

Comment and Reply

Resolving conflict: A response to Martin and Cheng (2006)

SHARON L. THOMPSON-SCHILL and
MATTHEW M. BOTVINICK
*University of Pennsylvania,
Philadelphia, Pennsylvania*

Martin and Cheng (2006) report the results of an experiment aimed at disentangling the effects of association strength from those of competition on performance on a verb generation task. Their experiment is situated at the center of a putative debate regarding the function of the left inferior frontal gyrus in language processing (see, e.g., Wagner, Paré-Blagoev, Clark, & Poldrack, 2001). Following in this tradition, Martin and Cheng purport to contrast two processes—selection between competing representations and controlled retrieval of weak associates—that we argue can be reduced to the same mechanism. We contend that the distinction between competition and association strength is a false dichotomy, and we attempt to recast this discussion within a Bayesian framework in an attempt to guide research in this area in a more fruitful direction.

Martin and Cheng (2006; henceforth MC06) report that performance on a verb generation task (measured by response times in young and elderly subjects and accuracy in a patient with left frontoparietal damage) was better for nouns either with one dominant verb associate (e.g., APPLE—eat) or with multiple strong verb associates (e.g., DOOR—open, DOOR—close) than for nouns with multiple weak verb associates (e.g., RUG—roll, RUG—lay). From this pattern, the authors conclude that “competition appeared to play no role in the verb generation task” (p. 400). This conclusion is predicated on the notion that association strength (between a noun and a verb) is unrelated to—or at least, partially dissociable from—competition, thus allowing these factors to be unconfounded. In this article, we challenge that assumption.

We begin with a brief review of “selection” and “controlled retrieval” hypotheses of the function of the left in-

ferior frontal gyrus (LIFG). We contend that terms such as *selection* and *controlled retrieval* have outlived their usefulness in this research enterprise, and that investigators would be better served by developing a more explicit framework in which the effects of manipulations such as association strength can be more formally assessed. To this end, we propose a Bayesian framework based on a small set of independently motivated assumptions. In addition, we present a simulation that addresses an observation from MC06 that would seem to pose difficulties for the proposed framework. It is beyond the scope of this comment to lay out a detailed defense of the model; rather, our goal is to present a rough sketch of a unifying theoretical perspective that is meant to begin the work of grounding this research in a more explicit computational context.

Verb Generation and Ventrolateral Prefrontal Cortex

Beginning with some of the earliest cognitive neuroimaging studies, the verb generation task has been utilized as a tool for exploring the mechanisms, and the neural substrates, involved in the retrieval of long-term world (or semantic) knowledge. Given the complexity of this task, it comes as no surprise that the interpretation of data from studies of verb generation remains somewhat controversial. Thompson-Schill, D’Esposito, Aguirre, and Farah (1997; henceforth TS97) challenged the claim that LIFG activation during verb generation was the result of semantic retrieval, because the magnitude of activation in LIFG increased under conditions thought to increase competition during selection among incompatible representations. Although this conclusion was based on three different manipulations of competition, we limit our discussion here to the verb generation task, since that is the focus of MC06.

Items used in the TS97 verb generation task were divided into two categories: “In the High Selection condition, items were nouns with many appropriate associated responses without any clearly dominant response; in the Low Selection condition, items were nouns with few associated responses or with a clear dominant response. . . . A ratio of the relative frequency of the most common completion to the relative frequency of the second-most common completion was calculated as a measure of response strength” (pp. 14792–14793). Activation in LIFG during verb generation was greater in response to high-selection than to low-selection items. A subsequent study reported that patients with LIFG lesions exhibited impairments generating verbs in response to items in the high-selection condition but not in the low-selection condition, and the magnitude of the impairment was correlated with the percentage of affected tissue in posterior LIFG (Thompson-

The preparation of this article was supported by the following grants from the National Institutes of Health: MH60414, MH67008, MH70850 (to S.L.T.-S.), and MH062196 (to M.M.B.). We thank Randi Martin for providing us with her stimuli in addition to very helpful comments on an earlier draft of this article; Doug Nelson for an equally useful set of comments on that draft; and Russell Epstein, Mike Kahana, Ken Norman, and Marius Usher for valuable conversations about these ideas in various stages of their genesis. Correspondence relating to this article may be sent to S. L. Thompson-Schill, Department of Psychology, University of Pennsylvania, 3720 Walnut Street, Philadelphia, PA 19104-6241 (e-mail: thompson@psych.upenn.edu).

Schill, Swick, Farah, D'Esposito, Kan, & Knight, 1998). These data indicated the necessity of LIFG for the process or processes that were varying with this manipulation of response dominance, which we described as selection among competitors.

