Evaluating the distinction between semantic knowledge and semantic access: Evidence from semantic dementia and comprehension-impaired stroke aphasia

Curtiss A. Chapman¹ · Omar Hasan² · Paul E. Schulz² · Randi C. Martin¹

Published online: 28 January 2020

© The Psychonomic Society, Inc. 2020

Abstract

Theories of semantic memory based on neuropsychological findings have posited a distinction between stored semantic representations and the mechanisms used to access and manipulate them (e.g., Lambon Ralph, Jefferies, Patterson, & Rogers, 2017; Warrington & Cipolotti, 1996). The most recent instantiation of this view, the controlled semantic cognition theory (Lambon Ralph et al., 2017), is supported by findings suggesting that multimodal (i.e., both verbal and nonverbal) semantic deficits may result from qualitatively different impairments: on the one hand, damage to a semantic access mechanism related to executive control, which is observed in semantic aphasia (SA), and on the other, damage to semantic representations, which is observed in semantic dementia (SD) (Jefferies & Lambon Ralph, 2006). In this study we compared SA and SD patients on several phenomena previously used to support these distinctions. Contrary to the prior results, we found that (1) overall, cross-task consistency was equivalent for the two groups; (2) neither patient group showed consistency driven by item identity across different semantic tasks; (3) correlations among task performance were not obviously driven by the semantic control demands of different tasks; (4) both groups showed executive function deficits; and (5) both groups showed strong effects of distractor interference in a synonym judgment task. Furthermore, we investigated the components of executive ability that could underlie semantic control deficits by correlating performance on updating, shifting, and inhibition tasks with performance on tasks testing semantic abilities. We found that updating was related to semantic processing generally, whereas shifting and inhibition were not. These results also suggest that complex executive function tasks relate to semantic tasks through their shared relationship with language abilities. Overall, evidence from SA and SD patients does not differentiate representations and access mechanisms in the semantic system, as has previously been suggested. Implications for the storage–access distinction are discussed.

Keywords Aphasia · Semantic dementia · Semantic memory · Executive function

Semantic memory is a crucial component of everyday human functioning that helps us to understand sounds, interact with objects and people, and engage with the environment in which we live. When semantic memory is impaired, as it is in patients with semantic dementia (SD) or comprehension-impaired aphasia due to stroke (referred to here as *semantic aphasia*; SA), interacting with the world via language and

meaning becomes difficult. Because SD and SA patients are often impaired to a similar degree on tests of semantic ability but have different sites of brain damage, researchers have long been interested in whether there are functional cognitive differences between these patients' deficits.

Early investigations comparing SA and SD patients suggested that SD patients suffer from a selective loss of semantic memory, a “storage” deficit, whereas SA patients suffer from the inability to reliably retrieve stored representations, a refractory “access” deficit (McCarthy & Kartsounis, 2000; Warrington, 1975; Warrington & Cipolotti, 1996; Warrington & McCarthy, 1983, 1987; Warrington & Shallice, 1979). A striking aspect of refractory access deficits is that these patients show poorer performance on repeated testing of the same item, particularly when the items are selected from the same semantic category. Warrington and colleagues hypothesized that once an item was accessed it went into a temporary refractory state. This

Electronic supplementary material The online version of this article (<https://doi.org/10.3758/s13423-019-01706-6>) contains supplementary material, which is available to authorized users.

✉ Randi C. Martin
rmartin@rice.edu

¹ Department of Psychology, Rice University, Houston, TX, USA

² Department of Neurology, McGovern Medical School, University of Texas Health Science Center at Houston, Houston, TX, USA

refractory state is brief for healthy individuals but longer for these patients, causing difficulty in subsequent semantic access. In these investigations, SA and SD patients were demonstrated to be differentially sensitive to four major task factors related to their storage or refractory access deficits (Warrington & Cipolotti, 1996):

- (1) *presentation rate*: SD patients were unaffected by how quickly stimuli were presented, whereas SA patients showed significantly improved performance with slowed presentation, as a longer time increased the likelihood of exiting the refractory state;
- (2) *consistency*: SD patients showed consistent performance with the same items across repetitions within the same task, whereas SA patients showed inconsistent performance attributed to random fluctuation in the likelihood of entering a sustained refractory state;
- (3) *word frequency*: SD patients were highly sensitive to word frequency, whereas SA patients were insensitive to word frequency, as highly frequent words were assumed to have more robust representations in the knowledge store but frequency was assumed to be unrelated to the probability of being in a refractory state; and
- (4) *semantic relatedness*: SA patients performed more poorly on tasks such as spoken-word, written-word matching when the arrays of words to select from were semantically close relative to arrays with semantically distant stimuli, which was attributed to the refractory state spreading to semantically related concepts.

More recent studies have maintained the distinction between storage and access deficits, though with some modifications to the conceptualization of each. As in the earlier approach, storage deficits are assumed to be due to a loss of semantic knowledge per se. According to the controlled semantic cognition (CSC) framework put forward by Jefferies and Lambon Ralph and colleagues (Jefferies & Lambon Ralph, 2006; Lambon Ralph, Jefferies, Patterson, & Rogers, 2017), this storage deficit is assumed to result from damage to a semantic hub, that is, a region that brings together disparate semantic information about a concept (such as sensory features from all modalities and nonperceptual knowledge such as function for natural or man-made objects). Damage to the hub will thus affect processing of that concept, irrespective of the specific features assessed or the task demands of any task. Access deficits are assumed to derive from a disruption of executive control that affects the semantic system—a “semantic control” deficit rather than a refractory access deficit (Jefferies & Lambon Ralph, 2006; Lambon Ralph et al., 2017). Semantic control, as defined by the CSC framework is the ability to flexibly direct activation through the semantic system and to select correctly from among

competing representations. This ability is claimed to be crucial, for instance, in focusing on uncommon meanings of ambiguous words (e.g., the “wine” meaning of “port”) or recontextualizing the function of an object, as is done when using a kitchen knife not only to cut a sandwich but also to open a package. Thus, this approach does not assume that items enter a refractory state after being accessed, but instead assumes that these items may be difficult to access on the first attempt. Moreover, there is no assumption that items will become increasingly difficult to access upon repeated presentation within a short time frame. Because there is no assumption that items enter a time-dependent refractory state, there is likewise no assumption that the presentation rate of stimuli will affect performance.

In SA patients, the CSC framework suggests that impaired semantic control is reflected in a number of behavioral phenomena: (1) poor performance on executive tasks due to the executive nature of the deficit and/or lesions that overlap with brain regions involved in executive control; (2) poor performance on more demanding semantic tasks due to the tasks’ presumed executive demands; (3) poor ability to suppress the influence of semantically related distractor items due to the competition inherent in selecting among multiple active representations; (4) strong influence of the diversity of semantic contexts in which a word appears on comprehension because the diverse sets of semantic features associated with different contexts cause greater competition in concept selection; (5) lack of benefit for high frequency stimuli because frequency and semantic diversity are positively correlated and therefore poorer performance with highly diverse words negates the typical benefit of high frequency; (6) inconsistent performance with the same items across tasks with different semantic control demands because access to representations is a function of the executive demands of the task; (7) the presence of associative errors in picture naming (e.g., saying ‘nail’ for a picture of a hammer) due to errors of a faulty selection mechanism when multiple, competing representations are activated; and (8) strong benefit from helpful cues and strong detriment from miscues due competition among representations being alleviated in the former case and exacerbated in the latter case (Hoffman, Jefferies, & Lambon Ralph, 2011a; Hoffman, Rogers, & Lambon Ralph, 2011b; Jefferies & Lambon Ralph, 2006; Jefferies, Patterson, & Lambon Ralph, 2008b; Noonan, Jefferies, Corbett, & Lambon Ralph, 2010; Noonan, Jefferies, Garrard, Eshan, & Lambon Ralph, 2013a). Some of these phenomena have been directly contrasted with SD patients’ performance. As is described in the refractory access literature, SD patients in these comparisons tend to show exaggerated benefits of high word frequency; additionally, SD patients do not tend to show associative errors or benefits of correct phonemic cues in picture naming (Jefferies & Lambon Ralph, 2006; Jefferies, Patterson, & Lambon Ralph, 2008b).

Other contrasting behaviors, including consistency and executive performance are described in more detail below.

The behavioral phenomena thought to reflect semantic control deficits in the CSC framework partially overlap with those claimed by ‘refractory access’ account cited above, but researchers have argued that semantic control and refractory access deficits are dissociable. A recent study argued that refractory effects (i.e., increasingly poor performance with repeated presentation, strong effects of stimulus presentation rate, and strong effects of semantic relatedness among stimuli) may be found in SA patients with frontal lobe damage, whereas semantic control deficits without refractory effects may be seen in patients with temporoparietal damage (Jefferies, Baker, Doran, & Lambon Ralph, 2007). However, this argument is problematic for two reasons. First, none of the six patients with frontal lesions in this study consistently showed refractory effects of presentation rate and repetition that were significant by traditional standards ($p < .05$). What distinguished the six frontal patients from the two nonfrontal patients was the more tenuous fact that the frontal patients *sometimes* showed significant effects of presentation rate or repetition, and nonfrontal patients *rarely* (rather than never) showed these effects. However, the small sample size, the lax significance criteria (i.e., one-tailed tests, $p < .10$), and the failure to test whether the frontal patients showed significantly greater refractory effects than nonfrontal patients suggest that these results are not reliable for disentangling the anatomy of refractory versus semantic control phenomena. Second, another recent study shows that symptoms of a refractory access deficit may be generated by damage to regions outside the frontal lobe. Campanella, Mondani, Skrap, and Shallice (2009) presented evidence from six temporal lobe glioma patients whose lesions overlapped in the left posterior middle and superior temporal gyri and retrosubicular cortex. These patients, despite nonfrontal damage in four of six patients, showed several indices of access deficits (i.e., inconsistency with the same item over repeated testing, absent frequency effects, strong semantic distance effects, and effects of presentation rate). Like Jefferies et al. (2007), this study’s sample was quite small. However, it shows in principle that refractory deficits are not exclusive to patients with frontal brain damage. Thus, a distinctive neural basis of refractory access versus semantic control is not firmly established.

In support of the contention that SA patients have semantic control deficits, researchers have pointed to these patients’ brain damage being localized in regions argued to underlie executive control and competitive concept selection, including the left inferior frontal lobe and temporoparietal cortex (Badre & Wagner, 2002; Duncan, 2010; Jefferies & Lambon Ralph, 2006; Noonan, Jefferies, Visser, & Lambon Ralph, 2013b; Thompson-Schill, D’Esposito, Aguirre, & Farah, 1997; Wagner, Paré-Blagoev, Clark, & Poldrack, 2001). These lesions contrast with the predominant locations of

cortical atrophy in SD patients, the anterior temporal lobes, which are typically associated with stored semantic representations (Gorno-Tempini et al., 2011). Although some early studies implicated the temporal poles as the most critical regions underlying the deficits in SD patients and associated them with the site of the semantic hub (Mummery et al., 2000; Patterson, Nestor, & Rogers, 2007), more recent studies have instead pointed to a more inferior and posterior region as the localization of the semantic hub, the anterior fusiform gyrus (Lambon Ralph et al., 2017; Mion et al., 2010; see also Forseth et al., 2018, for data from electrocorticography and direct cortical stimulation). Studies that create virtual lesions using TMS have provided further evidence that the temporal lobes support semantic knowledge and frontal and temporoparietal regions support semantic selection or executive control, though stimulation sites in the temporal lobe have not been situated at proposed hubs in the temporal pole or the anterior fusiform gyrus (e.g., Pobric, Jefferies, & Lambon Ralph, 2010; Whitney, Kirk, O’Sullivan, Lambon Ralph, & Jefferies, 2011).

The suggestion that executive-function-related difficulties underlie the distinct behavior of SA patients is appealing in its convergence with numerous studies that claim executive functions have an important role in retrieving linguistic representations from semantics (e.g., Biegler, Crowther, & Martin, 2008; Schnur, Schwartz, Brecher, & Hodgson, 2006; Sikora, Roelofs, Hermans, & Knoors, 2016). However, although the distinctive pattern of behavioral deficits and the correlations with different brain regions seem compelling, a closer look at the empirical data supporting the differentiation between the semantic deficits of SD and SA patients reveals some findings that do not fit easily with the storage–access distinction, as well as significant gaps in the available data. The sections below outline these limitations as a basis for motivating the present study, which addresses these inconsistencies and gaps in the literature.

Item and task consistency

Recent studies investigating semantic control versus storage deficits have examined consistency in two ways: consistency of performance on the same items across different semantic tasks (item consistency; e.g., picture naming and word–picture matching on the same items), and consistency between individuals’ mean performance across items in various semantic tasks (task consistency; e.g., Bozeat, Lambon Ralph, Patterson, Garrard, & Hodges, 2000b; Corbett, Jefferies, Eshan, & Lambon Ralph, 2009a; Jefferies & Lambon Ralph, 2006). Note that item consistency in these semantic control studies is similar to that studied in refractory access studies, albeit measured across tasks and modalities rather than across repeated presentations of the same items in the same task.

Task consistency, however, is independent of item identity and may thus be found regardless of item consistency (i.e., individuals may show a similar level of performance on two tasks even though they missed different items in the two tasks).

SA patients have been claimed to show item consistency only across multiple modalities of the same task (e.g., word and picture versions of the Camel and Cactus Test, which taps associative knowledge—e.g., matching a battery to a flashlight) but not across tasks claimed to require different types of semantic processing (e.g., Camel and Cactus test and picture naming) (e.g., Jefferies & Lambon Ralph, 2006). The same pattern has been claimed in SA patients' task consistency, where strong task correlations have been observed primarily across multiple modalities of the same task and not in other cases (Jefferies & Lambon Ralph, 2006). Researchers have interpreted the pattern of item and task consistency effects to mean that for SA patients, success depends on the semantic control demands of the task. In contrast to SA patients, SD patients have been claimed to show significant item and task consistency in nearly all cases, regardless of the task demands (Jefferies & Lambon Ralph, 2006). This pattern is argued to derive from a loss of conceptual representations, in that performance for items with disrupted representations should be impaired irrespective of access demands.

However, the proposed differences in item and task consistency between SA and SD patients is problematic in several regards. Some of the studies showing high item consistency in SD patients come from the refractory-access literature, which examined repetitions of the same items in the same task (e.g., Coughlan & Warrington, 1981; Warrington & Cipolotti, 1996; Warrington & Shallice, 1984). According to the CSC approach, SA patients, like SD patients, would be expected to show consistency with the same items on the same tasks, given that the semantic control demands would remain the same. Studies showing cross-task consistency for SD patients are few in number and have involved few patients—8–12 per study (Corbett, Jefferies, Ehsan, & Lambon Ralph, 2009; Garrard & Carroll, 2006; Jefferies & Lambon Ralph, 2006). Moreover, several studies have shown no cross-task item consistency in SD (Garrard & Carroll, 2006; Snowden et al., 2018; Snowden, Thompson, & Neary, 2004; Warrington & Shallice, 1984). Thus, the evidence for cross-task item consistency for SD patients is not particularly strong.

At the same time, SA patients have not only shown consistency across tasks with similar cognitive demands (e.g., picture–word matching and environmental sound–picture matching; Jefferies & Lambon Ralph, 2006), but also have shown consistent performance across tasks that do not have similar cognitive demands—for example, word–picture matching and object use (Corbett, Jefferies, Eshan, & Lambon Ralph, 2009a), picture naming and matching of actions to pictures (Corbett, Jefferies, Eshan, & Lambon Ralph, 2009a), word–picture matching and the word Camel and

Cactus Test (Jefferies & Lambon Ralph, 2006), and written word–picture matching with the picture Camel and Cactus Test (Robson, Sage, & Lambon Ralph, 2012). Because there is no clear relationship between the tasks' semantic control demands in these cases, it is not clear how such consistency can be accommodated. It should be noted, furthermore, that it is difficult to know whether consistency across any particular pair of tasks is predicted for SA patients, given that the means for assessing the cognitive control demands of semantic tasks have not been specified.

Even though SA patients might show consistent performance across tasks with different demands, one might claim that their level of consistency is nonetheless lower than that of SD patients. However, whereas SD patients have shown stronger consistency than SA patients in some comparisons (Corbett, Jefferies, Eshan, & Lambon Ralph, 2009a; Jefferies & Lambon Ralph, 2006), SD patients are not more consistent than SA patients in every case, but only in roughly half or less of all comparisons (Bozeat, Lambon Ralph, et al., 2000b; Corbett, Jefferies, Eshan, & Lambon Ralph, 2009a; Jefferies & Lambon Ralph, 2006; Thompson et al., 2018). These results suggest that the true difference in consistency between the groups is not overwhelmingly large, if it exists.

One factor likely to influence the degree of consistency across tasks for SA patients is concomitant language-specific deficits that affect performance in some tasks and not others. That is, patients with aphasia rarely have deficits to only a single process. So SA patients may have, for example, disruptions in perceiving or retrieving phonological or orthographic forms beyond any disruption related to semantic processing that would affect their performance on tasks that depend on spoken or written word processing (e.g., picture naming or picture–word matching). Because the degree to which language modalities are affected would be expected to differ across patients, these language-specific difficulties would reduce the size of cross-task correlations, especially with small sample sizes. For example, several patients in Jefferies and Lambon Ralph (2006) had global aphasia, which is often associated with phoneme discrimination deficits (Basso, Casati, & Vignolo, 1977; Gainotti, Miceli, Silveri, & Villa, 1982; Varney, 1984). Thus, their word–picture-matching performance was likely impaired by these speech discrimination deficits, as well as by their disrupted semantics. Also, deficits in picture naming may result from phonological deficits, in line with models of lexical access that include links between lexical and phonological representations that may be affected by brain damage (see Foygel & Dell, 2000). Schwartz, Dell, Martin, Gahl, and Sobel (2006), using a model with semantic–lexical and lexical–phonological links, found that picture-naming performance for 94 aphasic individuals was modeled well by lesioning both types of links. The importance of phonological–lexical connections was shown by the fact that 68% of patients who scored at or below 90% on the

Philadelphia Naming Test (Roach, Schwartz, Martin, Grewel, & Brecher, 1996; control average 97% in Dell et al., 1997) had greater damage to phonological than to semantic weights as a proportion of the weights for control subjects.

