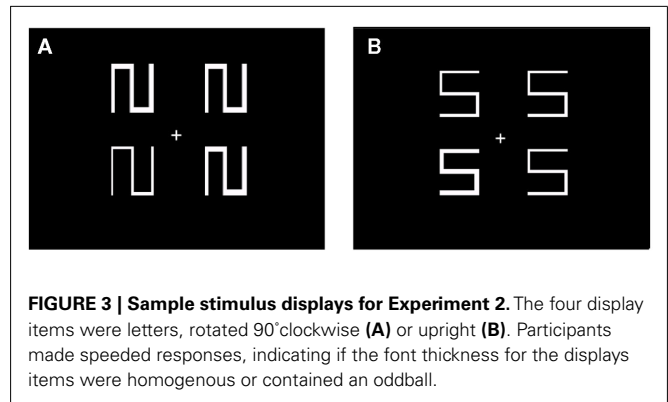

FIGURE 3 | Sample stimulus displays for Experiment 2. The four display items were letters, rotated 90˚clockwise **(A)** or upright **(B)**. Participants made speeded responses, indicating if the font thickness for the displays items were homogenous or contained an oddball.

RESULTS
Participants were slower when the characters were upright compared to when they were rotated, $F(1, 11) = 7.67$, $p < 0.01$. The mean RT was 375 ms for the upright displays and 348 ms for the rotated displays, for an average cost of 27 ms ($SE_{diff} = 5.6$ ms). Participants averaged 92% correct, and there was no evidence of a speed accuracy trade-off.

We designed this experiment under the assumption that the upright displays would produce automatic and rapid activation of the lexical codes associated with the numbers, and that these task-irrelevant representations would disrupt performance on the thickness judgments. We can envision at least two distinct ways in which linguistic codes might disrupt performance. Perceptually, linguistic encoding encourages holistic processing. If parts of a number are thick, there is a tendency to treat the shape in a homogenous manner, perhaps reflecting the operation of categorization (Fuchs, 1923; Prinzmetal and Keysar, 1989; Khurana, 1998). This bias may be reduced for the less familiar, rotated shapes, which may be perceived as separate lines.

Alternatively, the linguistic codes could provide potentially disruptive input to decision processes (e.g., response selection). This hypothesis is similar to the theoretical interpretation of the Stroop effect (MacLeod, 1991). In the classic version of that task, interference is assumed to arise from the automatic activation of the lexical codes of word names when the task requires judging the stimulus color, at least when both the relevant and irrelevant dimensions map onto similar response codes (e.g., verbal responses). In the current task, this interference would be more at a conceptual level (Ivry and Schlerf, 2008). Given that the four items in the display were homogenous, we would expect priming of the concept "same", relative to the concept "different", and that this would occur more readily for the upright condition where the items are readily recognized as familiar objects.

DISCUSSION
In the current study, we set out to sharpen the focus on how language influences perception. This question has generated considerable interest, reflecting the potential utility for theories of embodied cognition to provide novel perspectives on the psychological and neural underpinnings of abstract thought (Gallese and Lakoff, 2005; Feldman, 2006; Barsalou, 2008). An explosion of

empirical studies have appeared, providing a wide range of intriguing demonstrations of how behavior (reviewed in Barsalou, 2008) and physiology (Thierry et al., 2009; Landau et al., 2010; Mo et al., 2011) in perceptual tasks can be influenced by language. We set out here to consider different ways in which language might influence perceptual performance.

As a starting point, we chose to revisit a study in which performance on a visual search task was found to be markedly improved when participants were instructed to view the search items as linguistic entities, compared to when the instructions led the participants to view the items as abstract shapes (Lupyan and Spivey, 2008). The authors of that study had championed an interpretation and provided a computational model in which over-learned associative links between linguistic and perceptual representations allowed top-down effects of a linguistic cue to sharpen perceptual analysis.

While this idea is certainly plausible, we considered an alternative hypothesis, one that shifts the focus away from a linguistic modulation of perceptual processes. In particular, we asked if the benefit of the linguistic cues might arise because language, as a ready form of efficient coding, might reduce the burden on working memory. We tested this hypothesis by using identical search displays, with the one addition of a visual cue, assumed to minimize the demands on working memory. Under these conditions, we failed to observe any performance differences between participants given linguistic and non-linguistic prompts. These results present a challenge for the perceptual account, given the assumption that top-down priming effects would be operative for both the cued and non-cued versions of the task. Instead, the working memory hypothesis provides a more parsimonious account of the results, pointing to subtle ways in which performance entails a host of complex operations.

Our emphasis on how language might influence performance at post-perceptual stages of processing is in accord with the results from studies employing a range of tasks. In a particularly clever study, Mitterer et al. (2009) showed that linguistic labels bias the reported color of familiar objects. When presented with a picture of a standard traffic light in varying hues ranging from yellow to orange, German speakers were more likely to report the color as "yellow" compared to Dutch speakers, a bias consistent with the labels used by each linguistic group. Given the absence of differences between the two groups in performance with neutral stimuli, the authors propose that the effect of language is on decision processes, rather than by directly influencing perception.