Although additional evidence for this interpretation began to accumulate (see Thompson-Schill, 2005), Wagner and colleagues countered that LIFG “guides *controlled semantic retrieval* irrespective of whether retrieval requires selection against competing representations” (Wagner, Paré-Blagoev, Clark, & Poldrack, 2001, p. 329; emphasis in original). In support of this hypothesis, they reported greater LIFG activity during semantic similarity judgments of weakly associated word pairs (e.g., candle–halo) than of strongly associated word pairs (e.g., candle–flame). This association strength effect, which was assumed not to affect competition, was taken as evidence that LIFG mediates “a top-down bias signal that is recruited to the extent that the recovery of meaning demands controlled retrieval” (p. 329).

These two descriptions have been juxtaposed as mutually exclusive alternatives in subsequent research on this topic, as in: “Alternative models propose that [LIFG] guides top-down (controlled) retrieval of knowledge from long-term stores or selects goal-relevant products of retrieval from among competitors” (Badre, Poldrack, Paré-Blagoev, Inslar, & Wagner, 2005, p. 907). The most recent entry into this ostensible debate is MC06, as is clearly indicated by the article’s title. According to MC06, the verb generation effects reported in TS97 and Thompson-Schill et al. (1998) do not discriminate between the “selection hypothesis” and the “controlled retrieval hypothesis,” because response dominance is confounded with association strength in those studies. To break that confound, MC06 include a condition “in which there were two verb responses of relatively equal strength for a noun, both of which had high association strength” (p. 397; e.g., DOOR–open, DOOR–close). The logic for the inclusion of the DOOR condition¹ (and the resulting interpretation of the data) hinges on the assumption that selection demands are equated in the two high-selection conditions by virtue of the fact that response dominance is matched across the two: “If the difficulties in the high-selection condition demonstrated in the Thompson-Schill studies were caused by selection demands, we would expect no difference between the two high-selection conditions; both should be more difficult than the low-selection condition” (p. 397). Thus, the observed difference between RUG and DOOR conditions (and the similarity between DOOR and APPLE conditions) leads the authors to reject the hypothesis of TS97 in favor of the alternative, controlled retrieval hypothesis.

The Competition Hypothesis: Assumptions and Implications

MC06 conclude that the absence of a response dominance effect in their verb generation data refutes the hypothesis introduced in TS97. In so doing, they have failed to distinguish between the specific way in which that hypothesis was operationally defined in TS97 and the more

general hypothesis that our manipulation was designed to address. It is perhaps useful to reiterate our position here: Following from the seminal work of Desimone and Duncan (1995), who described visual attention as a top-down signal that biases competitive interactions between mutually inhibitory neurons, we proposed that LIFG functions bias competitive interactions among neurons representing conceptual knowledge (Kan & Thompson-Schill, 2004). That is, the role of LIFG in language tasks, including the verb generation task, is to resolve competition that occurs during selection among incompatible representations. The test of this hypothesis is provided by any manipulation that increases the activation levels of competitors.

Although this hypothesis has since been described as the “selection hypothesis,” this shorthand (for which we plead guilty!) perhaps highlights the wrong aspect of the original TS97 hypothesis. In subsequent articles, we have attempted to shift the focus back onto competition (e.g., Kan & Thompson-Schill, 2004; Thompson-Schill, 2005; Thompson-Schill, Bedny, & Goldberg, 2005), and thus we refer to our hypothesis here as the *competition hypothesis*. Our question here is thus whether an effect of association strength (in the absence of an effect of response dominance) refutes the competition hypothesis.

We propose an answer to this question that arises from three basic assumptions. First, we assume that the presentation of a stimulus triggers a pattern of activation across a set of response representations and that this pattern can be interpreted as representing a probability distribution, with the activation of each response representation reflecting the posterior probability that the relevant response is correct given the stimulus. This idea, which derives from independent psychological (e.g., Usher & McClelland, 2001) and neuroscientific (e.g., Pouget, Dayan, & Zemel, 2003) research, has the important consequence that the steady-state activation across output representations is normalized (i.e., it must sum to a constant).²

Second, we assume that a second process translates the probability distribution just stipulated into a single response (for relevant empirical data, see Gold & Shadlen, 2001). The probability of a response being selected is assumed to vary monotonically with its estimated posterior probability within the initial probability distribution, and the average latency of a response is assumed to vary inversely with this probability. The latter assumption formalizes the well-established monotonic relation between the time taken to make a response and the strength of the signal favoring that response (see, e.g., Murdock, 1985).