Executive impairments and correlations with semantic performance

One of the more compelling pieces of behavioral evidence that SA patients have semantic *control* deficits is the correlation of semantic to executive task performance for these patients. SA patients are often impaired on tasks of working memory and attention (i.e., digit span, elevator counting in the Test of Everyday Attention), pattern completion (i.e., Raven's matrices), or complex executive function (i.e., Wisconsin Card Sort, Brixton Spatial Rule Attainment) (e.g., Almaghyuli, Lambon Ralph, & Jefferies, 2012). These impairments have been taken as evidence of general executive problems that are associated with damage to the frontal and temporoparietal brain regions, areas typically damaged in these patients (Jefferies & Lambon Ralph, 2006; Lambon Ralph et al., 2017). Moreover, some researchers have found correlations between semantic and executive task performance for patients with semantic deficits and for patients recruited for showing executive deficits (Allen, Martin, & Martin, 2012; Corbett, Jefferies, Eshan, & Lambon Ralph, 2009a; Corbett, Jefferies, & Lambon Ralph, 2009b; Jefferies & Lambon Ralph, 2006; Robson et al., 2012; Thompson et al., 2018), but not for SD patients (Corbett, Jefferies, Eshan, & Lambon Ralph, 2009a; Jefferies & Lambon Ralph, 2006). These differing results for the two groups have been taken to support a causal link between semantic and executive deficits, showing that semantic cognition is "impaired as a consequence of attentional/executive deficits" (Jefferies & Lambon Ralph, 2006, p. 2144) in SA but not in SD patients. Executive deficits are thus argued to be symptomatic of deficits in semantic control.

Interpreting executive correlations Even when semantic and executive task performance correlate in SA, it is difficult to discern from the observed correlations which aspect of executive function relates to semantic access. Previously reported correlations have not focused on more clearly interpretable basic executive functions, such as inhibition, updating, and shifting (Miyake et al., 2000), but instead have focused on complex tasks that seem either to require a variety of executive functions or to tap problem solving, reasoning and intelligence (i.e., the Wisconsin Card Sort, Brixton Spatial Rule Attainment, Raven's Progressive Colored Matrices). The Wisconsin Card Sort and Raven's matrices have been associated with factors other than executive function that may also explain their correlation with semantic tasks: Both tasks correlate with language comprehension and word repetition

(Baldo et al., 2005), and the Wisconsin Card Sort with phonological short-term memory (Allen et al., 2012). Consequently, it is possible that language abilities underlie the correlations between executive function (EF) tasks and semantic task performance. Moreover, it should be noted that some case studies of patients claimed to have semantic control deficits have revealed that some of the patients did not have deficits on general executive tasks but were still claimed to have a semantic control deficit (Hoffman, Jefferies, Haffey, Littlejohns, & Lambon Ralph, 2013a).

Difficulties interpreting existing semantic–executive correlations might be alleviated by correlating semantic performance with executive tasks that are more easily interpretable. One recent study showed that performance on simpler executive tasks did not correlate with semantic performance in patients with varying degrees of semantic and phonological short-term memory deficits (Allen et al., 2012). Despite the fact that the patients in that study were not selected for showing multimodal semantic deficits, they may be relevant to the issue of semantic control, as Hoffman, Jefferies, and Lambon Ralph (2011a) claimed that patients with semantic short-term memory deficits have mild semantic control deficits.

Executive impairment in semantic dementia In contrast with SA patient performance, researchers have argued that SD patients' semantic deficits are unrelated to executive processing (e.g., Lambon Ralph et al., 2017). SD patients tested on phonological working memory and pattern completion tasks (e.g., Raven's matrices) tend to show no or relatively little impairment (Cerami et al., 2017; Gorno-Tempini et al., 2004; Jefferies & Lambon Ralph, 2006), and performance on executive tasks has not correlated with semantic performance (Corbett, Jefferies, Eshan, & Lambon Ralph, 2009a; Jefferies & Lambon Ralph, 2006). However, semantic–executive correlations are problematic as a marker of behavioral difference between SA and SD patients because, when the two groups are compared, SD patients are not typically tested on the range of executive and attention tasks given to SA patients (e.g., Wisconsin Card Sorting Test, Brixton Spatial Rule Attainment, the elevator counting task in the Test of Everyday Attention). Where they have been tested on executive/reasoning tasks in these comparisons, a restricted range of scores has sometimes limited the ability to detect correlations. For example, in Jefferies and Lambon Ralph (2006), all but one of their patients scored 100% correct on the Raven's Colored Progressive Matrices task (Raven, 1962), which was developed for testing children or adults with neurological disorders. It is impossible to conclude that SD and SA patients differ in the presence of semantic–executive behavior correlations when such correlations have not been adequately assessed for the same tasks in both patient groups.

One might claim that the limited executive testing of SD patients in studies comparing SA and SD patients might be

due to SD patients' difficulties understanding complex task instructions (Hodges & Patterson, 2007); however, there are some problems with such a proposal. First, although one might expect individuals with semantic deficits to have difficulty understanding complex task instructions, it is unclear why SD patients would have more difficulty than SA patients when they are often matched in performance on semantic tasks. Second, many studies outside of the storage–access literature have reported success in testing SD patients on various executive tasks (e.g., Bozeat, Gregory, Lambon Ralph, & Hodges, 2000a; Butts et al., 2015; Desgranges et al., 2007; George & Mathuranath, 2005; Gorno-Tempini et al., 2004; Libon et al., 2007; Macoir, Lavoie, Laforce, Brambati, & Wilson, 2017; Perry & Hodges, 2000). Furthermore, some of these studies have shown SD patients to have executive deficits (e.g., on trail-making and Stroop tasks; Butts et al., 2015; Desgranges et al., 2007; George & Mathuranath, 2005; Gorno-Tempini et al., 2004; Libon et al., 2007; Macoir et al., 2017). Such results motivate a closer examination of the relationship between executive function and semantic task performance in SD patients.

Studies documenting executive deficits in SD patients have suggested that SD patients may be more similar to SA patients on EF than is typically claimed. That is, although atrophy in SD occurs predominately in the anterior temporal lobes and spreads posteriorly as the disease progresses (Collins et al., 2017; Galton et al., 2001; Gorno-Tempini et al., 2011), frontal atrophy is found in SD in both cross-sectional (Cerami et al., 2017; Mummery et al., 2000) and longitudinal (Brambati et al., 2015; Brambati et al., 2009; Kumfor et al., 2016; Macoir et al., 2017) studies, even relatively early in disease progression (Borroni et al., 2007; Gorno-Tempini et al., 2004; Rosen et al., 2002). Fronto-temporal white matter tracts also appear to be typically affected in SD (Agosta et al., 2010; Brambati et al., 2015), and connectivity studies have shown reduced strength in executive networks (Farb et al., 2013). If frontal damage is relatively common in SD patients and executive network connectivity is affected, then one could reasonably expect semantic and executive abilities to correlate in SD patients as they do in SA patients.

Recruiting for executive deficits The CSC's interpretation of executive deficits as causing SA patients' semantic control deficits predicts that patients selected for showing executive deficits should also show symptoms of a semantic control deficit (Lambon Ralph et al., 2017). The extent to which executive and semantic deficits generally co-occur is difficult to assess in most of the published literature, given that semantic abilities are not typically highlighted in studies of executive dysfunction (e.g., Lhermitte, 1983; Stuss & Alexander, 2007). One recent study, however, did compare patients recruited for showing executive deficits to those recruited for showing semantic deficits (Thompson et al., 2018). In this study,

dysexecutive patients were claimed to have semantic control deficits because of their similarity to SA patients on a variety of behavioral measures: (1) poorer mean performance than controls on semantic measures, a relationship between executive and semantic performance that did not differ from that of SA patients; (2) consistency similar to that of SA patients and differences in consistency from a group of SD patients; (3) little or no effect of frequency or familiarity in semantic tasks; and (4) effects of semantic control manipulations similar to those among SA patients.

However, many aspects of this study's results did not support the authors' claims. First, dysexecutive patients showed consistently better performance than SA patients on semantic tasks. It is not clear why this should be the case if their semantic control abilities are truly like SA patients', as is suggested by their nearly identical effects of semantic distance, semantic ambiguity, and distractor interference. Second, item consistency was not measured outright for dysexecutive patients, but only in contrast with SA patients, leaving unclear whether dysexecutive patients were actually inconsistent or simply showed no significant difference from SA patients (a null effect) in these comparisons. Third, dysexecutive and SA patients showed no frequency effects on a synonym judgment task; however, controls likewise showed no frequency effects when tested on this task, suggesting that null frequency effects may be "normal" on this task rather than a sign of impaired performance (Jefferies, Patterson, Jones, & Lambon Ralph, 2009). Furthermore, dysexecutive patients showed differences in familiarity effects from SA patients on environmental sounds tests in the Thompson et al. (2018) study and showed no difference in familiarity effects from SD patients. Thus, there is little evidence that dysexecutive patients show atypically absent frequency or familiarity effects. Finally, dysexecutive patients showed no effect of miscues in an associative matching task. In this task, a probe word is matched with one of three words after being given a sentence that cues the dominant or subordinate meaning of the probe word (e.g., "The labourers cleaned out the pen," which leads one to match *PEN* with *PIG* rather than *PENCIL*). SA patients have previously shown difficulty overcoming the competition elicited by miscues in this task, consistent with a semantic control deficit per the CSC framework (Noonan et al., 2010). The fact that dysexecutive patients did not show this semantic control symptom further suggests that their executive function deficit did not impair semantic control in the fashion expected from the CSC.

Overall, the dysexecutive patients from Thompson et al. (2018) did not appear compellingly similar to SA patients across behavioral semantic control phenomena, contrary to predictions of the semantic control literature. Thus, some evidence suggests that patients recruited for showing executive deficits may not be as similar to SA patients in behavior as the CSC predicts.

Effects of distractor strength

One behavioral phenomenon thought to directly reflect semantic control is the ability to ignore irrelevant, distracting information. Noonan et al. (2010) tested this ability in SA patients in a synonym judgment task. Participants were asked to select the closest synonym of a probe word from among three words: a correct target, a distractor word, and an unrelated word. The distractor could be either more closely associated with the probe than the target (e.g., distractor *CAKE* for probe *PIECE* with target *SLICE*) or less closely associated to the probe than the target (e.g., distractor *HOUSE* for probe *LOFT* with target *ATTIC*). In this task, SA patients showed difficulty ignoring strong distractors, performing much worse with strong than weak distractors and showing a much stronger distractor strength effect than controls. Hoffman, Jefferies, and Lambon Ralph (2011a) found the same result in patients with semantic short-term memory deficits, which they argued derive from mild semantic control deficits. These results suggest that the size of distractor strength effects should be a robust reflection of semantic control ability.

Given that Hoffman, Jefferies, and Lambon Ralph (2011a) argued that semantic STM deficits derive from semantic control deficits, they also manipulated short-term memory demands in their distractor strength experiment by presenting material either visually or auditorily, with auditory presentation argued to place heavier burden on STM. They found that the SA patients performed more poorly with the auditory versions. Because both STM abilities and the ability to resist distractors were claimed to depend on the same control mechanism, one might have expected these two factors to interact, given the additive-factors logic (Sternberg, 1998); however, no statistical evidence was presented that supported the presence of an interaction.

Although the Noonan et al. (2010) findings on distractor strength effects are consistent with a semantic control deficit, it is important to note that SD patients' performance on such a task has not been reported. Several other task manipulations have also been claimed to directly gauge SA patients' semantic control abilities, such as the manipulations of semantic distance, semantic ambiguity, and distractor strength discussed above, but only one of these has been tested in SD patients—specifically, the effect of semantic distance. Thompson et al. (2018) showed that when given an associative matching task in which a probe (e.g., *HAT*) must be matched with one of three words, both SA and SD patients perform better when the target was more semantically “close” (e.g., *CAP*) than when the target is more semantically “distant” (e.g., *STOCKING*). Although this effect was stronger in SA than in SD patients, it was nevertheless present in SD patients to a stronger degree than in controls, bringing into question whether such a manipulation can truly distinguish damaged representations and damaged semantic control.

Because SD patients showed effects of the semantic distance manipulation, we might predict that other effects of semantic control might be present, or even typical, in SD patients. Demonstrating this would provide evidence either that SD patients show signs of semantic control deficits or that effects of these manipulations may derive from deficits other than those of semantic control.

The present study

The issues above motivated the present study of SD and SA patients. As we discussed, some results from the literature show dissociations in behavioral patterns that are argued to distinguish semantic knowledge and semantic control deficits, and other results do not corroborate these dissociations. True differences in consistency between the groups were unclear from the literature and were thus further investigated. Several studies have claimed dissociations between SD and SA patients despite individuals from the two groups not having been directly compared on the same executive or semantic control tasks (Corbett, Jefferies, Eshan, & Lambon Ralph, 2009a; Jefferies & Lambon Ralph, 2006; Thompson et al., 2018). Also, with regard to executive function, typically the authors comparing SA and SD patients have used only complex tasks for which it is difficult to determine the relevant factor underlying correlations with other measures. Following these concerns, in the present study we tested patients from the different groups on the same tasks and used a variety of EF tasks that tap simpler processes. Our goals were to (1) provide evidence regarding whether patterns of item and task consistency, executive performance, semantic–executive correlations, and semantic control manipulations distinguish SD and SA patient performance; (2) describe the executive capacities underlying semantic control more specifically; and (3) investigate further whether stroke patients recruited for having executive deficits show semantic control deficits. To these ends, we tested three patient groups and a control group on a set of semantic and executive tasks. The patient groups consisted of (1) SA patients, (2) stroke patients with deficits in phonological working memory (signifying the potential presence of an executive deficit), and (3) SD patients. The tasks on which patients were tested included tasks used in previous studies of SA and SD patients—complex executive tasks and semantic tasks that tap multiple input modalities—as well as tasks tapping simple executive functions—updating working memory, shifting between task goals, and inhibition of distractor interference (Miyake et al., 2000)—and a task manipulating distractor strength, to more directly assess semantic control abilities.

Participants

In this study we tested 16 stroke patients, six SD patients, and ten healthy, age- and education-matched controls. The stroke patients were recruited to our study from local Houston-area stroke rehabilitation centers and in cooperation with the Methodist Hospital, Memorial Hermann, and the University of Texas McGovern Medical School in Houston. The stroke patients were initially selected to meet one of two criteria: (1) having a multimodal semantic deficit (i.e., SA), as indicated by poor performance on both word and picture versions of the Camel and Cactus Test (Bozeat et al., 2000b; $n = 9$), or (2) having a phonological working memory deficit (i.e., non-SA), as determined by having a poor word span or digit span but not having a multimodal semantic deficit ($n = 7$; see Table 9 below). Patient demographic information is displayed in Table 1.

The stroke patients were native English-speakers who represented a range of overall severity, as reflected in their aphasia quotient scores on the Western Aphasia Battery (WAB; Kertesz, 2007; see Appendix Table 11 for scores on the WAB and other background measures). Stroke patients recruited for phonological working memory deficits showed no evidence of apraxia or dysarthria. All scored at the high end of the WAB fluency measures. Most patients did well on a

reading test, as assessed by PALPA 25, a visual lexical-decision task (Kay, Lesser, & Coltheart, 1992), or by a similar task reported in R. C. Martin, Breedin, and Damian (1999); each of these tasks included 60 words that varied in frequency and imageability, as well as 60 pronounceable nonwords. We scored subjects' accuracy in deciding whether each stimulus was a word or a nonword. As is shown in the Appendix Table 11, most scored at or above 90% correct on these tests. Of the three patients who scored below 90% correct on PALPA 25, two (P.B. & N.L.) did very well on the 30 high-imageability words (97% and 100% correct, respectively), and the other one (S.E.) scored 87% correct on these items. These items are the most relevant, because all of the written words in the semantic tasks were of high imageability. One patient, C.A., had particular reading difficulties, scoring at only 79% on the R. C. Martin et al. task (with similar performance for high- and low-imageability items), and his scores were left out of the group mean for the button Stroop task. C.A.'s score on the word CCT was retained in the group mean, as his performance was nearly identical across the picture and word versions of this task, and thus appears not to have been affected by his reading abilities.

The stroke patients had reasonably good low-level speech perception, as measured by a test of consonant discrimination (R. C. Martin & Breedin, 1992), and were also tested for basic sensory input problems, including hearing and vision problems. We compared patients' hearing to international standards for median hearing thresholds as a function of age (International Organization for Standardization, 2000), computing for each patient the standard score of their hearing threshold relative to the median and standard deviation from these age-adjusted norms. This was done at 500, 1000, 2000, 3000, and 4000 Hz for the left and right ears, averaged for each ear, and then averaged across ears to provide a single estimate of hearing loss (after Dial & Martin, 2017). Positive values indicate hearing thresholds greater than the normative value and negative numbers indicate thresholds less than the normative value. Stroke patient hearing thresholds ranged from -1.67 to 2.65 (see Appendix Table 11 for individual patient values). Three patients (K.A., H.O., and K.V.) had thresholds greater than 2.0 standard deviations above the normative value. K.A., but not the other two, demonstrated dramatic impairments on speech perception tasks, as did N.L., who did not have hearing loss. Therefore, K.A. and N.L.'s performance on tasks that required recognition of auditory input was not considered in comparison with the other patients in this study.¹ H.O. had previously been shown to score within the control range on a number of stringent tests of speech perception (Dial, McMurray, & Martin, 2019), and thus her hearing loss did not seem to have consequences relevant for

Table 1 Patient demographic information

Group	Patient	Age	Education	Years Post-Stroke
SA	C.A.	58	18	1
	D.Z.	45	20	9
	E.J.	77	18	1
	E.V.	57	16	18
	H.O.	55	16	1
	K.A.	57	—	13
	K.V.	56	16	14
	P.B.	45	16	6
	T.Q.	54	—	2
Non-SA	M.B.	67	13	13
	N.J.	46	14.5	5
	N.L.A.	55	14	4
	N.L.	71	17.5	8
	Q.O.	69	20	4
	S.E.	45	18	1
	S.J.	68	13	9
SD	E.G.	66	20	—
	E.K.	63	12	—
	K.E.	57	13	—
	I.R.	72	12	—
	T.A.	81	13	—
	T.F.	51	16	—

¹ These tasks were attempted with patient K.A., but, predictably, he performed very poorly.

picture–word matching. K.V. performed mildly worse than controls on a speech perception task, but when he was tested with matching a spoken word to a picture with a phonologically related name, he correctly rejected this phonological distractor 100% of the time. Thus, K.V.’s hearing impairment is unlikely to have impacted his performance on tasks involving words, but it may have impaired performance with non-word stimuli in the environmental sounds test.