It should be noted, however, that participants in the Mitterer et al. (2009) study were not required to make speeded responses; as such, this study may be more subject to linguistic influences at decision stages than would be expected in a visual search task. However, numerous visual search studies have also shown that RT in such studies is influenced by the degree and manner in which targets and distractors are verbalized (Jonides and Gleitman, 1972; Reicher et al., 1976; Wang et al., 1994). Consistent with the current findings, RTs are consistently slower when the stimuli are unfamiliar, an effect that has been attributed to the more efficient processing within working memory for familiar, nameable objects (e.g., Wang et al., 1994).

We recognize that language may have an influence at multiple levels of processing. That is, the perceptual and working memory accounts are not mutually exclusive, and in fact, divisions such as "perception" and "working memory" may in themselves be problematic given the dynamics of the brain. Nonetheless, we do think there is value in such distinctions since it is easy for our descriptions of task domains to constrain how we think about the underlying processes.

Indeed, this concern is relevant to some work conducted in our own lab. In a series of studies, we have shown that the effects of language on visual search is more pronounced in the right visual field (Gilbert et al., 2006, 2008). We have used a simple visual search task here, motivated by the goal of minimizing demands on memory processes and strategies. Our results, showing that task-irrelevant linguistic categories influence color discrimination, can be interpreted as showing that language has selectively shaped perceptual systems in the left hemisphere. Alternatively, activation of (left hemisphere) linguistic representations may be retrieved more readily for stimuli in the right, compared to left, visual field, and thus exert a stronger influence on performance. While the answer to this question remains unclear – and again, both hypotheses may be correct – the visual field difference disappears when participants perform a concurrent verbal task (Gilbert et al., 2006, 2008). This dual-task result provides perhaps the most compelling argument against a linguistically modified structural asymmetry in the perceptual systems of the two hemispheres. Rather, it is consistent with the post-perceptual account promoted here (see also Mitterer et al., 2009) given the assumption that the secondary task disrupted the access of verbal codes for the color stimuli, an effect that would be particular pronounced in the left hemisphere.

We extended the basic logic of our color studies in the second experiment presented here, designing a task in which language might hinder perceptual performance. We again used a visual search task, but one in which participants had to determine if a display item had a unique physical feature (i.e., font thickness). For this task, linguistic representations were irrelevant. Nonetheless, when the shapes were oriented to facilitate reading, a cost in RT was observed, presumably due to the automatic activation of irrelevant linguistic representations.

While linguistic coding can be a useful tool to aid processing, the current findings demonstrate that language can both facilitate and impede performance. Language can provide a concise way to categorize familiar stimuli; in visual search, linguistic coding would provide an efficient mechanism to encode and compare the display items (Reicher et al., 1976; Wang et al., 1994). However, when the linguistic nature of the stimulus is irrelevant to the task, language may also hurt performance (Brandimonte et al., 1992; Lupyan et al., 2010).

These findings provide a cautionary note when we consider how language and perception interact. No doubt, the words we speak simultaneously reinforce and compete with the dynamic world we perceive and experience. When language alters perceptual performance, is it tempting to infer a shared representational status of linguistic and sensory representations. However, even performance in visual search reflects memory, decision, and perceptual processes. We must be vigilant in characterizing the manner in which language and perception interact.

REFERENCES

Barsalou, L. W. (2008). Grounded cognition. *Annu. Rev. Psychol.* 59, 617–645.

Bartlett, J. C., Till, R. E., and Levy, J. C. (1980). Retrieval characteristics of complex pictures: effects of verbal encoding. *J. Verb. Learn. Verb. Behav.* 19, 430–449.

Borghi, A. M., and Pecher, D. (2011). Introduction to the special topic embodied and grounded cognition. *Front. Psychol.* 2:187. doi:10.3389/fpsyg.2011.00187

Boroditsky, L. (2001). Does language shape thought? Mandarin and English speakers' conceptions of time. *Cogn. Psychol.* 43, 1–22.

Brandimonte, M. A., Hitch, G. J., and Bishop, V. M. (1992). Verbal recoding of visual stimuli impairs mental image transformation. *Mem. Cognit.* 20, 449–455.

Ervin-Tripp, S. (1967). An Issei learns English. *J. Soc. Issues* 2, 78–90.

Feldman, J. (2006). From molecule to metaphor: A natural theory of language. Cambridge, MA: MIT Press.

Fuchs, W. (1923). "The influence of form on the assimilation of colours," in *A Source Book of Gestalt Psychology,* ed. W. D. Ellis (London: Routledge & Kegan Paul), 95–103.

Gallese, V., and Lakoff, G. (2005). The Brain's concepts: the role of the Sensory-motor system in conceptual knowledge. *Cogn. Neuropsychol.* 22, 455–479.