Our third assumption is that the probability distribution over response candidate representations is shaped not only by the stimulus, but also by a representation of the task. In the context of verb generation, the task representation would bias the distribution toward verb responses. One way of thinking about this effect is in terms of *prior probabilities*: The stimulus induces an initial probability distribution resembling the pattern of responses that would be observed in a free association task (but see note 2), and this distribution is then multiplied by a set of prior probabilities that scales up the probability of verbs and

scales down the probability of nonverbs, in keeping with the demands of the task. We will refer to the neural source of this control signal as LIFG, although the explanation of the basic framework need not assume any particular localization.

Consider the task of generating a verb related to the word *canoe* (one of the “high-selection, high-association” items used in MC06; see Figure 1, panel C, for free association norms for this item). We might roughly approximate the initial probability distribution induced by the stimulus *canoe* with these free association frequencies (where the probabilities for items represented by the response category “Other” would all have probabilities between 0 and .01). The presence of top-down input effectively adjusts this probability distribution, such that activation of units corresponding to *paddle*, *row*, and *swim* are raised and activation values of units corresponding to *boat*, *water*,

and so on, are lowered. The response latency to produce a verb would reflect the initial activation values, weighted by the task-relevant prior probabilities.

Because the effect of the task representation has different effects on task-relevant responses such as “row” and task-irrelevant responses such as “boat,” we can divide competitors into two classes: nonverb associates and weaker verb associates. Various experimental manipulations might affect one or both classes of competitors. For example, in Thompson-Schill, D’Esposito, and Kan (1999), LIFG activation was measured during a verb generation task in which each item could be primed with either the same task or an irrelevant task (i.e., color generation). The irrelevant-prime manipulation (which was associated with an increase in LIFG activation) explicitly targeted competition from nonverb associates (i.e., color). In contrast, in TS97, the manipulation was characterized