Patient D.Z. had extremely limited output, and her performance on tasks that required verbal output was not considered in comparison with the other patients in this study. Stroke patient S.E. was a Norwegian–English bilingual patient, but she had emigrated to the United States when she was five years old and has since predominately spoken English, so her bilingualism is unlikely to have strongly affected her performance. Given the evidence favoring a shared conceptual store across languages in bilinguals (see Francis, 2018), a semantic deficit should have been detectable in S.E.

Six SD patients were tested. Four of these patients were referred to our study by a neurologist on the basis of semantic deficiencies consistent with a diagnosis of semantic dementia. Two SD patients were referred for remote testing after previous participation in a study with M. Henry, who was then at the University of California, San Francisco. These patients all fulfilled the typical published criteria for SD: They had fluent speech with word-finding difficulties, impaired semantic knowledge and single-word comprehension, and intact phonology, syntax, and visual–spatial abilities (Hodges, Patterson, Oxbury, & Furnell, 1992). Dementia severity, shown in Appendix Table 11, was given by the party that referred the patients, with the qualification that characterizing severity for these patients is not straightforward. One SD patient, E.G., was recruited only for the semantic tasks and the synonym judgment portions of the study, and was not tested on any executive measures. Two SD patients, K.E. and E.K., were unable to understand or perform many of the executive tasks. After observing that K.E. and E.K. had extreme difficulty understanding how to perform simple executive tasks such as the one-back and trail-making tasks, we did not proceed with testing them to more difficult tasks, such as cued task-set shifting and the Wisconsin Card Sort Task. All testing for the SD patients took place within a span of less than 4 months, to reduce the impact of disease progression on behavior.

Brain damage was assessed in 12 of 16 stroke patients and 4 of 6 SD patients using volume data from *Freesurfer*’s *recon-all* command. Regions with more than 10% of voxels damaged were considered to have damage. Brain damage for the remaining 2 of 6 SD patients was assessed by comparison to age-matched healthy controls using voxel-based morphometry, controlling for age, sex, and total intracranial volume. Imaging was not available for the remaining stroke patients, due to contraindications for MRI scanning. Information on patient brain damage is presented in Table 2.

Stroke patients showed varying locations of damage, with some showing frontal or parietal damage that did not extend into temporal areas, and some showing only temporal damage. One stroke patient, H.O., showed damage restricted to the cerebellum. All but one SD patient (T.A.) showed broad temporal damage, predominantly in left anterior inferior temporal cortex. T.A. showed predominant right anterior inferior temporal damage in the fusiform gyrus, with some damage to the left superior temporal lobe. All SD patients also showed some degree of frontal and/or parietal atrophy in addition to anterior temporal atrophy. Such damage has not been described in previous comparisons of SA and SD patients. However, given the brief and cursory descriptions in most accounts, it is unclear whether the SD patients in previous studies did not have such damage or whether the authors simply noted the presence of temporal damage to confirm the behavioral diagnosis of SD. As we discussed earlier, studies of SD patients outside of those comparing SA and SD patients has documented such nontemporal damage (e.g., Agosta et al., 2010; Borroni et al., 2007; Brambati et al., 2015; Brambati et al., 2009; Cerami et al., 2017; Gorno-Tempini et al., 2004; Kumfor et al., 2016; Macoir et al., 2017). Interestingly, at least one study has shown that stimulation of the orbitofrontal cortex, an area damaged in 4 of 6 SD patients, may cause semantic paraphasias (Duffau et al., 2005). This finding supports the relationship between frontal damage in SD patients and their semantic deficits.

Power

Because of the inherent limitations of locating large numbers of stroke and dementia patients with multimodal semantic deficits and a lack of advance knowledge of the variability expected from such individuals on some of the tasks, it was difficult to estimate power in advance. However, the present study’s case series design, with multiple tests of semantic and executive abilities for each individual, provided confirmation of the consistency of the observed patterns.

Semantic assessments

Method

Semantic production and comprehension were assessed with five tasks in the Cambridge Battery (e.g., Jefferies & Lambon Ralph, 2006; <http://www.ftdrg.org/cambridge-semantic-battery/>), the first four of which contain the same set of 64 items, allowing for comparison of item-by-item performance across tasks: (1) *Picture naming*: Participants are asked to name line drawings, presented one at a time. (2) *Spoken word–picture matching*: Participants match a word presented over speakers to one of ten black and white line drawings, with distractors

Table 2 Patient brain lesion information

	Left Hemisphere					Right Hemisphere				
	Frontal	Temporal	Parietal	Occipital	Subcortical	Frontal	Temporal	Parietal	Occipital	Cerebellum
<i>SA</i>										
C.A.	–	–	–	–	–	–	–	–	–	–
D.Z.	inferior, middle, precentral gyrus	insula, middle, superior, Heschl's gyrus, pole	inferior, superior, supramarginal, angular gyri, postcentral gyrus	middle, superior	–	–	–	–	–	–
E.J.	–	insula, middle, superior, Heschl's gyrus	–	–	–	–	–	–	–	–
E.V.	middle, inferior	insula	–	–	–	–	–	–	–	–
H.O.	–	–	–	–	–	–	–	–	–	–
K.A.	inferior	insula, middle, superior, pole, Heschl's gyrus	–	–	–	–	insula, inferior, middle, superior, pole, Heschl's gyrus	supramarginal, angular gyrus	middle, superior	areas 4, 5, 6
K.V.	–	–	–	–	–	–	–	–	–	–
P.B.	inferior, middle, superior	insula, middle, superior, Heschl's gyrus, pole	inferior, superior, supramarginal, angular gyrus	–	–	–	–	–	–	–
T.Q.	–	–	–	–	–	–	–	–	–	–
<i>Non-SA</i>										
M.B.	–	–	inferior, superior, supramarginal, angular gyrus; post-central gyrus	middle, superior	–	–	–	–	–	–
N.J.	–	insula, middle, superior, Heschl's gyrus, pole	inferior, supramarginal, angular gyrus	–	–	–	–	–	–	–
N.L.A.	middle, inferior	insula, inferior, middle, superior, Heschl's gyrus, pole	inferior, angular gyrus	inferior, middle	–	–	–	–	–	–
N.L.	–	–	–	–	–	–	–	–	–	–
Q.O.	–	middle, superior	supramarginal	–	–	–	–	–	–	–
S.E.	inferior, precentral gyrus	insula, superior, Heschl's gyrus	inferior, superior, supramarginal, precuneus	–	putamen	–	–	–	–	–
S.J.	–	middle, superior, Heschl's gyrus	inferior, supramarginal, angular gyrus	–	–	–	–	–	–	–
E.G.	–	broad atrophy, greatest in ATL	mild superior parietal atrophy	–	–	–	broad atrophy, greatest in ATL	–	–	–
E.K.	inferior, middle, orbitofrontal, cingulate	insula, broad atrophy, fusiform	inferior, supramarginal, angular gyrus, precuneus	–	–	inferior, middle,	insula, broad atrophy, temporal atrophy, fusiform	inferior, supramarginal, lateral	–	–

Table 2 (continued)

Left Hemisphere					Right Hemisphere				
Frontal	Temporal	Parietal	Occipital	Subcortical	Frontal	Temporal	Parietal	Occipital	Cerebellum
K.E. inferior, middle, orbitofrontal, cingulate	insula, broad atrophy, fusiform, parahippocampal	inferior, supramarginal, precuneus	lingual gyrus	entorhinal	orbitofrontal, I, cingulate inferior, middle, orbitofrontal, I, cingulate	broad temporal atrophy, fusiform	angular gyrus, precuneus inferior, supramarginal, precuneus	—	—
I.R. —	broad atrophy, greatest in ATL, mild perisylvian atrophy superior	moderate perisylvian atrophy, mild superior parietal atrophy	—	—	—	broad temporal atrophy, greatest in ATL	—	—	—
T.A. inferior, medial orbitofrontal	—	—	—	—	orbitofrontal, frontal pole, cingulate	fusiform	—	—	—
T.F. inferior, middle, orbitofrontal, cingulate, paracentral lobule	insula, inferior, fusiform	posterior cingulate	—	entorhinal	inferior, middle, orbitofrontal, I, cingulate	insula, inferior, middle	supramarginal	lateral	—

drawn from the same category as the target. (3) *Camel and Cactus Test picture version (CCTp)*: Participants decide which of four colored drawings or pictures is most closely associated with the item presented at the top of the screen (e.g., is a camel more associated with a cactus, tree, sunflower, or rose?). (4) *Camel and Cactus Test word version (CCTw)*: This is the same as the CCTp, but the items are words rather than pictures. (5) *Sound–picture environmental sounds test*: Participants matched a given sound with one of ten pictures, with distractors drawn from the same category as the target.

In word–picture matching, participants could request to hear the audio file repeated if needed. Several items in the CCT tests were replaced to adapt the test for American participants, and several sounds from the environmental sounds task were replaced because of their high rate of error in controls. Replacement sounds were chosen by ensuring that two new control subjects identified them correctly. On the CCT, several items were missed consistently by controls. For these items, alternative answers were accepted (items and answers listed in Appendix Table 12).

We compared individual patient scores on each semantic task to control means using the modified *t* test of Crawford and Howell (1998). All statistical tests were two-tailed.

Results

The nine stroke patients selected for having a semantic deficit on the CCT word and picture tests (SA) and all six dementia patients showed impaired semantic performance ($p < .05$) across all tasks, except on the environmental sounds tests, in which the wide control range made it difficult to find significant differences (see Table 3). Patients H.O. and S.E. were impaired on one version of the CCT, but performed only marginally worse than controls ($p < .10$) on the other. Because H.O. showed significant impairment in both verbal and non-verbal tasks, but S.E. showed significant impairment only on tasks with a verbal component, we considered H.O. but not S.E. to have a multimodal semantic deficit.² SA patient S.E. showed no difficulty with word–picture matching. The seven stroke patients selected for having either word span or digit span deficits (non-SA) but preserved performance on one or both CCT tasks showed few or no semantic problems. Patient M.B. showed problems with picture naming because of a phonological output problem, which was evident by his large proportion of phonological errors in picture naming and spontaneous speech. Patient N.J. was impaired only on tasks that required use of lexical representations, as shown by his normal level of performance on the CCT picture and environmental sounds tasks.

² Group tests involving the SA and non-SA groups were also performed with S.E. included in the SA group. No substantial differences were found between this analysis and our reported analysis.

Table 3 Semantic task performance (number correct)

Participant		Picture Naming	Word–Picture Matching	CCT (pictures)	CCT (words)	Env. Sounds
SA	D.Z.	—	39*	26*	36*	20*
	P.B.	25*	52*	42*	34*	33
	E.J.	21*	53*	44*	45*	24*
	C.A.	38*	60*	34*	33*	38
	K.A.	57*	—	53*	45*	—
	H.O.	48*	37*	41*	55^	37
	K.V.	41*	59*	51*	39*	40
	T.Q.	54*	54*	51*	36*	37
	E.V.	57*	60*	52*	53*	42
Non-SA	N.J.	55*	58*	57	48*	40
	S.E.	52*	63	54^	48*	44
	Q.O.	62	64	56	59	45
	M.B.	58*	63	61	61	46
	S.J.	64	62	60	58	42
	N.L.A.	63	64	59	59	39
	N.L.	54*	—	57	—	—
SD	K.E.	36*	56*	25*	21*	19*
	E.K.	34*	53*	29*	26*	8*
	E.G.	37*	47*	31*	35*	25*
	T.F.	42*	51*	23*	35*	26^
	T.A.	54*	60*	29*	36*	20*
	I.R.	53*	55*	49*	43*	30
Control Mean (SD)		62.6 (1.2)	63 (1.2)	57.7 (2.0)	57.1 (1.1)	38.7 (5.4)
Control Range		61–64	61–64	55–61	55–59	30–44

*Impaired performance as compared to collected control data using Crawford and Howell's adjusted t test, $p < .05$. ^ $p < .10$. CCT = Camel and Cactus Test. Patients are ordered within their groups based on severity of semantic deficit as measured by their semantic composite score, described in Semantic–Executive Correlations section.

Significant differences in mean performance on these tasks were found in several group comparisons (see Table 4). SA and SD patients scored more poorly than controls on every task (all t s > 3.9 , $ps \leq .001$), except for SA patients on the environmental sounds test, in which controls showed particularly poor performance, $t(16) = 1.55$, $p = .14$. Non-SA patients performed more poorly than controls only at picture naming, $t(15) = 2.77$, $p = .01$. The SD group performed more poorly than the SA group on the picture CCT and the environmental sounds task, t s > 2.6 , $ps < .05$, and marginally more poorly on the word CCT, $t(13) = 2.16$, $p = .053$, but the groups performed similarly on picture naming and word–picture matching (t s < 1). The differences between the SD and SA groups on the CCT and the environmental sounds tasks are likely due to these tasks' difficulty. In the CCT tasks, one must discern the appropriate relation for matching the two pictures, and on the environmental sounds tests, the lack of familiarity with some of the sounds may have caused patients to eliminate possibilities through a reasoning process. Predictably, given the selection criteria, the non-SA group performed better than

both SA and SD groups on every task, t s > 2.7 , $ps < .05$. Our SD patients showed impairments comparable to those in a previous research on SA and SD (Jefferies & Lambon Ralph, 2006) on every task except picture naming, in which our patients scored higher, $t(13.7) = 2.33$, $p = .04$ (other $ps > .18$). Likewise, our SA patients were comparable to previously tested SA patients (Jefferies & Lambon Ralph, 2006), though they scored significantly higher than previous patients at picture naming, $t(15.8) = 2.73$, $p = .02$, and marginally higher at environmental sound–picture matching, $t(9.5) = 1.85$, $p = .096$ (other $ps > .12$).

Item and task consistency

Method

Item consistency for the 64-item tasks was analyzed using logistic regression, and Cohen's kappa (K) by comparing each task pair across groups and for each participant. In the logistic

Table 4 Group means across semantic tasks

	Picture Naming	Word–Picture Matching	CCT Picture	CCT Word	Environmental Sounds
SA	42.6 (14)	51.8 (9.1)	43.8 (9.2)	41.8 (8.2)	33.9 (7.9)
Non-SA	58.3 (4.8)	62.3 (2.3)	57.7 (2.4)	55.5 (5.9)	42.7 (2.8)
SD	42.7 (8.8)	53.5 (4.8)	31.0 (9.3)	32.7 (7.9)	21.3 (7.7)
Control	62.6 (1.2)	63 (1.2)	57.7 (1.9)	57.1 (1.1)	38.7 (5.4)
Control Range	61–64	61–64	55–61	55–59	30–44
Control vs. SA	***	**	***	***	—
Control vs. Non-SA	*	—	—	—	—
Control vs. SD	**	**	***	***	**
Non-SA vs. SA	*	*	**	**	*
Non-SA vs. SD	**	**	***	***	***
SA vs. SD	—	—	*	^	*

*** $p < .001$, ** $p < .01$, * $p < .05$, ^ $p < .10$

regression analysis, we included a normed measure of familiarity (Scott, Keitel, Becirspahic, Yao, & Sereno, 2019) as a control variable. Jefferies and Lambon Ralph (2006) used their own derived familiarity values in consistency regressions. For each pairwise comparison, a regression was run with each task in a pair as a predictor. We considered the tests consistent only if both regressions showed significance. In regressions comparing groups, additional predictors included group, the Group $\times$ Task interaction, and the Group $\times$ Familiarity interaction. We analyzed Cohen's K to measure consistency while controlling for overall levels of performance. That is, patients might be consistent simply because they did very poorly or very well overall, and Cohen's K adjusts for this. Values of K range from 0 to 1, representing fully inconsistent and fully consistent behavior, respectively. For this analysis, we treated tasks as raters of items for each individual patient.

As an additional analysis of group-wise item consistency we ran logistic mixed effects models for each task comparison. That is, for each task pair in each patient group, we predicted one task's accuracy from the other's, including group and familiarity³ as fixed effects and predicting random intercepts for subjects and items. These models were run with each task as a predictor. We considered effects significant if $z > 2$ in both models or marginal if $z > 1.65$ in both models. Because our group consistency comparisons involved few patients, we acknowledge that any differences observed or not observed must be taken as only preliminary evidence. However, we note that group consistency comparisons of SA and SD in past

studies have relied on comparatively small numbers of patients (i.e., 10–12) and should be treated with similar caution (Corbett, Jefferies, Eshan, & Lambon Ralph, 2009a; Jefferies & Lambon Ralph, 2006).

Besides task-pairwise measures of consistency, we also measured item consistency across all four 64-item tasks using Fleiss's generalized K (Fleiss, 1971) in each patient and group, again treating tasks as raters of items for each individual patient. The logic behind using Fleiss's K was the same as that for Cohen's K . Fleiss's K , however, allowed us to estimate consistency across all tasks rather than only one pair at a time. As with Cohen's K , values of Fleiss's K closer to 1 indicate stronger consistency than would be expected by chance, and values closer to 0 indicate consistency closer to chance.

Results

Item consistency

For each of the four semantic tasks in the consistency analyses, several patients were excluded for having near ceiling performance (i.e., scoring within the control range), which would make it difficult to detect inconsistencies. Given that the non-SA patients were selected for not having semantic deficits, many of them fell into the excluded group (see Appendix Table 13 for a list of those excluded on each task).

As we discussed above, some prior studies have claimed that SD patients show a strong degree of consistency across all semantic tasks with the same items, whereas SA patients show consistency only between tasks with the same cognitive demands (Corbett, Jefferies, Eshan, & Lambon Ralph, 2009a; Jefferies & Lambon Ralph, 2006; Thompson et al., 2018; though see the null effects of consistency in Garrard & Carroll, 2006; Snowden et al., 2018; Snowden, Thompson, & Neary, 2004; Warrington & Shallice, 1984). When analyzed across all 64-item tasks at the group level, a significant level of

³ In mixed models and regressions, we ran separate models including log SUBTLEX frequency (Brysbaert & New, 2009) as a control variable rather than familiarity. This was done because frequency effects are claimed to differ between SA and SD patients, and therefore word frequency seems to be more relevant to processing here than is familiarity, notwithstanding the typically high correlation between frequency and familiarity (Kuperman, Stadthagen-Gonzalez, & Brysbaert, 2012). These analyses showed essentially the same pattern of results as those based on familiarity.