Gilbert, A. L., Regier, T., Kay, P., and Ivry, R. B. (2006). Whorf hypothesis is supported in the right visual field but not the left. *Brain Lang.* 103, 489–494.

Gilbert, A. L., Regier, T., Kay, P., and Ivry, R. B. (2008). Support for lateralization of the Whorf effect beyond the realm of color discrimination. *Brain Lang.* 105, 91–98.

Goldstone, R. L., and Barsalou, L. W. (1998). Reuniting perception and conception. *Cognition* 65, 231–262.

Ivry, R. B., and Schlerf, J. E. (2008). Dedicated and intrinsic models of time perception. *Trends Cogn. Sci. (Regul. Ed.)* 12, 273–280.

Jonides, J., and Gleitman, H. (1972). A conceptual category effect in visual search: O as letter or as digit. *Percept. Psychophys.* 12, 457–460.

Kay, P., and Kempton, W. (1984). What is the Sapir-Whorf hypothesis? *Am. Anthropol.* 86, 65–79.

Khurana, B. (1998). Visual structure and the integration of form and color. *J. Exp. Psychol. Hum. Percept. Perform.* 24, 1766–1785.

Landau, A. N., Aziz-Zadeh, L., and Ivry, R. B. (2010). The influence of language on perception: listening to sentences about faces affects the perception of faces. *J. Neurosci.* 30, 15254–15261.

Loftus, G. R., and Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychon. Bull. Rev.* 1, 476–490.

Loftus, E. F., and Palmer, J. C. (1974). Reconstruction of automobile destruction: an example of the interaction between language and memory. *J. Verb. Learn. Verb. Behav.* 13, 585–589.

Luck, S. J., Chelazzi, L., Hillyard, S. A., and Desimone, R. (1997). Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *J. Neurophysiol.* 77, 24–42.

Lupyan, G. (2008). The conceptual grouping effect: categories matter (and named categories matter more). *Cognition* 108, 566–577.

Lupyan, G., and Spivey, M. (2008). Perceptual processing is facilitated by ascribing meaning to novel stimuli. *Curr. Biol.* 18, R410–R412.

Lupyan, G., and Spivey, M. (2010). Making the invisible visible: auditory cues facilitate visual object detection. *PLoS ONE* 5, e11452. doi:10.1371/journal.pone.0011452

Lupyan, G., Thompson-Schill, S. L., and Swingley, D. (2010). Conceptual penetration of visual processing. *Psychol. Sci.* 21, 682–691.

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

Mazer, J. A., and Gallant, J. L. (2003). Goal-related activity in V4 during free viewing visual search. Evidence for a ventral stream visual salience map. *Neuron* 40, 1241–1250.

Meteyard, L., Bahrami, B., and Vigliocco, G. (2007). Motion detection and motion verbs: language affects low-level visual perception. *Psychol. Sci.* 18, 1007–1013.

Mitterer, H., Horschig, J. M., Musseler, J., and Majid, A. (2009). The influence of memory on perception: it's not what things look like, it's what you call them. *J. Exp. Psychol. Learn. Mem. Cogn.* 35, 1557–1562.

Mo, L., Xu, G., Kay, P., and Tan, L. H. (2011). Electrophysiological evidence for the left-lateralized effect of language on preattentive categorical perception of color. *Proc. Nat. Acad. Sci. U.S.A.* 108, 14026–14030.

Paivio, A. (1971). *Imagery and Verbal Processes.* New York: Holt, Rinehart & Winston.

Pinker, S. (1997). *How the Mind Works.* New York: W.W. Norton.

Prinzmetal, W., and Keysar, B. (1989). A functional theory of illusory conjunctions and neon colors. *J. Exp. Psychol. Gen.* 118, 165–190.

Reicher, G. M., Snyder, C. R. R., and Richards, J. T. (1976). Familiarity of background characters in visual scanning. *J. Exp. Psychol. Hum. Percep. Perform.* 2, 522–530.

Rosch, E. H. (1973). Natural categories. *Cogn. Psychol.* 4, 328–350.

Thierry, G., Athanasopoulos, P., Wiggett, A., Dering, B., and Kuipers, J. R. (2009). Unconscious effects of language-specific terminology on preattentive color perception. *Proc. Nat. Acad. Sci. U.S.A.* 106, 4567–4570.

Treisman, A., and Gelade, G. (1980). A feature-integration theory of attention. *Cogn. Psychol.* 12, 97–136.

Wang, Q., Cavanagh, P., and Green, M. (1994). Familiarity and pop-out in visual search. *Percept. Psychophys.* 56, 495–500.

Whorf, B. L. (1956). "Science and linguistics," in *Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf,* ed. J. B. Carroll (Cambridge, MA: MIT Press), 207–219.

Wolfe, J. M. (1992). The parallel guidance of visual attention. *Curr. Dir. Psychol. Res.* 1, 124–128.