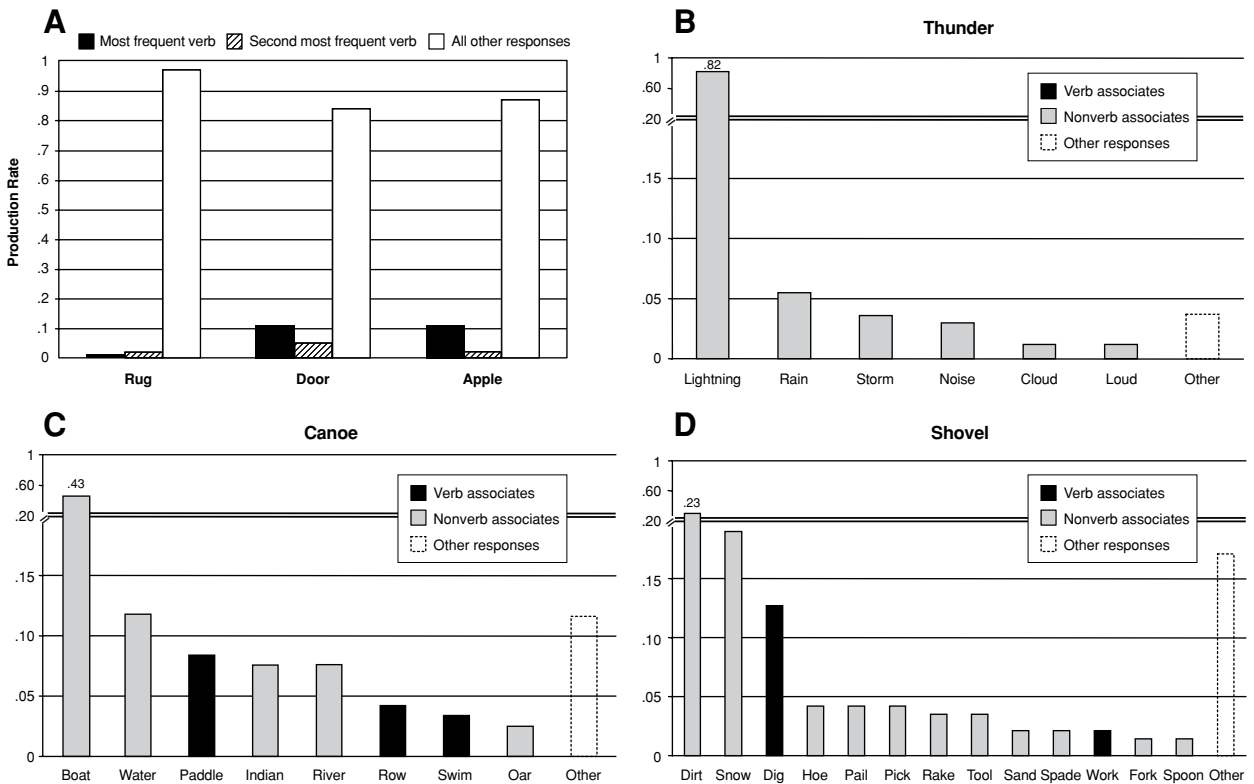


Figure 1. Free association rates from the South Florida word association norms (Nelson, McEvoy, & Schreiber, 1998) were obtained for the items used in MC06. Panel A indicates the mean proportions of subjects who produced the two verbs classified by MC06 as most frequent and second most frequent in their normative data; note that the rankings of the subjects and the South Florida norms do not always agree. The unfilled bar (“All other responses”) is the difference between the sum of those two proportions and 1.0. Panels B–D illustrate the free association rates for three items from MC06: *thunder* (high selection, low association; cf. RUG), *canoe* (high selection, high association; cf. DOOR), and *shovel* (low selection; cf. APPLE). Items were chosen that were closest to the mean in panel A with regard to both noun–verb association and verb₁/verb₂ ratio. The proportion of subjects who produced a particular target word (identified on the x-axis) in the presence of the cue word is listed for all responses reported in the norms. Responses given by fewer than 2 participants were deemed to be idiosyncratic and were excluded from the norms. The proportion of such responses (calculated by us) is plotted as “other.” Note that the scale on the y-axis shifts at the double horizontal line (at .20) to allow for a clearer picture of the patterns among the lower frequency items. These items illustrate a few simple observations: (1) As characterized by MC06, *canoe* and *shovel* have stronger verb associates than *thunder* does (which is also borne out in the average data in panel A), although we also see that “strong associate” is not the same as “strongest associate.” (2) *Thunder* has stronger nonverb associates than *canoe* and *shovel* do. (3) As characterized by MC06, *canoe* has stronger verb competitors than does *shovel*; this information cannot be estimated from the data for *thunder* because no verbs were produced in free association. However, (4) for both *canoe* and *shovel*, the strongest competitor is a nonverb.

as targeting competition among verb associates but may have also affected competition from nonverb associates.

Studies of verb generation in individuals with a sufficiently high error rate to enable an analysis of error types remind us of this often-overlooked source of competition in the verb generation task: One must focus on verb associates to the exclusion of nonverb associates that will inevitably spring to mind. MC06 report that a majority of M.L.'s errors were productions of nonverb responses. We observed a similar pattern in the data we described in Thompson-Schill et al. (1998); we suspect that this reflects the fact that items with a low response dominance ratio have weaker associates (as MC06 observed), and thus stronger nonverb competitors.

Once the door is opened (pun intended) to the influence of nonverb competitors, the assumption that RUG and DOOR are matched in terms of competitor strength becomes suspect. The three conditions in MC06 were matched on the association strength of the strongest nonverb associate (see their Table 1). This matching inadvertently created a potentially important difference between RUG and DOOR trials: Because the noun–verb association strength is on average higher in the DOOR condition than in the RUG condition (and lower in both than the nonverb association strength), the relative difference between the strongest verb and nonverb associates is lower for DOOR than for RUG trials.³ Thus, competition from other responses is greater for RUG trials than it is for DOOR (or for APPLE) trials. This is evident in Figure 1, in which we report the free association response rates for the items used in MC06 (both on average and for a representative item in each condition).

Indeed, once one has shifted to a distribution-based framework for thinking about the neural characterization of a stimulus (and its relationship to possible responses), the distinction between competition and association strength vanishes, and the argument above becomes unnecessary. If the likelihoods assigned to candidate responses are constrained to sum to 1 (which they must, if they are to be thought of as probabilities), then items with weaker associations *must* be accompanied by stronger associations for competing responses. Association strength and competition are just two sides of the same coin.

The story we have told thus far is fairly straightforward: RUG trials provide weaker evidence for any given verb than do APPLE or DOOR trials. Weaker associations translate to longer response times. LIFG functions to decrease activation of nonverb responses (in effect, changing the prior probability for the verbs), allowing a response word that would otherwise assume a low probability ranking to become the most highly activated response candidate. Patient M.L. is less able to make this adjustment, so he incorrectly produces nonverb responses to those items for which nonverb representations provide the strongest degree of competition (i.e., RUG items).

Thus, the differences reported in MC06 between RUG and DOOR trials emerge in a fairly straightforward way from the three computational assumptions we have just outlined. However, another potentially important aspect of MC06 is that performance for APPLE and DOOR trials

is equivalent. If nonverb competitors influence reaction times (as argued above), why don't verb competitors? In particular, once LIFG has adjusted prior probabilities in order to boost verbs relative to nonverbs, APPLE is left with no (significant) competitors, but DOOR has at least two representations fighting it out. According to our argument above, it would seem that the APPLE response should be faster than that of DOOR. To explain why intuition may fail here, we have turned to an explicit computational model of a response selection process.