Table 5 Group pairwise consistency performance

	SA			SD			Non-SA		
	Wald	Fleiss's <i>K</i>	Mixed Model <i>z</i>	Wald	Fleiss's <i>K</i>	Mixed Model <i>z</i>	Wald	Fleiss's <i>K</i>	Mixed Model <i>z</i>
PN–WPM	< 1	.03	< 1.1	> 8.8**	.18***	> 3.1*	–	–	–
PN–CCTp	< 1.8	.08^	< 1	< 2.5	.10*	< 1.4	–	–	–
PN–CCTw	> 5.6*	.11*	> 1.66^	< 1.9	.06	< 1.1	< 1.2	.14	< 1
WPM–CCTp	< 1	.01	< 1	< 1.5	.06^	> 1.8^	–	–	–
WPM–CCTw	< 1	.00	< 1	< 1.7	.05	< 1.3	–	–	–
CCTp–CCTw	19.48***	.21***	> 3.6*	11.0**	.16**	> 2.6*	–	–	–

PN = picture naming, WPM = word–picture matching, CCTp = Camel and Cactus pictures, CCTw = Camel and Cactus words, *** $p < .001$ ** $p < .01$ * $p < .05$. ^ $p < .10$. Wald and Mixed model coefficients are from models in which familiarity was controlled.

consistency was observed in SA patients (Fleiss's $K = .04$, $z = 2.01$, $p = .04$) and in SD patients (Fleiss's $K = .07$, $z = 3.42$, $p = .001$). Consistency was numerically higher in SD patients than SA patients, but both groups showed low overall consistency. These overall findings are reflected in the low consistency observed in many pairwise comparisons of tasks (see Table 5). In logistic regression, Fleiss's K , and mixed model analyses, all patient groups showed strong consistency across the two CCTs, but only one or two other task pairs showed significant consistency. Thus, expectations were violated for SD patients, who did not show consistency across all tasks, and for SA patients, in whom consistency was observed across a pair of tasks with apparently different cognitive demands (i.e., picture naming – CCT word).

Regressions comparing group consistency that included familiarity as a predictor showed no evidence of group differences between SA and SD patients. Likewise, no group differences in familiarity effects were observed between SA and SD patients. These absent group differences are unsurprising, given the low power allowed by the small patient group sizes.

Individual patient analyses converged with group analyses in showing little difference between patient groups. Fleiss's K

indicated that one patient in each group showed significant consistency across all tasks: SA (patient H.O., $K = .20$, $z = 3.95$, $p < .001$), SD (patient I.R., $K = .17$, $z = 3.25$, $p = .001$), and non-SA (patient S.E., $K = .19$, $z = 2.66$, $p = .008$). One other value was marginal in the SD group (patient E.G., $K = .10$, $z = 1.93$, $p = .054$) and the non-SA group (patient Q.O., $K = .13$, $z = 1.83$, $p = .067$). Logistic regression and Cohen's K showed several patients in each group with significant or marginal consistency for the CCT pair and only one or two other patients for other comparisons with significant or marginal values (see Table 6).

Together, item consistency analyses suggest little item-driven consistency for any patient group and only rare and minor item consistency in individual patients. These results further suggest, contrary to some claims in the published literature (Garrard & Carroll, 2006; Jefferies & Lambon Ralph, 2006) though consistent with other findings (Corbett, Jefferies, Eshan, & Lambon Ralph, 2009a; Jefferies & Lambon Ralph, 2006), that SD patients are not always more consistent than SA patients across different semantic tasks with the same items and that SA patients may be consistent across tasks with little apparent overlap in task demands. Our

Table 6 Numbers of patients with significantly ($p < .05$) or marginally ($p < .10$) consistent performance by task pair in logistic regression and Cohen's K analyses

	SA		SD		Non-SA	
	Regression	Cohen <i>K</i>	Regression	Cohen <i>K</i>	Regression	Cohen <i>K</i>
PN–WPM	1/8	2/8	2/6	2/6	0/1	0/1
PN–CCTp	0/8	0/8	0/6	0/6	0/1	0/1
PN–CCTw	0/7	0/7	1/6	2/6	0/2	1/2
WPM–CCTp	1/9	1/9	1/6	1/6	0/1	0/1
WPM–CCTw	0/8	1/8	0/6	1/6	0/1	0/1
CCTp–CCTw	4/8	4/8	3/7	3/6	1/1	1/1

PN = picture naming, WPM = word–picture matching, CCTp = Camel and Cactus pictures, CCTw = Camel and Cactus words. Regression tallies come from models that controlled for word familiarity.

results are also in line with studies that have found a lack of cross-task item consistency in SD patients (Snowden et al., 2018; Warrington & Shallice, 1984).

Correlations between semantic tasks

As was the case for item consistency, some previous literature has suggested that SD patient performance should correlate between all semantic tasks, whereas SA patients should show correlations only where tasks have the same cognitive demands (Corbett, Jefferies, Eshan, & Lambon Ralph, 2009a; Jefferies & Lambon Ralph, 2006). In the present data, for both SA and SD groups, many semantic tasks showed moderate to strong correlations (see the data in Table 3 and the correlations in Table 7). Because of small sample sizes, few of these correlations were significant, but even so, many of the nonsignificant correlations were fairly strong ($r_s = .37-.70$). In the SA group, moderately high correlations were observed for tasks with ostensibly similar cognitive control demands (i.e., CCT picture – CCT word; word–picture matching – environmental sounds), but also for tasks with little apparent similarity in demands (i.e., picture naming with CCT picture and CCT word; see Fig. 1). For the SD group, two correlations were of marginal significance, an additional five ranged from .45 to .67, and the remaining correlations were close to zero.

To more easily compare the correlations in each group, we used Fisher's z transformation, computed the mean transformed scores for each group, and transformed these means back into correlation coefficients. Both mean correlations were moderately large, though SD patients showed a slightly higher mean correlation, of $r = .46$, than did SA patients, $r = .43$. As another means of comparing the degree of correlation among the tasks overall, we computed the percentage of variance accounted for by the first principal component based on the intertask correlations for each group. These were very similar (SA patients, 63%; SD patients, 58%). Thus, these

findings demonstrated little difference between the groups in the degree of cross-task consistency.

Executive function assessments

Method

We employed ten simple tests of executive function—including tests of updating, shifting, and inhibition—and three complex executive or reasoning tasks. Tasks tapping updating ability included letter one-back and two-back working memory tasks (Hull, Martin, Beier, Lane, & Hamilton, 2008), word span with one-syllable word lists (R. C. Martin, Lesch, & Bartha, 1999), and digit span (Wechsler, 1981) tasks. In the one-back task, participants saw a string of letters appear one at a time on the screen and were instructed to press a button whenever the letter on screen was the same as the letter on the immediately preceding trial. In the two-back version, participants were to press the button whenever the letter on screen was the same as the letter two trials before. In the word span task, participants were read short lists of words and were asked to repeat them back in order. List length began at two and increased until participants scored 50% or less on a given list length. Digit span was administered similarly to word span, except that prerecorded auditory digit stimuli were presented and responses were typed back. In both span tasks, spans were calculated as the estimated list length at which 50% of a list would be repeated back correctly, using linear interpolation between the two lengths spanning 50%. In the digit span task, SA patient T.Q. and SD patient E.K. repeated their answers aloud, because neither could reliably match the spoken digits to the written digits on the keyboard keys.

Tasks tapping shifting ability included a modified trail-making test (Zhao et al., 2013) and a cued task-set shifting task (Monsell, 2003). In part A of the modified trail-making

Table 7 Semantic task correlations for SA and SD patients

	Pic Naming	Word–Pic Matching	CCT Pic	CCT Word
<i>SA</i>				
Word–pic matching	.01	–	–	–
CCT pic	.57	.56	–	–
CCT word	.39	–.24	.38	–
Env. sounds	.81*	.54	.66^	.25
<i>SD</i>				
Word–pic matching	.58	–	–	–
CCT pic	.55	.09	–	–
CCT word	.77^	–.06	.67	–
Env. sounds	.54	–.19	.45	.73^

* $p < .05$. ^ $p < .10$.

Fig. 1 Scatterplot of stroke and semantic dementia (SD) patient performance across two pairs of semantic tasks. (a) Correlations between tasks with similar demands: Camel and Cactus picture and

word versions. (b) Correlations between tasks with different demands: Camel and Cactus word version and picture naming

task, participants connected randomly scattered numbers around the page in order, from 1 to 25. Numbers were within either a circle or a square. In part B, participants connected numbers in order, but they also had to alternate between shapes as they counted (e.g., move from square 2 to circle 3, and NOT square 3). In the cued task-set-shifting task, participants pressed a button corresponding to the color or shape of an item, as directed by the cue printed above them. The item had two potential colors and shapes. In the critical block, the cue switched every two trials. The dependent measure was the local switch cost, which is the difference in log RTs between correct trials in which the response dimension switched and in which it repeated. Any correct trial following an incorrect trial was discarded from the calculation of the switch cost, as an incorrect response may have affected the following RT.

Tasks tapping inhibition of distractor interference or prepotent response included a standard verbal Stroop (adapted from Miyake et al., 2000), a Stroop test with button press response, horizontal Stroop (e.g., Hamilton & Martin, 2005), and a picture–word interference test (e.g., Allen et al., 2012). In the verbal Stroop, participants spoke the name of the color of a string of characters presented onscreen while ignoring the meaning of the written word. Six colors were used. Character strings could be neutral (i.e., a string of asterisks), congruent, or incongruent relative to the color in which they were displayed. The task included 77 neutral trials, 12 congruent trials, and 65 incongruent trials. The button press Stroop was similar to the verbal Stroop, but color responses were made via button press, only four colors were used, and the number of trials was larger: 192 neutral, 48 congruent, and 144 incongruent. In the horizontal Stroop, participants responded to the direction of a left- or right-pointing arrow, which could appear in the left, center, or right side of a

horizontal rectangle. In congruent trials, the direction of the arrow was the same as its position in the rectangle; in incongruent trials, these conflicted; in neutral trials, the arrow was in the center of the rectangle. For all Stroop tasks, the dependent measure was the Stroop effect, calculated as the difference in log RTs between correct incongruent and neutral trials.

In the picture–word interference task, participants were asked to name line drawings with words superimposed on top of them. Words were presented simultaneously with the image and could be either semantically related or unrelated to the image. The dependent measure was the interference effect, calculated as the log RT difference between correctly named related and unrelated trials.

Complex tasks included a computerized version of the Wisconsin Card Sort Task (WCST; adapted from Kimberg, D’Esposito, & Farah, 1997), Raven’s Colored Progressive Matrices (CPM; Raven, 1962), and Raven’s Standard Progressive Matrices (SPM; Raven, 1965). These two Raven’s tasks were included because SD patients showed ceiling effects in previous testing on Raven’s CPM, which was designed for children and adults with brain damage and is thus a simpler task than Raven’s SPM (e.g., Jefferies & Lambon Ralph, 2006). In the WCST, four category cards were arrayed at the top of the screen. These cards varied in color (red, green, blue, yellow), shape (star, circle, triangle, cross), and number of items (1, 2, 3, 4). Participants were given a single card at a time and asked to sort it under one of these category cards by a single dimension. The program provided feedback on whether the card was sorted by the correct dimension. If eight cards were sorted correctly, the sorting dimension changed. The task ended when either 15 categories were sorted or 288 cards were sorted. The

dependent variables included the number of categories sorted and percentage of total erroneous moves.⁴ In Raven's CPM, participants were asked to pick one of six tiles to complete a colored visual pattern in a 2×2 matrix. In Raven's SPM, patterns black and white and trials include both 2×2 and 3×3 pattern matrices. For the latter, participants chose from a set of eight tiles.

As was done with semantic tasks, we compared individual patient scores on each executive task in order to control means using the modified *t* test of Crawford and Howell (1998), except for Raven's SPM, for which norms were available, and for span measures, as comparisons are generally assessed relative to the control range rather than the control average. All statistical tests were two-tailed. For span measures, we took scores below the range of controls to indicate impairment. In tasks for which RTs were collected, outliers were trimmed at 2.5 standard deviations by condition. In executive tasks, RTs were log-transformed to normalize RT distributions. Controls all sorted the maximum number of categories on the WCST, so any patient sorting less than 15 categories was considered impaired relative to controls at number of categories sorted.⁵

Results

Some prior research comparing SA and SD patients has claimed that SA patients tend to have executive deficits and that SD patients do not (e.g., Jefferies & Lambon Ralph, 2006; Thompson et al., 2018). However, these comparisons have relied on complex EF tests and have included only very limited testing of SD patients on EF tasks. In contrast to these claims, some studies suggest that executive deficits are common in SD patients, especially later in the progression of the disease (Macoir et al., 2017), with deficits for SD patients being documented on tasks such as trail making (e.g., George & Mathuranath, 2005) and verbal Stroop (Macoir et al., 2017). In addition, tasks tapping simpler components of executive function have typically not been used in either patient group when they are compared, but performance on these tasks should be revealing regarding the nature of the EF deficit presumed to exist for SA patients (and not SD).

Performance on simple executive tasks

Given the small number of SD patients tested on some of the executive tasks, statistical tests were not carried out contrasting the performance between the groups. Statistical tests were limited to comparing individuals against controls. At a descriptive level in terms of group means, no patient group showed consistently worse performance than controls across all simple executive tasks. However, all patient groups showed poorer performance than controls on several updating tasks, and both SA and SD patients showed poorer performance than controls on multiple inhibition tasks (see Table 8). SD patients appeared to have generally higher digit spans than did SA and non-SA patients, consistent with previous studies showing superior phonological STM in SD patients (Jefferies, Hoffman, Jones, & Lambon Ralph, 2008a; Jefferies & Lambon Ralph, 2006). There did not appear to be any large differences in executive scores between SA, SD, and non-SA patients on any other simple executive tasks.

At the single-subject level, no patient performed poorly on all simple executive tasks, but several patients performed poorly across multiple updating or inhibition tasks (see Table 9; the full results are available in Appendix Table 14). The patients with scores available in Table 9 showed high accuracy in terms of *d'* or percent correct on all or most of the simple EF tasks, demonstrating that these patients understood these tasks even though they may have shown exaggerated effects in terms of RTs. As is shown in Appendix Table 15, all stroke patients and the two SD patients who completed several of these tasks scored at a very high level on the inhibition tasks that did not involve verbal output (i.e., button Stroop and nonverbal Stroop) and on the Trail A and B shifting task. All but one was highly accurate on the one-back task. Five SA patients and five non-SA patients showed significant or marginal impairment on three of four updating tasks. Additionally, two SD patients who were unable to perform *N*-back tasks were also impaired at both span tasks. On inhibition tasks, four SA patients and two non-SA patients were significantly or marginally impaired on two or more of the four tasks. Most SD patients were unable to understand the instructions for at least one of the color Stroop tasks, so it is also noteworthy that 3 of 4 SD patients who were able to perform one of the color Stroop tasks showed exaggerated interference effects. No patient in any group performed poorly on both shifting tasks, and few in any group were impaired on either task. These results are consistent with the group means in showing little difference in the overall pattern of executive impairment between SA and SD patients. Naturally, these claims are tempered by the fact that many of our SD patients were unable to perform several executive tasks. Nonetheless, these results suggest poor executive performance may be present in many SD patients, as every SD patient struggled with at least one simple executive task.

⁴ Numbers of perseverations were also counted. However, perseverations followed essentially the same pattern as percentage errors, so they are not reported.

⁵ We compare patients to our control group for the WCST and for Raven's CPM because public norms were not available for these tasks. Published norms exist only for the traditional WCST, which allows only 128 cards or six categories to be sorted. The computerized version of the WCST allows a maximum of 288 cards or 15 categories to be sorted. Raven's CPM is typically employed on children, so no adult norms exist for this test.

Table 8 Group executive function task performance

Group	Word Span	Digit Span	One-back (% correct)	Two-back (% correct)	Cued task-set shifting (log ms)	Trail Making (B/A, s)	Verbal Stroop (log ms)	Button Stroop (log ms)	Nonverbal Stroop (log ms)	Picture–Word Interference (log ms)	Raven's CPM	Raven's SPM	WCST (categories sorted)	WCST (% error)
	Updating				Shifting		Inhibition				Complex EF			
SA	2.93	3.01	95.1	74.1	0.205	3.79	.450	.293	.120	.032	30.9	41.7	11.1	43.9
Non-SA	3.09	3.99	98.8	82.8	0.205	3.23	.390	.199	.116	.023	33.3	45.8	14.7	25.7
SD	3.51	5.23	100	–	–	3.54	.343	.434	.136	–.081	–	–	7	48.0
Control mean	4.86	6.94	99.3	92.0	0.116	2.74	.191	.175	.116	.028	32	–	15	25.1
Control Range	3.8–8.0	4.4–8.8	97–100	85–97	–0.11–0.41	1.7–4.0	.11–.32	.10–.24	.05–.30	–.002–.04	24–36	–	15–15	18–31

Performance on complex executive tasks

Wisconsin Card Sorting Task At a descriptive level in terms of group means, SA and SD patients were generally worse than controls at the WCST (in both categories sorted and percentages of errors), whereas non-SA patients were not. For those SD patients who showed trouble understanding more basic executive tasks, we did not attempt testing them on the WCST and Raven's matrices. One SD patient attempted the WCST but was unable to understand the instructions. All the patients with scores available in Table 9 showed good enough understanding of the WCST to sort at least four categories correct, which would be difficult without understanding the task's instructions.

At the individual patient level, 4 of 9 SA, 1 of 6 non-SA, and 2 of 2 SD patients sorted fewer than 15 categories correctly, reflecting impaired performance. Two of eight SA patients and 2 of 6 non-SA patients were able to sort 15 categories but showed at least a marginally higher percentage of errors than did controls ($ps < .10$). Patients from all groups had great difficulty with this task. SD patient T.A. made frequent perseverations and could not stay on task.

Raven's Progressive Matrices At the group mean level, SA and non-SA patients scored at similarly high levels on Raven's CPM and Raven's SPM. At the individual patient level, on Raven's SPM, 1 of 6 SA patients was below the 10th percentile of normed scores, and all other patients were in the 25th to 75th percentile. One control obtained quite a low score (24 out of 36), but even when her score was discarded from the control data, only 1 of 6 SA patients was below the range of the controls. We were surprised to discover that our SA patients performed so well on these tasks, given that the Raven's test has been a primary task for observing semantic–executive correlations in SA patients (Almaghyuli et al., 2012; Corbett, Jefferies, Eshan, & Lambon Ralph, 2009a; Corbett, Jefferies, & Lambon Ralph, 2009b; Jefferies & Lambon Ralph, 2006; Noonan et al., 2013; Robson et al., 2012; Thompson et al., 2018). Unfortunately, limited testing time did not allow us collect data from most of our SD patients on the Raven's tasks. The one SD patient who was tested, however, performed very well on Raven's CPM and at the 25th percentile on Raven's SPM, roughly the mean for the SA group. This SD patient's performance is consistent with SD patients' generally good performance on Raven's Matrices tests in the literature.

Executive task summary

No patient group showed a general executive function deficit that spanned updating, shifting, inhibition, and complex executive tasks, but at the single-subject level, patients in both the SA and SD groups showed executive deficits on updating

Table 9 Individual patient executive function task performance

Group	Patient	Word Span	Digit Span	One-Back (% correct)	Two-Back (% correct)	Cued Task-Set Shifting (local switch cost, log ms)	Trail Making (B/A, s)	Verbal Stroop (log ms)	Button Stroop (log ms)	Nonverbal Stroop (log ms)	Picture-Word Interference (interference effect, log ms)	Raven's CPM	Raven's SPM	WCST (categories sorted)	WCST (% error)
Updating															
Inhibition															
Shifting															
Complex EF															
SA	D.Z.	x	*	—	*	—	—	x	*	*O	x	n/a	n/a	n/a	n/a
	P.B.	*	*	*	*	—	—	n/a	—	—	n/a	—	—	—	*
	E.J.	*	*	^	*	x	—	*	^	—	—	—	—	*	*
	C.A.	—	—	—	^	—	—	x	—	—	x	—	—	—	—
	K.A.	x	x	^	*	—	^	*	—	*O	*O	n/a	n/a	*	*
	H.O.	*	*	—	x	—	—	*	—	*O	*	—	—	—	^
	K.V.	x	*	^	*	*	—	*	—	—	*O	n/a	n/a	*	*
	T.Q.	*	*	*	*	—	*	*	*	—	—	—	—	*	*
	E.V.	*	*	—	*	—	—	*	*	*	*	—	—	—	—
Non-SA	N.J.	*	*	—	*	—	—	—	*O	—	*	—	—	—	—
	S.E.	*	*	—	*	—	—	—	*O	—	*O	—	—	—	*
	Q.O.	*	—	—	—	—	—	—	—	—	—	n/a	n/a	—	—
	M.B.	*	*	—	*	—	*	—	*O	—	*O	—	—	*	*
	S.J.	*	*	*	—	—	—	*	*	—	—	n/a	n/a	—	—
	N.L.A.	*	*	—	^	—	*	*	*	—	*	—	—	—	—
	N.L.	x	x	x	x	*	—	—	x	*O	—	n/a	n/a	*	*
SD	K.E.	*	*	x	x	x	x	x	x	x	x	n/a	n/a	n/a	n/a
	E.K.	*	*	x	x	x	x	x	x	x	*O	n/a	n/a	n/a	n/a
	E.G.	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	T.F.	—	—	x	x	x	^	—	x	—	*O	n/a	n/a	n/a	n/a
	T.A.	—	—	—	x	x	*	x	*	—	—	n/a	n/a	*	*
	I.R.	*	—	—	—	*	—	*	—	—	*O	—	—	*	*
Controls	Mean	4.86	6.94	99.3	92	0.116	2.74	0.191	0.175	0.117	0.028	32	n/a	15	25.1
	Range	3.8–8.0	4.4–8.8	97–100	85–97	–0.11–0.41	1.7–4.0	.11–.32	.10–.24	.05–.30	–.002–.04	24–36	n/a	15–15	18–31

* $p < .05$ by Crawford and Howell modified t test. For span tasks, * indicates patient is below control range. ^ $p < .10$. x = patient unable to perform task. o = significant in opposite of expected direction. “—” = unimpaired. “n/a” = no score available.

tasks, on one inhibition task (verbal Stroop), and on the WCST. In terms of group mean performance, non-SA patients did not show executive deficits relative to controls beyond the updating deficits for which they were recruited. However, several individual non-SA patients were impaired on the WCST or on multiple inhibition tasks without showing a multimodal semantic deficit. For example, patient M.B. made many errors on the Wisconsin Card Sort but performed well on all semantic tasks but picture naming and patient N.L.A. performed poorly on three of four inhibition measures but performed well on all semantic tasks. These non-SA patients thus display a dissociation of executive and semantic deficits that suggests executive deficits do not necessitate semantic deficits. SA patients were expected to show executive deficits across simple and complex tasks, and our results confirm this expectation. SD patients, on the basis of some claims in the published literature, were not predicted to show executive deficits, but in fact did show them, in line with other findings (George & Mathuranath, 2005; Gorno-Tempini et al., 2004; Libon et al., 2007; Macoir et al., 2017). The generalizability of our results is limited by the small number of SD patients tested on these tasks, and the severe difficulties some patients experienced in performing the tasks was especially relevant to this small sample. Nevertheless, our findings are consistent with literature suggesting that SD patients have executive problems, perhaps linked to frontal pathology (Brambati et al., 2015; Brambati et al., 2009; Kumfor et al., 2016; Macoir et al., 2017). Thus, these findings motivated further testing of executive function in SD patients to ascertain the relationship between executive function and semantic performance.

Semantic–executive correlations

Method

In semantic–executive correlations, SA and non-SA patients were grouped together ($n = 16$). This was done because semantic control theory suggests that poor executive performance should cause poor semantic performance and our stroke patients were recruited for showing potential for either semantic or executive deficits. One potentially problematic aspect of combining the two groups would be that non-SA patients might create the appearance of a correlation simply because of outlying good performance on all tasks. However, this did not appear to be an issue in our patients, in that there was substantial overlap between SA and non-SA performance on most executive function tasks.

Results

Before correlations were analyzed between semantic and executive tasks, the data were condensed into composite scores

to reduce the number of correlations that were computed. Provided that the tasks going into a composite are correlated with each other, the composites should reflect the underlying executive or semantic dimension better than any individual task, since the influence of any specific task features would be minimized (Nunnally & Bernstein, 1994; Winer, 1971). We derived composites from tasks with correlations greater than $r = .30$ by creating z -scores relative to the control group and averaging across relevant tasks (EF task correlations reported in Appendix Table 16). To understand the relationship between the semantic composite and each executive composite, we also looked at the correlations between the individual semantic tasks and the executive composites. Because of the small size of the SD group and many missing EF scores for this group, correlations between semantic and executive measures were only performed for the stroke group, combining the data from SA and non-SA patients.

The five semantic measures were combined in order to form a semantic composite; the word span, digit span, and two-back scores were combined in order to form an updating composite (primarily reflecting phonological working memory); and the log RT local shift cost for cued task-set shifting and the B–A difference score from trail-making were combined to form a shifting composite. Because the correlations were small between inhibition tasks (range = .07–.25), we chose to use the button Stroop as a representative inhibition measure rather than create a composite. This task contained the largest number of data points and was the most likely of our inhibition tasks to correspond to the inhibition of irrelevant information hypothesized in semantic control. Because all controls sorted categories at ceiling on the WCST, we used percentage of errors on the WCST for correlation with semantic tasks.⁶ Correlations between the two Raven's tests and the WCST variables were too low or comprised too few observations to combine them into a complex EF composite. We derived an executive principal component via the correlations between our executive composites, log button Stroop scores, WCST percent error, and Raven's CPM scores. Three missing scores (D.Z.'s picture naming score, K.V.'s word span, and E.J.'s cued task-set shifting score) were imputed to create a composite for these subjects using multivariate imputation by chained equations (Van Buuren & Groothuis-Oudshoorn, 2011). Because K.A. was missing scores for two semantic tasks, his semantic performance was not included in the comparisons. Note that the signs ($\pm$) of the Stroop costs, shifting costs, and WCST percent error scores were reversed for correlations, so that more positive scores always indicated better performance. This reversal allowed for easier interpretation of

⁶ The results for a composite relative to the stroke group—combining the number of categories sorted, percentage of errors, and a third measure, number of perseverations—showed essentially the same results as those reported here for percentages of errors.

the correlations, in that positive correlations would indicate that better performance on the EF tests was related to better performance on the semantic tasks. Correlations among the executive and semantic measures can be seen in Table 10. The significance of these correlations was not corrected for family-wise Type I errors, because we wished to have a liberal criterion for determining whether such predicted relations would be obtained.

Of our simple executive measures, only updating correlated significantly with our semantic composite, $r(11) = .60$, $p = .03$, suggesting a general relationship between phonological working memory and semantic processing. Though updating correlated moderately to strongly with all semantic tasks, it correlated most strongly with word–picture matching, $r(12) = .85$, $p = .06$, and the environmental sounds test, $r(12) = .71$, $p = .01$. These correlations were likely driven by the large number of related distractors in these tasks, as individuals would have to maintain the spoken word or sound while searching through a large set of items to find the matching picture. The shifting composite showed a moderate but nonsignificant correlation with the semantic composite, but this correlation was not evident across most tasks—only the correlation with CCT word reached significance. The significant correlation between shifting and the CCT word task is difficult to explain, because there is no clear shifting requirement in the CCT word task, and the correlation of shifting with CCT picture was very small. Both CCT word and the shifting task required lexical processing, but other semantic tasks with a clear lexical component did not correlate with shifting (e.g., word–picture matching, picture naming). Our inhibition measure, which we believed to be the executive function component most likely to be related to semantic control as it is described in the literature, showed no correlation with the semantic composite or individual semantic tasks. Thus, the results for the simple EF tasks gave little support to the notion that semantic processing abilities relate to executive function abilities outside of phonological working memory.

As we noted in the introduction, the previous literature has focused on the relation between semantic abilities and complex tasks thought to tap executive function, such as the WCST and

Raven's, documenting moderate correlations for SA patients (Almaghyuli et al., 2012; Corbett, Jefferies, & Lambon Ralph, 2009b; Jefferies & Lambon Ralph, 2006). None of our complex executive and reasoning measures correlated significantly with our semantic composite, though the WCST error percentage, both Raven's tests, and the executive principal component all showed moderate correlations with the semantic composite and with various semantic tasks (all $r_s \geq .30$).

Overall, non-SA patients were less impaired than SA patients on both semantic and EF measures. To control for the possibility that overall aphasia severity is the source of the observed correlations between complex EF tasks and semantic abilities, we used the Western Aphasia Battery Aphasia Quotient (WAB AQ; Kertesz, 1982) as a measure of severity, and partialled that out of the correlations. WAB AQ correlated strongly with our semantic composite, $r(12) = .84$, $p < .01$, updating composite, $r(11) = .86$, $p < .01$, EF principal component, $r(7) = .58$, $p = .10$, WCST percentage of errors, $r(12) = -.61$, $p = .02$, and to a lesser extent with Raven's SPM, $r(8) = .52$, $p = .12$, and Raven's CPM, $r(9) = .44$, $p = .17$. When controlling for WAB AQ, all semantic–EF correlations decreased substantially, with the updating, EF component, and Raven's SPM correlations becoming negative—updating, $r'(11) = -.42$, $p = .17$; EF principal component, $r'(7) = -.21$, $p = .62$; Raven's SPM, $r'(8) = -.05$, $p = .90$ —and the WCST and Raven's CPM correlations having small positive values that were far from significance—WCST, $r'(11) = .05$, $p = .88$; Raven's CPM, $r'(8) = .18$, $p = .65$. These results suggest that all of our observed semantic–executive correlations are driven by overall aphasia severity rather than a relationship between semantic processing and executive function per se.

Sensitivity to distractors in a synonym judgment task

Previous studies have shown that SA patients show exaggerated interference relative to controls for tasks that require resistance to closely related semantic distractors (Noonan et al., 2010). Such findings are consistent with a deficit in semantic

Table 10 Semantic and executive task correlations for SA and non-SA stroke patients

	Executive Principal Component	Updating Composite	Shifting Composite	Button Stroop Effect (log ms)	Wisconsin Card Sort (% errors)	Raven's CPM	Raven's SPM
Semantic composite	.30	.60*	.34	.19	.48	.42	.35
Picture naming	.21	.47	.28	–.20	.37	.55^	.50
Word–picture matching	.65^	.85*	–.10	.11	.35	.35	.25
CCT picture	.03	.48^	.11	.27	.17	.26	.00
CCT word	.27	.41	.58*	.17	.44	–.03	.10
Env. sounds	.50	.71*	–.03	.15	.40	.40	.74*

* $p < .05$; ^ $p < .10$. CPM = Colored Progressive Matrices; SPM = Standard Progressive Matrices.

control. However, such prior studies have not included SD patients and thus it is unclear whether they might show similar effects. To address this issue, we included both SA and SD patients in an examination requiring participants to select a word synonymous with a target word in the presence of a strong versus a weak distractor. This task was based off of a similar task used by Hoffman, Jefferies, and Lambon Ralph (2011a). In addition to the distractor manipulation, we also included a manipulation of the short-term memory load for the synonym judgment task. This was done because Hoffman and colleagues have argued that reduced ability to maintain semantic information in short-term memory is also a reflection of a semantic control deficit. If so, then we should expect the effects of distractor strength and STM load to interact for the SA patients, with especially poor performance in high short-term memory (STM) load, strong-distractor conditions.

Method

A subset of patients were tested on the synonym judgment task, including eight SA patients and two SD patients, each of whom were compared to ten control participants. Our synonym judgment task was a modified version of that used by Hoffman, Jefferies, and Lambon Ralph (2011a; originally from Samson, Connolly, & Humphries, 2007). The task required either choosing the synonym to match a probe or choosing which two words were synonyms. In the task, we manipulated both the strength of distracting information and the STM load required during performance. Both manipulated factors have been argued to draw on semantic control (Hoffman, Jefferies, & Lambon Ralph, 2011a), and thus, according to an additive-factors logic (Sternberg, 1998), one would predict an interaction between STM load and distractor strength.

In the low-STM-load (one-choice) condition of the task, participants were provided with a probe word and were asked to choose which of three words was a synonym to the probe. The answer choices included the correct target, a distractor word, and an unrelated word. In the weak-distractor condition, the probe word (e.g., *hog*) and the correct target (e.g., *pig*) were more strongly related by word association norms (Nelson, McEvoy, & Schreiber, 1998) than were the probe and distractor (e.g., *mud*). In the strong-distractor condition, the probe (e.g., *pilot*) and the target (e.g., *captain*) were more weakly associated than the probe and distractor (e.g., *plane*).

In the high-STM-load (two-choice) condition, participants were provided with a 2×2 grid of words and were asked to choose which two words were synonyms. Here, the four choices included two correct targets, a distractor, and an unrelated word. Example stimuli are provided below, with the correct answer in bold and the distractor underlined:

One-choice, weak distractor:		One-choice, strong distractor:	
	1 <u>mud</u>		1 <u>plane</u>
hog	2 pig	pilot	2 captain
	3 toll		3 dust
Two-choice, weak distractor:		Two-choice, strong distractor:	
1	2	1	2
hog	<u>mud</u>	pilot	<u>plane</u>
3	4	3	4
pig	toll	captain	dust

The low-load condition had a lower STM requirement, because one word of the synonym pair was given to the participant and a maximum of three comparisons had to be made (i.e., between the probe and each other stimulus), whereas in the high-load condition, a maximum of six comparisons had to be made to find which two words were most strongly related. The number of comparisons was crucial to this manipulation, because participants must hold in mind the identities of the words currently being compared, the set of comparisons that have already been made, and the strength of the semantic association of likely answers.

Our materials were modified in two ways from those of Hoffman, Jefferies, and Lambon Ralph (2011a). First, we chose to create a different high STM condition. In the version of the task administered by Hoffman, Jefferies, and Lambon Ralph, the high-STM condition was an auditory version of the low-STM condition. We opted to use a visual high-STM-load condition to avoid confounding modality of presentation with the STM load manipulation. Second, we used a different set of stimuli than did Hoffman, Jefferies, and Lambon Ralph, in order to (1) use more contemporary word association strength norms (Nelson et al., 1998) and (2) employ stimuli normed on American English speakers. To ensure that the selected words were familiar to subjects age-matched to our patients, we asked five healthy older controls to rate 600 words from these norms on a scale of 1 (*unfamiliar*) to 5 (*highly familiar*). Only words with an average rating of 4+ were included as probes, targets, or distractors. From these materials, we created 45 trials with strong distractors and 45 trials with weak distractors. An additional 46 trials with words rated as less familiar (23 with strong, 23 with weak distractors) were used as filler trials, for a total of 136 trials.

Due to the limited sample of words in the association norms, we were unable to match the frequency of the target words across distractor conditions. The target words in the weak-distractor condition had a higher log SUBTLEX frequency than did those in the strong-

distractor condition (mean strong, 2.81; mean weak, 3.37; $t(87.2) = 4.67$, $p < .001$).⁷

The stimuli in the synonym task were nouns presented in a white font on a black computer screen. Answer choices in each position on the screen had an equal chance of being the correct answer across trials. Testing took place over two sessions that occurred at least a week apart. The 136 trials were divided randomly into two trial sets with equal numbers of strong- and weak-distractor trials. In each experimental session, half of the trials were presented in the low-STM condition and half in the high-STM condition such that after both sessions participants had seen all stimuli in both STM load conditions. The order in which the low-load and high-load blocks were presented was counterbalanced across participants and patient groups. The dependent measure in the task was accuracy in the four STM Load $\times$ Distractor conditions (i.e., one-choice, weak distractor; one-choice, strong distractor; two-choice, weak distractor; two-choice, strong distractor).