Simulation of Response Competition Effects

Usher and McClelland (2001) proposed a simple but elegant network model of response selection processes. In line with the perspective just introduced, this model assumes that a stimulus input induces a probability distribution across possible responses. In the simple case that Usher and McClelland focus on, there are two available responses whose likelihoods are represented in the connection weights from the stimulus input, which are constrained to sum to 1 (see Figure 2). When the input unit is activated, both response units are activated to a degree determined by the connection weights from the input, by reciprocal inhibitory connections between the two response units, and by inherent noise. A response is understood to be produced when one of the response units crosses an activation threshold, and the number of processing cycles at which this occurs is identified as the reaction time.

The Usher and McClelland (2001) model can be used to model verb generation if the input is understood to be a noun and the response units to be response words.⁴ In order to investigate the relationship between the DOOR and APPLE conditions in MC06, we assumed that the noun stimuli in each case mapped to only two verb responses (one can think of nonverb associates of the noun input as inhibited by top-down input, and therefore irrelevant during response selection). In order to capture the difference between the two conditions, in the DOOR case the weights

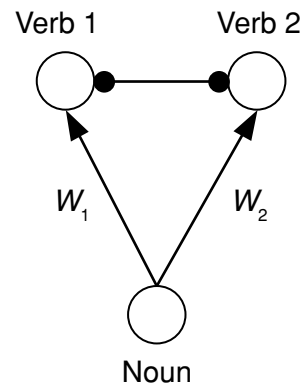


Figure 2. The Usher and McClelland (2001) model, as applied to verb generation. Connection weights w_1 and w_2 were set to .9 and .1 in order to simulate the APPLE condition, and to .5 and .5 in order to simulate the DOOR condition. The strength of the reciprocal inhibitory connection weights between the verb units varied as shown in Figure 3.

of each response unit were matched (.5, .5), and in the APPLE case they were assumed to be highly asymmetrical (.9, .1).⁵ These values were selected with the aim of exploring the extreme case, for maximal clarity, rather than of fitting data from MC06 quantitatively.

With such connection weights, it might seem inevitable that average response times for APPLE would be smaller than those for DOOR. However, the observed behavior of the model was not so simple. The relative speeds of APPLE and DOOR responses turned out to be quite parameter dependent, and there were parameter values well within the range explored by Usher and McClelland (2001) under which the two conditions yielded nearly identical reaction times, or in which DOOR was actually faster on average than APPLE. The relative speed of the two conditions depended, in particular, on the strength of the inhibitory weights running between the two response units—that is, on the degree of response competition (see Figure 3).

The behavior of the model depends on a critical assumption of the Usher and McClelland (2001) model, one shared by many accounts of speeded response selection (Ratcliff & Rouder, 1998; Shadlen & Newsome, 2001)—namely, that at the moment of stimulus onset the response system is not precisely at rest, but is instead in some possibly random initial state. In our verb generation simulation, one of the verb response units tended to have slightly more activation than the other at the onset of the noun input. In the APPLE condition, this could have the effect of delaying response production if the weaker verb associate (connection weight .1) was initially more active, since the stronger associate had then to overcome this initial imbalance in order to eventually win out. The assumption of a slight (and random) initial bias thus gave the DOOR condition a reaction time advantage over the APPLE condition, which weighed against the fact that APPLE involved one very strong connection weight. The balance between these

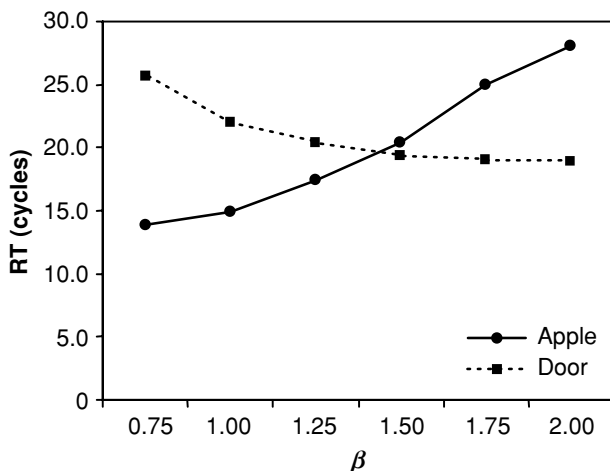


Figure 3. Mean response times (in processing cycles) for the model depicted in Figure 2, as the strength of reciprocal inhibition (β) increased from 0.75 to 2.00. Each data point is based on a sample of 1,000 trials.

two factors was found to depend on the parameter controlling the strength of reciprocal inhibition between the two response units. The larger this parameter becomes, the larger the impact of the initial state on the APPLE condition, and the larger the average reaction time in this condition relative to DOOR (see Figure 3).

Needless to say, this simulation involves many simplifications and can only be considered a rough and preliminary sketch of a proper computational account. Nevertheless, the simulation suffices to undermine the conclusion MC06 draw from their finding of similar reaction times for the DOOR and APPLE conditions, which is that “competition appeared to play no role in the verb generation task.” Our simulation illustrates how the data might in fact support the opposite conclusion, since it shows how response competition could, in principle, be directly responsible for the parity between APPLE and DOOR.

LIFG and Conflict Resolution

The impetus for this whole debate (which we have tried here to reframe) is the question of the role of LIFG in the verb generation task. We have clearly articulated one role, which would be best characterized as a top-down (i.e., stimulus-independent) adjustment on the prior probabilities of each response candidate. Although we have not implemented this idea in a working model, one can imagine a biasing process that adjusts the probability distributions so that all of the verbs are higher than all of the nonverbs (imagine pushing and pulling the bars of each plot in Figure 1). A failure of this biasing mechanism would result in an error rate that is correlated with the strength of nonverb competition, which is a plausible characterization of the pattern of data from Patient M.L.⁶

Is this a “selection” mechanism or a “controlled retrieval” mechanism? We argue that it is both, or rather, one and the same. Consider these two descriptions of putatively competing hypotheses: First, Thompson-Schill (2005) wrote, “In response to the target ‘scissors,’ the strongly-associated action ‘cut’ might be activated from the input. In contrast, in response to the target ‘cat,’ the activation of many weakly associated actions (e.g., ‘scratch,’ ‘purr’) and/or of a strongly associated non-action (e.g., dog) might fail to produce sufficient activation to select any action representation. Both of these situations (underdetermined representations and prepotent representations)⁷ can induce conflict among active representations in working memory that requires top-down intervention. . . . We suggest that this intervention comes in the form of a modulatory signal from prefrontal cortex that aids in the selection of an appropriate representation” (pp. 177–178). Second, from Wagner and colleagues: “the manipulation of associative strength, which we previously argued to impact controlled retrieval demands without consequences for selection, . . . can also result in variable competition because of the presence of irrelevant competitors or an ‘underdetermined response’ on weak trials” (Badre et al., 2005, p. 914). The similarity between these accounts illustrates what we see as a convergence on a unifying framework that we have attempted to articulate here.

Summary

Selection versus controlled retrieval. Association strength versus competition. Although these terms have played a useful heuristic role, they have increasingly come to seem too vague and pliable. Indeed, when placed within the context of a probabilistic framework like the one we have sketched in the present response, these terminological contrasts break down: Any manipulation of association strength must also be a manipulation of competition. Selection and controlled retrieval both involve the integration of top-down with bottom-up information sources.

In our opinion, the contributions of MC06 to this literature are of a far more interesting nature than the ostensible debate between selection and controlled retrieval. Their findings challenge a simple mapping between response dominance and competition and reveal counterintuitive aspects of behavior that may have interesting theoretical and empirical implications.⁸ We note with great interest that M.L., under this framework, would be described as having a problem adjusting the priors on verbs but not regulating conflict that might emerge between multiple, incompatible representations. In contrast, other patients do indeed seem to have a problem with the latter type of conflict (see, e.g., Robinson, Blair, & Cipolotti, 1998). This distinction may bear on a pressing question concerning the functional organization of LIFG, and of prefrontal cortex more generally, with regard to cognitive control mechanisms (cf. Badre et al., 2005).

We hope to see an end to debates about the “selection hypothesis” versus the “controlled retrieval hypothesis.” Instead, we have offered an explanation of the data in MC06 that relies on a Bayesian account of the verb generation task in conjunction with a simple simulation of response conflict: The former explains the relationship between DOOR and RUG trials, and the latter, which is consistent with and complementary to the first, explains the relationship between DOOR and APPLE trials. These explanations cut across entrenched distinctions in the literature, to offer instead a unifying conceptual framework and numerous points of departure for further research. Regardless of whether these particular hypotheses bear fruit, we believe that a consideration of competition, conflict, and control that is more computationally explicit will be essential to advancing our understanding of the functions of prefrontal cortex.