To analyze the synonym judgment task, 2 (STM load) $\times$ 2 (distractor strength) repeated measures analyses of variance (ANOVAs) were performed on each individual patient's accuracy. Limits of word association norms precluded the creation of stimuli in which the same item could serve as both a strong and weak distractor.⁸ Because each trial was presented in both STM conditions but trials differed across distractor strength conditions, STM load condition (low or high) was analyzed as a within-subjects factor, and distractor relatedness (weak or strong) was analyzed as a between-subjects factor, both across items. Comparisons of the distractor effects in low- and high-load conditions by t test were also planned a priori for the individual patient ANOVAs. Finally, we tested whether each patient's accuracy was above chance in each of the four task conditions, with chance being 1/3 in the low-load (one-choice) condition and 1/6 in the high-load (two-choice) condition. Because this task was introduced late in the study, only two SD patients were tested on this measure.

Results

Control group accuracy and individual patient accuracy are shown in Fig. 2. Controls performed near ceiling in the low-load condition (97% correct overall) and were somewhat

worse in the high-load condition (88%). They showed evidence of a distractor strength effect, particularly in the high-load condition (85% strong vs. 91% weak distractor). An ANOVA on control performance confirmed that they showed a significant distractor effect, $F(1, 494) = 8.7$, $p = .003$; a significant STM load effect, $F(1, 494) = 67.84$, $p < .001$; and a significant interaction between the two, $F(1, 494) = 4.48$, $p = .03$, where the distractor effect in the low-STM load condition was limited by a ceiling effect.

Nearly all SA and SD patients—with the exception of SA patient H.O.—appeared to show poorer performance than controls, especially in the strong-distractor conditions. Statistical tests confirmed these impressions: All SA and SD patients were substantially less accurate than controls ($ps < .05$), except SA patient H.O. Two SA patients showed no difference from chance performance in one or more condition: T.Q. in both strong-distractor conditions, and P.B. in the low-load, strong-distractor condition ($p > .05$).

Every patient appeared to show a strong-distractor effect, and many appeared to do more poorly with a high STM load. However, only one patient (E.J.) appeared to show the expected interaction, with a stronger effect of distractor strength in high- versus low-STM conditions. ANOVAs on individual patient accuracy confirmed that most semantically impaired patients showed strong-distractor effects, but only one showed an STM load effect, and no patient showed the expected interaction effect. Both SD patients and all but one SA patient (H.O.) showed distractor strength effects ($F_s > 5.11$, $ps < .03$), all of which were significantly stronger than the effects among controls ($ts > 3.18$, $ps < .05$). Only SA patient H.O. showed a main effect of STM load, $F(1, 88) = 14.59$, $p < .001$, but this effect was no larger than that among controls [$t(10) = 0.09$, $p = .93$]. Two SA patients (K.V. & H.O.) showed significant interactions between STM load and distractor strength, $F_s > 5.02$, $ps < .03$, and one SA patient (K.A.) showed a marginally significant interaction, $F(1, 88) = 3.25$, $p = .07$. K.V. and K.A. showed distractor strength effects in the low-load condition, $F_s > 14.13$, $ps < .001$, but not in the high-load condition, $F_s < 1.20$, $ps > .27$ —the reverse of the expected pattern. H.O. showed the same interaction pattern as controls, with strong distractors hurting performance only in the high-load condition, $F(1, 88) = 4.15$, $p = .04$, due to a ceiling effect in the low-load condition. Like most SA patients, neither SD patient showed effects of STM load or an interaction between factors ($F_s < 1$).

Thus, the patient-level results show exaggerated effects of distractor strength across both SA and SD patients, but no exaggerated STM effects or predicted interactions of distractor strength and STM condition for any patients. These results call into question the claim that SA and SD patients differ in their abilities to inhibit strong distractors due to the executive control deficits underlying the semantic processing deficits of SA but not SD patients. Though our small SD patient sample

⁷ The stimuli from Hoffman, Jefferies, and Lambon Ralph (2011a) also showed frequency differences between distractor conditions for the probe words [$t(39.8) = 2.21$, $p = .03$], and target words [$t(39.7) = 3.56$, $p = .001$], and a marginal difference for the distractor words [$t(38.7) = 2.01$, $p = .05$]. Though the study that had originally used these stimuli had controlled for the influence of frequency (Samson, Connolly, & Humphreys, 2007), other studies that have used this task did not (Hoffman, Jefferies, & Lambon Ralph, 2011a; Noonan et al., 2010).

⁸ The stimuli of Hoffman, Jefferies, and Lambon Ralph (2011a) shared this property. However, their study had statistically examined only the distractor manipulation, presumably because separate control groups were tested on the different STM load conditions.

Fig. 2 Synonym judgment control group and individual patient accuracies, across short-term memory (STM) load and distractor strength conditions. Error bars represent ± 1 SEM

limits the strength of our conclusions, our result is consistent with that of another previous study showing a “semantic control” deficit in SD patients (Thompson et al., 2018). Our results also challenge the notion that semantic STM and the ability to resist strong distractors derive from the same control deficit, given that no interaction between the STM load and distractor strength factors was obtained.

General discussion

The primary purpose of this study was to investigate whether patterns of semantic and executive performance for SD and SA patients that have previously been claimed to distinguish semantic knowledge and semantic control deficits would be upheld when testing these patient groups on the same tasks. The tasks we employed included examinations of semantic knowledge, basic executive function, and the ability to inhibit responses to strong distractors. Other phenomena purported to distinguish these semantic deficits and patient groups—such as picture naming errors, effects of phonemic cueing, frequency/familiarity effects (see the introduction)—were not examined in the present study. Here we will briefly

summarize our results with respect to the phenomena addressed in this study before moving on to more general theoretical considerations.

1. *Item and task consistency.* Contrary to some previous evidence (Corbett, Jefferies, Eshan, & Lambon Ralph, 2009a; Jefferies & Lambon Ralph, 2006; Thompson et al., 2018), but in line with other findings (Garrard & Carroll, 2006; Snowden et al., 2018; Snowden, Thompson, & Neary, 2004; Warrington & Shallice, 1984), our study showed little difference between SA and SD patients for item and task consistency. These results may seem surprising, given prior reports of consistent performance in SD patients; however, consistency in SD has most often been demonstrated for performance with repeated presentations of the same items in the same task (Coughlan & Warrington, 1981; Snowden et al., 2004; Warrington, 1975; Warrington & Cipolotti, 1996; Warrington & Shallice, 1984). According to the CSC framework, such consistency should also be expected for SA patients with repeated presentations of items in a task, given that the task demands would remain identical across presentations.

2. *Executive task performance.* When tested on the same EF tasks, the SA and SD patients in the present study performed similarly. Patients in both groups performed at a high

level on the Raven's Progressive Matrices tasks but were impaired on other EF tests. The similarity between groups was observed for simple and for more complex EF tasks—though comparisons on the more complex tasks were often difficult to make because many of the SD patients were unable to perform them. The one apparent difference between the SA and SD groups was in updating ability, but this difference was in tasks that primarily tapped phonological retention. Given SD patients' typically good phonological abilities (e.g., Woollams, Cooper-Pye, Hodges, & Patterson, 2008) and the impairment of phonological processing in individuals with aphasia (Dial & Martin, 2017), such a difference would be expected.

3. Relation of EF to semantic task performance. In general, EF performance for simple tasks, including inhibition tasks, did not relate to semantic performance for SA patients. The only suggestion of a correlation for SA patients was for two complex tasks—the WCST and Raven's tasks—but these correlations appeared to be mediated mainly through patients' general language impairments, given that the correlations were far from significant when aphasia severity was included as a control. Two SD patients who could be tested on the complex WCST also showed impairments.

4. Effects of distractor strength. Although the SA patients showed exaggerated effects of strong distractors, so did the two SD patients who were tested on the same task. As we noted in the introduction, typically SD patients have not been tested on the same tasks given to SA patients to assess semantic control abilities, such as the ability to resist distraction. In the one manipulation on which they have been tested, SD patients have likewise shown exaggerated effects (i.e., with the ability to match word meanings across longer vs. shorter semantic distances; Thompson et al., 2018). The present study is the first to provide data on SD patients' performance with varying levels of distractor strength, and is only the second study to provide data on purported semantic control manipulations in SD patients.

5. Semantic deficits of non-SA patients and SA patients. We found that even though the non-SA patients at the individual level sometimes had EF deficits similar to those of SA patients, they did not show semantic deficits like those of the SA patients.

Thus, overall it is hard to argue from these data that the SD and SA have semantic deficits deriving from highly distinct sources and that the semantic deficits of the SA group derive from a semantic control mechanism that is related to executive function deficits.

One might counter that there are other differences in the performance of the two groups that we have failed to address—specifically, differential effects of word frequency, cueing, and error types in naming (Almaghyuli et al., 2012; Jefferies & Lambon Ralph, 2006; Jefferies, Patterson, & Lambon Ralph, 2008b; Noonan et al., 2010). For instance, as noted in the introduction, the CSC literature claims that

SA patients fail to show typical frequency effects, due to the greater semantic diversity of more frequent items causing greater competition in meaning selection (Hoffman, Lambon Ralph, & Rogers, 2013b; Hoffman, Rogers, & Lambon Ralph, 2011b). However, a recent study from our lab using a large-scale database of individuals with aphasia found that those with and without multimodal semantic deficits showed indistinguishable effects of frequency and semantic diversity on their word processing performance (Chapman & Martin, 2018). We also note that there is controversy regarding the extent to which SA and SD patients show differential error types in naming (e.g., Budd et al., 2010). Moreover, even if these additional distinctions could be found to generally hold, it would remain a challenge to explain how these effects arose in the absence of differential effects of consistency, relations to executive function, and strength of distractors in terms of an explanation based on semantic control versus storage deficits.

One might also argue that the SD patients that we tested in the present study are qualitatively different from the SD patients previously compared with SA patients (e.g., Corbett, Jefferies, Eshan, & Lambon Ralph, 2009a; Jefferies et al., 2007; Jefferies, Hoffman, Jones, & Lambon Ralph, 2008a; Jefferies & Lambon Ralph, 2006; Jefferies, Patterson, & Lambon Ralph, 2008b; Jefferies, Rogers, Hopper, & Lambon Ralph, 2010; Noonan et al., 2013a; Thompson et al., 2018), given that those in the present study had frontal or parietal brain damage. However, it is not clear that the patients from previous studies did not have such damage, and it would be surprising if they, in fact, did not. In previous studies comparing SA and SD, the method of delineating SD patient damage was not described, and it may be that subtle frontal damage was overlooked if atrophy was assessed only by visual inspection, because frontal damage in these patients is less prominent than temporal atrophy. It may also be that the goal of the authors was not to fully describe the patients' atrophy but only to affirm that the behavioral diagnosis of SD coincided with the typical anterior temporal damage—a perfunctory step for establishing an etiological diagnosis of SD. Both of these explanations are plausible, as frontal damage appears to be typical in SD, even if it is less prominent than temporal damage (e.g., Agosta et al., 2010; Borroni et al., 2007; Brambati et al., 2015; Brambati et al., 2009; Cerami et al., 2017; Gorno-Tempini et al., 2004; Kumfor et al., 2016; Macoir et al., 2017; Mummery et al., 2000; Rosen et al., 2002). Previous studies comparing SA and SD patients (e.g., Corbett, Jefferies, Eshan, & Lambon Ralph, 2009a; Jefferies & Lambon Ralph, 2006; Thompson et al., 2018) may have provided an additional indicator of extratemporal damage by including more extensive executive testing in both SA and SD patients, but, as we described above, the executive testing in these studies was limited, especially in SD patients. Without clear delineation of SD patients' brain damage from previous comparative studies, it is difficult to conclude that

they did not show similar damage to the patients in the present study. It is worth noting that even the SD patients in our study without frontal damage (I.R. and E.G.) were similar to our SA patients—they showed broad semantic deficits; only one showed significant consistency across all tasks; neither showed consistency in all semantic task pairings; one patient showed deficits on several executive tasks (including the WCST); and both showed strong effects of distractor interference like our SA patients. This finding shows that even “typical” SD patients may appear very similar to SA patients.

Implications for storage–access distinction

The present findings suggest that the distinction between storage and control in semantic processing is not as strongly supported by contrasting patterns of performance for the SD and SA patient groups as has been previously argued. In general, these groups’ performance on a number of dimensions was quite similar. Of course, the question then arises regarding whether the distinction of storage versus access should be maintained. There are general theoretical grounds for wanting to maintain this distinction. In the perception domain, a vast literature drawing on behavioral and neuroimaging studies of healthy and brain damaged individuals has shown dissociations between the brain regions involved in representing sensory features and those that direct attention to those representations, (Corbetta & Shulman, 2002; Duncan, 2010; Hahn, Ross, & Stein, 2006; Ruff et al., 2006; Tamber-Rosenau, Esterman, Chiu, & Yantis, 2011). By analogy, one might expect to find a similar dissociation in the semantic domain, even though the representations are more complex and abstract and components of semantic representations may be distributed across many brain regions (Binder, Desai, Graves, & Conant, 2009; A. Martin & Chao, 2001; Patterson et al., 2007). In fact, as we discussed in the introduction, numerous studies have pointed to regions in the left inferior frontal gyrus as being involved in controlled retrieval of semantic representations and the resolution of competition between semantic representations—most typically this is assessed in tasks involving verbal representations (Badre & Wagner, 2007; Snyder, Banich, & Munakata, 2011; Thompson-Schill et al., 1997). Given these theoretical and empirical factors, one needs to consider why the present results were obtained, and what different approaches might be taken that might provide support for the division between semantic storage and control. Below we discuss two possible reasons for the present results—specifically, overlap in anatomical damage and the use of inappropriate behavioral tests—and suggest possible means of addressing these issues.

Overlap in anatomical damage obscures group differences

Although individuals with aphasia typically have left middle cerebral artery infarcts that primarily affect perisylvian regions

and spare the temporal pole and inferior temporal and fusiform gyri, lesions in these regions do occur. For instance, in the Walker et al. (2011) voxel-based lesion–symptom mapping study of semantic errors in naming for individuals with aphasia, at least 5 of 64 had damage to voxels in the temporal pole and inferior temporal gyri. In fact, that study revealed that the region correlating most highly with semantic errors in naming was the left temporal pole.⁹ On the other hand, it is evident that those diagnosed with SD have atrophy that extends into the frontal lobes (Brambati et al., 2015; Brambati et al., 2009; Macoir et al., 2017), with the extent of frontal atrophy related to the severity of the disease (Brambati et al., 2015; Brambati et al., 2009; Hodges, Graham, & Patterson, 1995). Consequently, it is possible that the contrast between the two groups is obscured by at least some individuals in both groups having substantial damage to both a semantic hub in the temporal lobe and to fronto-parietal regions involved in semantic control. The present study provides confirmation that damage extending into these regions thought to be involved in semantic control (and executive function in general) is evident for those diagnosed as SD patients. Again, although our sample of SD patients was small, it provides a clear example that the extent of damage and the degree of executive ability are important to consider when comparing SA and SD.

To address this concern, one might limit comparisons to SA patients who have frontal or parietal damage but not damage to anterior inferior temporal regions and to SD patients who have primarily inferior temporal lobe atrophy but minimal frontal lobe atrophy. Such an approach takes for granted the existence of a semantic hub in inferior temporal regions, though it should be acknowledged that there is considerable debate about which brain areas are involved in semantic representation and whether a semantic hub in anterior temporal lobe regions is critical for all types of semantic knowledge (e.g., Binder et al., 2009; Schwartz et al., 2009; Simmons, Reddish, Bellgowan, & Martin, 2010). Also, limiting damage among SD patients to those with minimal frontal damage would not necessarily leave semantic control mechanisms intact, even if such functions are localized to frontal and temporo-parietal regions. The atrophy associated with SD affects both gray matter and white matter tracts in the temporal lobe including tracts connecting temporal regions to frontal regions (e.g., Brambati et al., 2015; Whitwell et al., 2010). Such tracts have been argued to be involved in semantic control (de Zubizaray, Rose, & McMahon, 2011; Harvey, Wei, Ellmore, Hamilton, & Schnur, 2013), and there is evidence that such is the case in at least one study that has found disrupted resting state functional connectivity of frontal

⁹ It should be noted that this relation persisted even when factoring out semantic comprehension deficits, suggesting that the relation was not due to semantic representation disruption per se, but instead to the mapping between semantics and word representations for output.

executive systems due to anterior temporal atrophy in SD (Farb et al., 2013). Consequently, evidence of semantic control deficits could be observed for the SD patients due to the disruption of these tracts.

Because of these complications, it would be ideal to have sufficient neural and behavioral data on a large number of individuals from both groups. With these data, the degree of damage to various regions could be assessed to sort out possible disruptions of specific semantic knowledge domains and the contribution of white matter damage to performance. It should be noted, however, that such an approach would seem to be particularly challenging for the SD patients, given the rarity and progressive nature of their disease and consequently lesser time available for extensive testing.

The present study provides evidence that executive deficits—although typically less severe than semantic deficits (Brambati et al., 2015)—are more common in SD patients than is often discussed. Although our findings are limited by the small sample of SD patients tested, we believe such a finding should encourage authors to provide specific descriptions of patient atrophy and executive ability in order to provide clear comparisons between SA and SD patients. As it stands, the previous literature comparing SA and SD is ambiguous concerning the frontal or parietal damage of its SD patients.

Tests for distinguishing semantic knowledge and control deficits are inappropriate Of course, to determine whether any anatomical division supports the distinction between semantic knowledge and semantic control, it is necessary to have the appropriate empirical tests to make this distinction. This section evaluates the appropriateness of the tests discussed in this article, which appear not to distinguish well between the deficits, and suggests possible alternatives.

Consistency criteria In a previous review of the distinction between storage and refractory access deficits, Rapp and Caramazza (1993) raised a number of theoretical concerns that can be echoed here for the storage versus semantic control distinction. Regarding item consistency, it is not clear why markedly different patterns are predicted for storage versus control deficits. The theoretical prediction of consistency for SD patients seems to rest on the assumption that knowledge of concepts degrades in a complete fashion, such that if a concept is lost, all aspects of its meaning are lost. If this were true, then it follows that performance should be impaired for those items irrespective of the task involved. However, considerable evidence suggests that features of concepts are lost over time as the disease progresses, with distinctive features that differentiate close category members being lost earlier and features of concepts more typical for its semantic category being better preserved until later (Bozeat et al., 2003; Hodges et al., 1995; Lambon Ralph, Graham, & Patterson, 1999; Lambon Ralph

& Patterson, 2008; Patterson & Erzinçioğlu, 2008). A selective loss of distinctive features could lead to inconsistency in performance. For example, if the features that distinguish apples and pears are degraded, such that when seeing a picture of a pear, both names are activated to a similar degree, then naming might be correct on one trial but incorrect on another, depending on random noise in the system that determined which name won out during selection.