REFERENCES

- BADRE, D., POLDRACK, R. A., PARÉ-BLAGOEV, E. J., INSLER, R. Z., & WAGNER, A. D. (2005). Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron*, *47*, 907-918.
- BOTVINICK, M. M., BRAVER, T. S., BARCH, D. M., CARTER, C. S., & COHEN, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*, 624-652.
- DESIMONE, R., & DUNCAN, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193-222.
- GOLD, J. I., & SHADLEN, M. N. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends in Cognitive Sciences*, *5*, 10-16.
- HAMILTON, A. C., & MARTIN, R. C. (2005). Dissociations among tasks involving inhibition: A single-case study. *Cognitive, Affective, & Behavioral Neuroscience*, *5*, 1-13.
- KAN, I. P., & THOMPSON-SCHILL, S. L. (2004). Selection from perceptual and conceptual representations. *Cognitive, Affective, & Behavioral Neuroscience*, *4*, 466-482.
- MARTIN, R. C., & CHENG, Y. (2006). Selection demands versus association strength in the verb generation task. *Psychonomic Bulletin & Review*, *13*, 396-401.
- MURDOCK, B. B., JR. (1985). An analysis of the strength-latency relationship. *Memory & Cognition*, *13*, 511-521.
- NELSON, D. L., DYRDAL, G. M., & GOODMON, L. B. (2005). What is preexisting strength? Predicting free association probabilities, similarity ratings, and cued recall probabilities. *Psychonomic Bulletin & Review*, *12*, 711-719.
- NELSON, D. L., MCEVOY, C. L., & SCHREIBER, T. A. (1998). *The University of South Florida word association, rhyme, and word fragment norms*. Available at w3.usf.edu/FreeAssociation/.
- POUGET, A., DAYAN, P., & ZEMEL, R. S. (2003). Inference and computation with population codes. *Annual Review of Neuroscience*, *26*, 381-410.
- RATCLIFF, R., & ROUDER, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, *9*, 347-356.
- ROBINSON, G., BLAIR, J., & CIPOLOTTI, L. (1998). Dynamic aphasia: An inability to select between competing verbal responses? *Brain*, *121*, 77-89.
- SHADLEN, M. N., & NEWSOME, W. T. (2001). Neural basis of a perceptual decision in the parietal cortex (Area LIP) of the rhesus monkey. *Journal of Neurophysiology*, *86*, 1916-1936.
- THOMPSON-SCHILL, S. L. (2005). Dissecting the language organ: A new look at the role of Broca's area in language processing. In A. Cutler (Ed.), *Twenty-first century psycholinguistics: Four cornerstones* (pp. 173-189). Mahwah, NJ: Erlbaum.
- THOMPSON-SCHILL, S. L., BEDNY, M., & GOLDBERG, R. F. (2005). The frontal lobes and the regulation of mental activity. *Current Opinion in Neurobiology*, *15*, 219-224.
- THOMPSON-SCHILL, S. L., D'ESPOSITO, M., AGUIRRE, G. K., & FARAH, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences*, *94*, 14792-14797.
- THOMPSON-SCHILL, S. L., D'ESPOSITO, M., & KAN, I. P. (1999). Effects of repetition and competition on activity in left prefrontal cortex during word generation. *Neuron*, *23*, 513-522.
- THOMPSON-SCHILL, S. L., SWICK, D., FARAH, M. J., D'ESPOSITO, M., KAN, I. P., & KNIGHT, R. T. (1998). Verb generation in patients with focal frontal lesions: A neuropsychological test of neuroimaging findings. *Proceedings of the National Academy of Sciences*, *95*, 15855-15860.
- USHER, M., COHEN, J. D., SERVAN-SCHREIBER, D., RAJKOWSKI, J., & ASTON-JONES, G. (1999). The role of locus coeruleus in the regulation of cognitive performance. *Science*, *283*, 549-554.
- USHER, M., & MCCLELLAND, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, *108*, 550-592.
- VIGLIOCCO, G., VINSON, D. P., DAMIAN, M. F., & LEVELT, W. (2002). Semantic distance effects on object and action naming. *Cognition*, *85*, B61-B69.
- WAGNER, A. D., PARÉ-BLAGOEV, E. J., CLARK, J., & POLDRACK, R. A. (2001). Recovering meaning: Left prefrontal cortex guides controlled semantic retrieval. *Neuron*, *31*, 329-338.

NOTES

1. For the reader's ease, we will refer to the conditions in MC06 by their exemplars: APPLE (i.e., low selection, high association), DOOR (i.e., high selection, high association), and RUG (i.e., high selection, low association). The conditions are defined with reference both to the response dominance ratio (to establish high or low selection, following TS97) and to association strength.