On the other side of the issue, one could argue for cross-task consistency for individuals with a semantic control deficit. If one task requires greater semantic control than another (e.g., the Camels and Cactus Task as compared to word–picture matching), then performance on that task would be predicted to be lower overall for those with greater control deficits; however, the rank ordering of individuals' performance on the two tasks could be the same and relate to individuals' degree of semantic control deficit, resulting in a correlation between performance on the tasks. In fact, it is somewhat unclear how one would predict no relation between performance on two tasks if the deficit is purely one of cognitive control. That is, how could some patients with a control deficit perform relatively better on a task with greater control demands whereas others perform worse? Such an outcome would be required in order for no correlation to be obtained. For the lack of correlation to hold, one might contend that different patients have different types of control deficits that differentially affect performance on different tasks, but this would considerably complicate the theoretical claims. It is more likely that the observed absent correlations in the literature (e.g., Jefferies & Lambon Ralph, 2006) are due to deficits beyond semantic control that impact various tasks differently (e.g., speech perception or word production difficulties).

Because both consistency and inconsistency could be predicted from either SA or SD, item and cross-task consistency appear to lack diagnosticity in distinguishing storage and access deficits.

Influence of distractors Regarding the influence of highly associated distractors, again it seems that deficits of degraded storage or semantic control could result in exaggerated interference effects. In the synonym judgment task used in the present study, based on that from Hoffman, Jefferies, and Lambon Ralph (2011a), association strength was manipulated between two pairs of words: the probe and the distractor and the probe and the target. Thus, in the high interference condition a strongly associated distractor and a weakly associated synonym were paired with a probe (e.g., plane vs. captain for pilot) and in the low interference condition a weakly associated distractor and a strongly associated synonym were paired with a probe (mud vs. hog for pig).

Although little experimental work has been directed at understanding how subjects make choices in the synonym judgment task, some prior work has shown that RTs for synonymy judgments depend on the semantic distance between synonyms as measured by LSA (Bullinaria & Levy, 2007). Thus, one might expect that the relative semantic distance of the target versus the distractor to the probe would predict performance. The logic of the Hoffman et al. conclusion was that greater semantic control would be needed in the high interference condition to suppress the selection of the highly associated distractor and thus a deficit in control could be the source of exaggerated effects. However, on the principle that brain damage will have a greater effect on weaker semantic relations (e.g., Rogers et al., 2004), greater interference would also be expected for individuals with semantic knowledge deficits in the high interference condition since the weak relation between the target and probe will be more affected by brain damage than the strong semantic relation between the probe and highly associated distractor. That is, the difference in the semantic distance of the probe to the target versus distractor in the high-interference condition will be larger for those with knowledge deficits. Consequently, an exaggerated effect of a distractor should appear whether the disruption is to semantic knowledge or control.

Relation of semantic performance to executive function

Several issues here make problematic the evaluation of executive function's relations to semantic task performance. First, even if one assumed that those with storage deficits do not have executive function deficits, one might expect that premorbid but persisting individual differences in executive function would nonetheless impact performance on semantic tasks if performance on the task depends on biasing attention toward certain semantic features and away from others. One might expect such abilities to be particularly relevant when semantic knowledge is damaged such that attention must be directed to weakened links between the features of concepts. Thus, it is unclear why semantics and EF would be not be found to correlate in the SD group, unless they are performing at ceiling on a particular EF test, as was the case with Raven's matrices in Jefferies and Lambon Ralph (2006).

Another problematic issue related to the claim that the EF problems of the SA group are the source of their semantic deficits is that, in the present study, the non-SA group who demonstrated EF deficits did not show general semantic deficits. Similarly, the patients recruited for EF problems in Thompson et al. (2018) showed semantic control problems similar to those among SA patients, but consistently better performance on semantic tasks. Moreover, prior results, even some of those supporting

the CSC framework, have shown that some SA patients who showed evidence of strong-distractor effects on semantic tasks did not show deficits on a variety of EF tasks. For instance, in Hoffman et al. (2013a), of three patients showing evidence of a semantic control deficit on tasks with strong distractors, one (A.B.U.) showed no deficit on any EF task, and the two other patients showed deficits on only a few tasks and not on others. Particularly for A.B.U., the authors argued that a specific semantic control function can be damaged irrespective of damage to other EF functions. If so, then it is unclear whether other EF deficits play a causal role in semantic processing deficits. That is, the specific semantic-control function may be both necessary and sufficient. Deficits in other EF components may be irrelevant—though they may be associated with a disruption of semantic control, due to anatomical proximity of the brain regions subserving semantic control and EF. The findings from our non-SA patients would seem to support that conclusion.

Other possible empirical tests One set of findings from the earlier literature on the semantic knowledge–access distinction came from tasks that supported the notion of refractory access—that is, poorer performance with repeated testing of the same item on the same task, which is exaggerated for items drawn from the same semantic category (Hamilton & Martin, 2010; Warrington & Cipolotti, 1996). These findings seem difficult to account for on the basis of a knowledge deficit and strongly suggest that a process involving access or control has been disrupted. Numerous recent findings have indicated that even healthy individuals show impaired performance with repeated testing in a task such as the blocked cyclic naming task, which involves repeated naming of a small set of items from the same semantic category (Belke, Meyer, & Damian, 2005; Crowther & Martin, 2014). The findings suggest that repeated selection involves increasing competition, particularly for semantically related items. Resolution of this competition appears to involve the left inferior frontal lobe, as damage to this region exaggerates the increase in errors across repetitions for related versus unrelated sets of items (Harvey & Schnur, 2015; Schnur et al., 2009).

Tests of SA patients have shown some evidence of exaggerated effects on such tasks (Jefferies et al., 2007—though in terms of exaggerated effects of semantic relatedness rather than in abnormally large increases in differences between related and unrelated items across repetitions); however, only very limited testing has been done with SD patients to determine if they show a lesser degree of interference buildup that is more in line with that of controls subjects. At least one SD patient has been tested on cyclic word–picture matching, and he showed a large effect of semantic relatedness, to a degree similar to

that among SA patients (Jefferies et al., 2007). It should be noted that these blocked cyclic tests have been limited to those involving either picture naming or word–picture matching, and thus involve mapping verbal representations to meaning (e.g., Biegler et al., 2008; Harvey & Schnur, 2015). It seems that a semantic task like this, which does not involve words, has not been developed.¹⁰

With respect to EF, the present results indicate that SA patients are not uniformly impaired on EF tests, because they performed within the normal range on shifting and nonverbal inhibition tasks. Although some patients showed impairment on the complex WCST, the previous and current evidence suggests that this deficit may be related to their impaired language abilities (Baldo et al., 2005). Future testing should focus on tasks that target semantic control specifically. Many such tasks have been proposed—for example, semantic distance, distractor strength, and miscuing manipulations (Noonan et al., 2010; Thompson et al., 2018)—but all of these tasks depend on the use of verbal representations, and their specific relationship to a semantic executive function has not been well-justified. It would be important to develop nonverbal semantic control tasks to ensure that the task taps into the semantic level rather than to the mapping between semantics and lexical items. Tasks that target inhibition at the semantic versus the phonological level could be used to verify that the difficulty is specific to the semantic level.

An intriguing possibility would be to use a task such as those developed by Hasher and colleagues, which would show that older adults tend to form more extraneous associations than do younger adults—referred to as *hyper-binding* (Campbell & Hasher, 2018; Campbell, Hasher, & Thomas, 2010). For instance, when given a one-back task on pictures superimposed with irrelevant words and then surprised with a paired-associate memory task, older adults tend to show a memory advantage for picture–word pairs preserved from the one-back task, and a disadvantage for disrupted pairs. Such a task could easily be changed to be purely nonverbal. If SA patients showed enhanced performance relative to SD patients for preserved pairs of pictures under such conditions, it would be hard to argue that the results for the SA patients derived from a disruption of semantic knowledge per se. Naturally, we would also expect SA patients to show a greater difference between preserved and disrupted pairs than would healthy older adults, signaling their decreased inhibitory abilities.

¹⁰ Belke (2013) used a semantic classification task involving manmade/natural judgments of pictured objects and showed facilitative effects of sequential presentations of semantically related items. However, in such a task, competition for the selection of a response does not occur, since all related items would typically lead to the same response.

Conclusions

Previous work had suggested that semantic deficits in stroke patients are caused by executive deficits, whereas the semantic deficits in semantic dementia are caused by knowledge loss. Our study assessed the behavioral patterns that have supported these distinctions, for patients from both groups tested on the same tasks. We found that item consistency, semantic task correlations, performance on executive tasks, and effects of distractor strength did not clearly distinguish these two semantic disorders as was previously suggested. Furthermore, with our study we attempted to better define what executive capacities underlie semantic deficits in stroke, by including tasks tapping more specific cognitive abilities. To this end, we found that complex executive tasks and phonological working memory correlate to some degree with semantic abilities, but other simple components of executive function, such as inhibition and shifting, do not. Further investigations of complex executive–semantic correlations suggest that language abilities, rather than executive function per se, are the source of the observed relationship. Overall it appears that, in general, several features of performance that have been argued to distinguish the semantic storage–control distinction do not map in a clear fashion onto the two patient groups.

The lack of a distinction between the SA and SD patients in our study suggests that one should not use syndromic categories, such as “semantic aphasia” and “semantic dementia,” as predictors of task behavior. It is very likely that a variety of deficits contribute to the semantic problems that we see in any patient. Concept loss and executive deficits were both potential contributors to semantic deficits in the SA and SD patients we tested. A better method of grouping patients might be based on lesion location, which could show the source of commonalities or differences more accurately than a description of their behavior alone, and might explain the similarities seen between SA and SD in the present study. A detailed investigation of individual patient behavior would likewise provide much deeper insight into the specific factors affecting and affected by a patient’s semantic deficits. Although our results do not argue against the possible existence of semantic storage and semantic access deficits, they do suggest that other data are needed to support these distinctions. Moreover, new empirical approaches are needed that would provide more convincing tests of storage versus access deficits.

Acknowledgments This work was supported by the T. L. L. Temple Foundation.

Open Practices Statement The data and materials discussed in this article are available from the authors upon request.

References

- Agosta, F., Henry, R. G., Migliaccio, R., Neuhaus, J., Miller, B. L., Dronkers, N. F., . . . Gorno-Tempini, M. L. (2010). Language networks in semantic dementia. *Brain*, 133, 286–299.
- Allen, C. M., Martin, R. C., & Martin, N. (2012). Relations between short-term memory deficits, semantic processing, and executive function. *Aphasiology*, 26, 428–461. doi:<https://doi.org/10.1080/02687038.2011.617436>
- Almaghyuli, A., Lambon Ralph, M. A., & Jefferies, E. (2012). Deficits of semantic control produce reverse frequency effects: Evidence from neuropsychology and dual task methodology. *Neuropsychologia*, 50, 1968–1979. doi:<https://doi.org/10.1016/j.neuropsychologia.2012.04.022>
- Badre, D., & Wagner, A. D. (2007). Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia*, 45, 2883–2901.
- Baldo, J. V., Dronkers, N. F., Wilkins, D., Ludy, C., Raskin, P., & Kim, J. (2005). Is problem solving dependent on language? *Brain and Language*, 92, 240–250. doi:<https://doi.org/10.1016/j.bandl.2004.06.103>
- Basso, A., Casati, G., & Vignolo, L. A. (1977). Phonemic identification defect in aphasia. *Cortex*, 13, 85–95.
- Belke, E. (2013). Long-lasting inhibitory semantic context effects on object naming are necessarily conceptually mediated: Implications for models of lexical-semantic encoding. *Journal of Memory and Language*, 69, 228–256.
- Belke, E., Meyer, A. S., & Damian, M. F. (2005). Refractory effects in picture naming as assessed in a semantic blocking paradigm. *Quarterly Journal of Experimental Psychology*, 58, 667–692.
- Biegler, K. A., Crowther, J. E., & Martin, R. C. (2008). Consequences of an inhibition deficit for word production and comprehension: Evidence from the semantic blocking paradigm. *Cognitive Neuropsychology*, 25, 493–527. doi:<https://doi.org/10.1080/02643290701862316>
- Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, 19, 2767–2796. doi:<https://doi.org/10.1093/cercor/bhp055>
- Borroni, B., Brambati, S. M., Agosti, C., Gipponi, S., Bellelli, G., Gasparotti, R., . . . Padovani, A. (2007). Evidence of white matter changes on diffusion tensor imaging in frontotemporal dementia. *Archives of Neurology*, 64, 246–251.
- Bozeat, S., Gregory, C. A., Lambon Ralph, M. A., & Hodges, J. R. (2000a). Which neuropsychiatric and behavioural features distinguish frontal and temporal variants of frontotemporal dementia from Alzheimer's disease? *Journal of Neurology, Neurosurgery and Psychiatry*, 69, 178–186.
- Bozeat, S., Lambon Ralph, M. A., Graham, K. S., Patterson, K., Wilkin, H., . . . Hodges, J. R. (2003). A duck with four legs: Investigating the structure of conceptual knowledge using picture drawing in semantic dementia. *Cognitive Neuropsychology*, 20, 27–47.
- Bozeat, S., Lambon Ralph, M. A., Patterson, K., Garrard, P., & Hodges, J. R. (2000b). Non-verbal semantic impairment in semantic dementia. *Neuropsychologia*, 38, 1207–1215. doi:[https://doi.org/10.1016/S0028-3932\(00\)00034-8](https://doi.org/10.1016/S0028-3932(00)00034-8)
- Brambati, S. M., Amici, S., Racine, C. A., Neuhaus, J., Miller, Z., Ogar, J., . . . Gorno-Tempini, M. L. (2015). Longitudinal gray matter contraction in three variants of primary progressive aphasia: a tensor-based morphometry study. *NeuroImage Clinical*, 8, 345–355.
- Brambati, S. M., Rankin, K. P., Narvid, J., Seeley, W. W., Dean, D., Rosen, H. J., . . . Gorno-Tempini, M. L. (2009). Atrophy progression in semantic dementia with asymmetric temporal involvement: A tensor-based morphometry study. *Neurobiology of Aging*, 30, 103–111. doi:<https://doi.org/10.1016/j.neurobiolaging.2007.05.014>
- Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, 41, 977–990. doi:<https://doi.org/10.3758/BRM.41.4.977>
- Budd, M. A., Korte, K., Cloutman, L., Newhart, M., Gottesman, R. F., Davis, C., . . . Hillis, A. E. (2010). The nature of naming errors in primary progressive aphasia versus acute post-stroke aphasia. *Neuropsychologia*, 24, 581–589.
- Bullinaria, J. A., & Levy, J. P. (2007). Extracting semantic representations from word co-occurrence statistics: A computational study. *Behavior Research Methods*, 39, 510–526. doi:<https://doi.org/10.3758/BF03193020>
- Butts, A. M., Machulda, M. M., Duffy, J. R., Strand, E. A., Whitwell, J. L., & Josephs, K. A. (2015). Neuropsychological profiles differ among the three variants of primary progressive aphasia. *Journal of International Neuropsychological Society*, 21, 429–435.
- Campanella, F., Mondani, M., Skrap, M., & Shallice, T. (2009). Semantic access dysphasia resulting from left temporal lobe tumours. *Brain*, 132, 87–102.
- Campbell, K. L., & Hasher, L. (2018). Hyper-binding only apparent under fully implicit test conditions. *Psychology and Aging*, 33, 176–181.
- Campbell, K. L., Hasher, L., & Thomas, R. C. (2010). Hyper-binding: A unique age effect. *Psychological Science*, 21, 399–405.
- Cerami, C., Dodich, A., Greco, L., Iannaccone, S., Magnani, G., Marcone, A., . . . Perani, D. (2017). The role of single-subject brain metabolic patterns in the early differential diagnosis of primary progressive aphasia and in prediction of progression to dementia. *Journal of Alzheimer's Disease*, 55, 183–197.
- Chapman, C., & Martin, R. C. (2018). *Semantic diversity affects word processing similarly for PWA with and without semantic deficits: Evidence against the semantic control hypothesis*. Platform presentation at the meeting of the Academy of Aphasia, Montreal, Quebec, Canada.
- Collins, J. A., Montal, V., Hochberg, D., Quimby, M., Mandelli, M. L., Makris, N., . . . Dickerson, B. C. (2017). Focal temporal pole atrophy and network degeneration in semantic variant primary progressive aphasia. *Brain*, 140, 457–471.
- Corbett, F., Jefferies, E., Eshan, S., & Lambon Ralph, M. A. (2009a). Different impairments of semantic cognition in semantic dementia and semantic aphasia: Evidence from the non-verbal domain. *Brain*, 132, 2593–2608. doi:<https://doi.org/10.1093/brain/awp146>
- Corbett, F., Jefferies, E., & Lambon Ralph, M. A. (2009b). Exploring multimodal semantic control impairments in semantic aphasia: Evidence from naturalistic object use. *Neuropsychologia*, 47, 2721–2731. doi:<https://doi.org/10.1016/j.neuropsychologia.2009.05.020>
- Corbetta, M., & Shulman, G. L. (2002). Control of goal directed and stimulus-driven attention in the brain. *Nat Rev Neuroscience*, 3, 201–215.
- Crawford, J. R., & Howell, D. C. (1998). Comparing an individual's test score against norms derived from small samples. *Clinical Neuropsychologist*, 12, 482–486. doi:<https://doi.org/10.1076/clin.12.4.482.7241>
- Crowther, J. E., & Martin, R. C. (2014). Lexical selection in the semantically blocked cyclic naming task: the role of cognitive control and learning. *Frontiers in Human Neuroscience*, 8, 9. doi:<https://doi.org/10.3389/fnhum.2014.00009>
- de Zubicaray, G. I., Rose, S. E., & McMahon, K. L. (2011). The structure and connectivity of semantic memory in the healthy older adult brain. *NeuroImage*, 54, 1488–1494.
- Desgranges, B., Matuszewski, V., Piolino, P., Chételat, G., Mézenge, F., Landeau, B., . . . Eustache, F. (2007). Anatomical and functional alterations in semantic dementia: A voxel-based MRI and PET