2. The posterior probability distribution can be thought of as the strength of evidence for a response, which is not equivalent to the probability of a response, although they are clearly related to each other by some function. This is implied in the division we assume between a stage of *representation* that computes the posterior probability of each response word being the appropriate response in the present context, and

a *response selection* stage that translates this distribution into a single response (our second assumption). We draw this distinction for two reasons: First, Nelson and colleagues (e.g., Nelson, Dyrdal, & Goodman, 2005) have explored the relationship between cue–target association strength and competitor strength in free association data. In their work, competitor strength and association strength are negatively correlated, but imperfectly, since the response rates are not constrained to sum to 1 (because idiosyncratic responses deemed to be unreliable are not included in their analysis). In contrast, when the relationship between a stimulus and a response is characterized as a probability distribution, the strength between a stimulus and any one response must be perfectly negatively correlated with the strength between that stimulus and all other responses (i.e., competitor strength). Second, we use the terms *response* and *response word* here somewhat imprecisely. Although we have implemented verb generation with two layers (input and output), most models of language production assume multiple stages of processing. For example, one could think of our response units as competing semantic representations, which could feed into a subsequent stage of processing (e.g., phoneme selection).

3. Note that the figures given in Table 1 of MC06 with regard to noun–verb association strength and noun–nonverb association strength cannot be directly compared, because the former are drawn from verb association data collected by MC06 and the latter are drawn from normative free association data. However, the relative differences between the strongest verb and the strongest noun associate in each condition can be inferred from these data.

4. The activations in the response layer of the Usher and McClelland model relate more closely to the response selection stage we described in our second core assumption. We interpret the noun-to-verb weights as a proxy for the likelihood distribution over response words that would, in a fuller model, be represented in a pattern of activation over an intermediate layer (see, e.g., Usher, Cohen, Servan-Schreiber, Rajkowski, & Aston-Jones, 1999).

5. The remainder of the simulation was conducted precisely as prescribed by Usher and McClelland (2001), with the following parameters: $\alpha = 0$, $\sigma = 0.5$, $\theta = 1.057$, $\tau = 0.1$, $k = 0.2$, $i_0 = 0$, preparatory cycles = 25, total cycles per trial = 1,000. The β parameter varied between 0.75 and 2.00, as shown in Figure 3.

6. This impairment could be akin both to M.L.'s tendency to false alarm to familiar (but incorrect) items on a short-term memory test and to his exaggerated interference effect for the classic Stroop task, two deficits that were previously hypothesized to result from a failure of

inhibition in the verbal domain (Hamilton & Martin, 2005). This interpretation appears to be more consistent with the present proposal than with the characterization of M.L.'s deficit provided in MC06.

7. Thompson-Schill (2005) borrowed the distinction between prepotent response override and underdetermined responding from work on conflict monitoring (Botvinick, Braver, Barch, Carter, & Cohen, 2001). Response override creates conflict in situations that require one to override a strong response in favor of a weaker (but task-appropriate) response; naming the ink color on an incongruent trial of the Stroop task is a paradigmatic example of a situation requiring response override. Underdetermined responding creates conflict in situations in which the stimulus does not uniquely specify the appropriate response. Conflict emerges among a set of weakly activated responses, none of which is more compelling than the others. Both types of conflict are likely to occur on the verb generation task.

8. Although the present model is too provisional to permit us to make novel predictions with much confidence, we are intrigued by several implications of the β -parameter effect.

First, one might predict that two responses that are themselves closely related (e.g., *paddle* and *row*, the two strongest verb responses to CANOE) would inhibit each other more than two that are distantly related (e.g., *sit* and *eat*, the two strongest verb responses to TABLE). With other word production tasks (e.g., semantic interference during picture naming; Vigliocco, Vinson, Damian, & Levelt, 2002), the amount of interference is greater between concepts that are more closely related; by extension, we might predict faster RTs for items like CANOE than for items like TABLE.

Second, the consequences of varying β in the simulation raise the possibility that β may be an adjustable parameter that reflects another form of cognitive control: Increasing lateral inhibition between all responses (i.e., increasing β) could serve to reduce the effects of competition in cases of indeterminacy.

Finally (and moving from mild to rampant speculation), one could then predict that neurophysiological changes affecting the magnitude of lateral inhibition would have consequences for the pattern of response latencies for DOOR and APPLE trials, so we cannot resist commenting on what appears to be a trend in MC06 for an effect of aging in this regard (for their older controls, mean difference between DOOR and APPLE trials = 350 msec, $p = .22$).

(Manuscript received October 18, 2005;
revision accepted for publication April 6, 2006.)