- study. *Neurobiology of Aging*, 28, 1904–1913. doi:<https://doi.org/10.1016/j.neurobiolaging.2006.08.006>
- Dial, H., & Martin, R. C. (2017). Evaluating the relationship between sublexical and lexical processing in speech perception: Evidence from aphasia. *Neuropsychologia*, 96, 192–212.
- Duffau, H., Gatignol, P., Mandonnet, E., Peruzzi, P., Tzourio Mazoyer, & Capelle, L. (2005). New insights into the anatomo-functional connectivity of the semantic system: A study using cortico-subcortical electrostimulations. *Brain*, 128, 797–810.
- Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: Mental programs for intelligent behaviour. *Trends in Cognitive Sciences*, 14, 172–179. doi:<https://doi.org/10.1016/j.tics.2010.01.004>
- Farb, N. A. S., Grady, C. L., Strother, S., Tang-Wai, D. F., Masellis, M., Black, S., . . . Chow, T. W. (2013). Abnormal network connectivity in frontotemporal dementia: Evidence for prefrontal isolation. *Cortex*, 49, 1856–1873.
- Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76, 378–382.
- Forseth, K. J., Kadipasaoglu, C. M., Conner, C. R., Hickok, G., Knight, R. T., & Tandon, N. (2018). A lexical semantic hub for heteromodal naming in middle fusiform gyrus. *Brain*, 141, 2112–2126. doi:<https://doi.org/10.1093/brain/awy120>
- Francis, W. S. (2018). Shared core meanings and shared associations in bilingual semantic memory: Evidence from research on implicit memory. *International Journal of Bilingualism*. Advance online publication. doi:<https://doi.org/10.1177/1367006918814375>
- Gainotti, G., Miceli, G., Silveri, M. C., & Villa, G. (1982). Some anatomo-clinical aspects of phonemic and semantic comprehension disorders in aphasia. *Acta Neurologica Scandinavica*, 66, 652–665. doi:<https://doi.org/10.1111/j.1600-0404.1982.tb04530.x>
- Galton, C. J., Patterson, K., Graham, K., Lambon Ralph, M. A., Williams, G., Antoun, N., . . . Hodges, J. R. (2001). Differing patterns of temporal atrophy in Alzheimer's disease and semantic dementia. *Neurology*, 57, 216–225.
- Garrard, P., & Carroll, E. (2006). Lost in semantic space: A multi-modal, non-verbal assessment of feature knowledge in semantic dementia. *Brain*, 129, 1152–1163. doi:<https://doi.org/10.1093/brain/awl069>
- George, A., & Mathuranath, P. S. (2005). Primary progressive aphasia: A comparative study of progressive nonfluent aphasia and semantic dementia. *Neurology India*, 53, 182–186.
- Goodglass, H., & Kaplan, E. (1983). Boston Diagnostic Aphasia Exam. Philadelphia, PA: Lea and Febiger.
- Gorno-Tempini, M. L., Dronkers, N. F., Rankin, K. P., Ogar, J. M., Phengrasamy, L., Rosen, H. J., . . . Miller, B. L. (2004). Cognition and anatomy in three variants of primary progressive aphasia. *Annals of Neurology*, 55, 335–346. doi:<https://doi.org/10.1002/ana.10825>
- Gorno-Tempini, M. L., Hillis, A. E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S. F., . . . Grossman, M. (2011). Classification of primary progressive aphasia and its variants. *Neurology*, 76, 1006–1014.
- Hahn, B., Ross, T. J., & Stein, E. A. (2006). Neuroanatomical dissociation between bottom-up and top-down processes of visuospatial selective attention. *NeuroImage*, 32, 842–853.
- Hamilton, A. C., & Martin, R. C. (2005). Dissociations among tasks involving inhibition: A single case study. *Cognitive, Affective, & Behavioral Neuroscience*, 5, 1–13. doi:<https://doi.org/10.3758/CABN.5.1.1>
- Hamilton, A. C., & Martin, R. C. (2010). Inferring semantic organization from refractory access dysphasia: Further replication in the domains of geography and proper nouns but not concrete and abstract concepts. *Cognitive Neuropsychology*, 27, 614–635. doi:<https://doi.org/10.1080/02643294.2011.609541>
- Harvey, D. Y., & Schnur, T. T. (2015). Distinct loci of lexical and semantic access deficits in aphasia: Evidence from voxel-based lesion-symptom mapping and diffusion tensor imaging. *Cortex*, 67, 37–58. doi:<https://doi.org/10.1016/j.cortex.2015.03.004>
- Harvey, D. Y., Wei, T., Ellmore, T. M., Hamilton, A. C., & Schnur, T. T. (2013). Neuropsychological evidence for the functional role of the uncinate fasciculus in semantic control. *Neuropsychologia*, 51, 789–801.
- Hodges, J. R., Graham, N., & Patterson, K. (1995). Charting the progression in semantic dementia: Implications for the organization of semantic memory. *Memory*, 3, 463–495.
- Hodges, J. R., & Patterson, K. (2007). Semantic dementia: a unique clinicopathological syndrome. *Lancet Neurology*, 6, 1004–1014. doi:[https://doi.org/10.1016/S1474-4422\(07\)70266-1](https://doi.org/10.1016/S1474-4422(07)70266-1)
- Hodges, J. R., Patterson, K., Oxbury, S., & Furnell, E. (1992). Semantic dementia: Progressive fluent aphasia with temporal lobe atrophy. *Brain*, 115, 1783–1806. doi:<https://doi.org/10.1093/brain/115.6.1783>
- Hoffman, P., Jefferies, E., Haffey, A., Littlejohns, T., & Lambon Ralph, M. A. (2013a). Domain-specific control of semantic cognition: A dissociation within patients with semantic working memory deficits. *Aphasiology*, 27, 740–764. doi:<https://doi.org/10.1080/02687038.2012.751578>
- Hoffman, P., Jefferies, E., & Lambon Ralph, M. A. (2011a). Explaining semantic short-term memory deficits: evidence for the critical role of semantic control. *Neuropsychologia*, 49, 368–381. doi:<https://doi.org/10.1016/j.neuropsychologia.2010.12.034>
- Hoffman, P., Lambon Ralph, M. A., & Rogers, T. T. (2013b). Semantic diversity: A measure of semantic ambiguity based on variability in the contextual usage of words. *Behavior Research Methods*, 45, 718–730. doi:<https://doi.org/10.3758/s13428-012-0278-x>
- Hoffman, P., Rogers, T., & Lambon Ralph, M. A. (2011b). Semantic diversity accounts for the “missing” word frequency effect in stroke aphasia: Insights using a novel method to quantify contextual variability in meaning. *Journal of Cognitive Neuroscience*, 23, 2432–2446. doi:<https://doi.org/10.1162/jocn.2011.21614>
- Hull, R., Martin, R. C., Beier, M. E., Lane, D., & Hamilton, A. C. (2008). Executive function in older adults: A structural equation modeling approach. *Neuropsychology*, 22, 508–522. doi:<https://doi.org/10.1037/0894-4105.22.4.508>
- International Organization for Standardization. (2000). Acoustics—Statistical distribution of hearing thresholds as a function of age (Technical Standard ISO7029). Geneva, Switzerland: ISO.
- Jefferies, E., Baker, S. S., Doran, M., & Lambon Ralph, M. A. (2007). Refractory effects in stroke aphasia: A consequence of poor semantic control. *Neuropsychologia*, 45, 1065–1079. doi:<https://doi.org/10.1016/j.neuropsychologia.2006.09.009>
- Jefferies, E., Hoffman, P., Jones, R. W., & Lambon Ralph, M. A. (2008a). The impact of semantic impairment on verbal short-term memory in stroke aphasia and semantic dementia: A comparative study. *Journal of Memory and Language*, 58, 66–87. doi:<https://doi.org/10.1016/j.jml.2007.06.004>
- Jefferies, E., & Lambon Ralph, M. A. (2006). Semantic impairment in stroke aphasia vs. semantic dementia: A case-series comparison. *Brain*, 129, 2132–2147. doi:<https://doi.org/10.1093/brain/awl1153>
- Jefferies, E., Patterson, K., Jones, R. W., & Lambon Ralph, M. A. (2009). Comprehension of concrete and abstract words in semantic dementia. *Neuropsychology*, 23, 492–499. doi:<https://doi.org/10.1037/a0015452>
- Jefferies, E., Patterson, K., & Lambon Ralph, M. A. (2008b). Deficits of knowledge versus executive control in semantic cognition: Insights from cued naming. *Neuropsychologia*, 46, 649–658. doi:<https://doi.org/10.1016/j.neuropsychologia.2007.09.007>
- Jefferies, E., Rogers, T. T., Hopper, S., & Lambon Ralph, M. A. (2010). “Pre-semantic” cognition revisited: Critical differences between semantic aphasia and semantic dementia. *Neuropsychologia*, 48, 248–261. doi:<https://doi.org/10.1016/j.neuropsychologia.2009.09.011>

- Kay, J., Lesser, R., & Coltheart, M. (1992). PALPA: Psycholinguistic Assessments of Language Processing in Aphasia. Hove, UK: Psychology Press.
- Kertesz, A. (1982). Western Aphasia Battery. New York, NY: Grune & Stratton.
- Kimberg, D. Y., D'Esposito, M., & Farah, M. J. (1997). Effects of bromocriptine on human subjects depend on working memory capacity. *NeuroReport*, 6, 3581–3585. doi:<https://doi.org/10.1097/00001756-199711100-00032>
- Kumfor, F., Landin-Romero, R., Devenney, E., Hutchings, R., Grasso, R., Hodges, J. R., & Piguet, O. (2016). On the right side? A longitudinal study of left- versus right-lateralized semantic dementia. *Brain*, 139, 986–998. doi:<https://doi.org/10.1093/brain/awv387>
- Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, 44, 978–990. doi:<https://doi.org/10.3758/s13428-012-0210-4>
- Lambon Ralph, M. A., Graham, K. S., & Patterson, K. (1999). Is a picture worth a thousand words? Evidence from concept definitions by patients with semantic dementia. *Brain and Language*, 70, 309–335.
- Lambon Ralph, M. A., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*, 18, 42–55. doi:<https://doi.org/10.1038/nrn.2016.150>
- Lambon Ralph, M. A., & Patterson, K. (2008). Generalization and differentiation in semantic memory: Insights from semantic dementia. *Annals of the New York Academy of Sciences*, 1124, 61–76.
- Lhermitte, F. (1983). “Utilization behavior” and its relation to lesions of the frontal lobes. *Brain*, 106, 237–255. doi:<https://doi.org/10.1093/brain/106.2.237>
- Libon, D. J., Xie, S. X., Moore, P., Farmer, J., Antani, S., McCawley, G., . . . Grossman, M. (2007). Patterns of neuropsychological impairment in frontotemporal dementia. *Neurology*, 68, 369–375. doi:<https://doi.org/10.1212/01.wnl.0000252820.81313.9b>
- Macoir, J., Lavoie, M., Laforce, R., Brambati, S. M., & Wilson, M. A. (2017). Dysexecutive symptoms in primary progressive aphasia: Beyond diagnostic criteria. *Journal of Geriatric Psychology and Neurology*, 30, 151–161.
- Martin, A., & Chao, L. L. (2001). Semantic memory and the brain: Structure and processes. *Current Opinion in Neurobiology*, 11, 194–201.
- Martin, R. C., & Breedin, S. D. (1992). Dissociations between speech perception and phonological short-term memory deficits. *Cognitive Neuropsychology*, 9, 509–534.
- Martin, R. C., Lesch, M. F., & Bartha, M. C. (1999). Independence of input and output phonology in word processing and short-term memory. *Journal of Memory and Language*, 41, 3–29. doi:<https://doi.org/10.1006/jmla.1999.2637>
- McCarthy, R. A., & Kartsounis, L. D. (2000). Wobbly words: refractory anomia with preserved semantics. *Neurocase*, 6, 487–497. doi:<https://doi.org/10.1080/13554790008402719>
- Mion, M., Patterson, K., Acosta-Cabrero, J., Pengas, G., Izquierdo-Garcia, D., Hong, Y. T., . . . Nestor, P. J. (2010). What the left and right anterior fusiform gyri tell us about semantic memory. *Brain*, 133, 3256–3268. doi:<https://doi.org/10.1093/brain/awq272>
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, 41, 49–100. doi:<https://doi.org/10.1006/cogp.1999.0734>
- Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, 7, 134–140. doi:[https://doi.org/10.1016/S1364-6613\(03\)00028-7](https://doi.org/10.1016/S1364-6613(03)00028-7)
- Mummery, C. J., Patterson, K., Price, C. J., Ashburner, J., Frackowiak, R. S., & Hodges, J. R. (2000). A voxel-based morphometry study of semantic dementia: Relationship between temporal lobe atrophy and semantic memory. *Annals of Neurology*, 47, 36–45.
- Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms [Database]. Retrieved from <http://w3.usf.edu/FreeAssociation/>
- Noonan, K. A., Jefferies, E., Corbett, F., & Lambon Ralph, M. A. (2010). Elucidating the nature of deregulated semantic cognition in semantic aphasia: Evidence for the roles of prefrontal and temporoparietal cortices. *Journal of Cognitive Neuroscience*, 22, 1597–1613. doi:<https://doi.org/10.1162/jocn.2009.21289>
- Noonan, K. A., Jefferies, E., Garrard, P., Eshan, S., & Lambon Ralph, M. A. (2013a). Demonstrating the qualitative differences between semantic aphasia and semantic dementia: A novel exploration of non-verbal semantic processing. *Behavioural Neurology*, 26, 7–20. doi:<https://doi.org/10.1155/2013/941542>
- Noonan, K. A., Jefferies, E., Visser, M., & Lambon Ralph, M. A. (2013b). Going beyond inferior prefrontal involvement in semantic control: evidence for the additional contribution of dorsal angular gyrus and posterior middle temporal cortex. *Journal of Cognitive Neuroscience*, 25, 1824–1850. doi:https://doi.org/10.1162/jocn_a_00442
- Nunnally, J. C., & Bernstein, I. H. (1994). Psychometric theory (3rd ed.). New York, NY: McGraw-Hill.
- Patterson, K., & Erzinçlioğlu, S. W. (2008). Drawing as a “window” on deteriorating conceptual knowledge in neurodegenerative disease. In C. Lange-Küttner & A. Vintner (Eds.), *Drawing and the non-verbal mind: A life-span perspective* (pp. 281–304). New York, NY: Cambridge University Press.
- Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience*, 8, 976–987. doi:<https://doi.org/10.1038/nrn2277>
- Perry, R. J., & Hodges, J. R. (2000). Differentiating frontal and temporal variant frontotemporal dementia from Alzheimer's disease. *Neurology*, 54, 2277–2284.
- Rapp, B., & Caramazza, A. (1993). On the distinction between deficits of access and deficits of storage: A question of theory. *Cognitive Neuropsychology*, 10, 113–141. doi:<https://doi.org/10.1080/02643299308253458>
- Raven, J. C. (1962). Coloured progressive matrices sets A, AB, B. London, UK: H. K. Lewis.
- Raven, J. C. (1965). Advanced progressive matrices sets I & II. London, UK: H. K. Lewis.
- Roach, A., Schwartz, M. F., Martin, N., Grewel, R. S., & Brecher, A. (1996). The Philadelphia Naming Test: Scoring and rationale. *Clinical Aphasiology*, 24, 121–133.
- Robson, H., Sage, K., & Lambon Ralph, M. A. (2012). Wernicke's aphasia reflects a combination of acoustic-phonological and semantic control deficits: A case-series comparison of Wernicke's aphasia, semantic dementia and semantic aphasia. *Neuropsychologia*, 50, 266–275. doi:<https://doi.org/10.1016/j.neuropsychologia.2011.11.021>
- Rogers, T. T., Lambon Ralph, M. A., Garrard, P., Bozeat, S., McClelland, J. L., Hodges, J. R., & Patterson, K. (2004). Structure and deterioration of semantic memory: A neuropsychological computational investigation. *Psychological Review*, 111, 205–235. doi:<https://doi.org/10.1037/0033-295X.111.1.205>
- Rosen, H. J., Kramer, J. H., Gorno-Tempni, M. L., Schuff, N., Weiner, M., & Miller, B. L. (2002). Patterns of cerebral atrophy in primary progressive aphasia. *American Journal of Geriatric Psychiatry*, 10, 89–97.
- Ruff, C. C., Blankenburg, F., Bjoertomt, O., Bestmann, S., Freeman, E., Haynes, J.-D., . . . Driver, J. (2006). Concurrent TMS-fMRI and psychophysics reveal frontal influences on human retinotopic visual cortex. *Current Biology*, 16, 1479–1488. doi:<https://doi.org/10.1016/j.cub.2006.06.057>
- Schnur, T. T., Schwartz, M. F., Brecher, A., & Hodgson, C. (2006). Semantic interference during blocked-cyclic naming: evidence from

- aphasia. *Journal of Memory and Language*, 54, 199–227. doi:<https://doi.org/10.1016/j.jml.2005.10.002>
- Schnur, T. T., Schwartz, M. F., Kimberg, D. Y., Hirshorn, E., Coslett, H. B., & Thompson-Schill, S. L. (2009). Localizing interference during naming: convergent neuroimaging and neuropsychological evidence for the function of Broca's area. *Proceedings of the National Academy of Sciences*, 106, 322–327. doi:<https://doi.org/10.1073/pnas.0805874106>
- Schwartz, M. F., Dell, G. S., Martin, N., Gahl, S., & Sobel, P. (2006). A case-series test of the interactive two-step model of lexical access: Evidence from picture naming. *Journal of Memory and Language*, 54, 228–264. doi:<https://doi.org/10.1016/j.jml.2005.10.001>
- Schwartz, M. F., Kimberg, D. Y., Walker, G. M., Faseyitan, O., Brecher, A., Dell, G. S., & Coslett, H. B. (2009). Anterior temporal involvement in semantic word retrieval: Voxel-based lesion-symptom mapping evidence from aphasia. *Brain*, 132, 3411–3427. doi:<https://doi.org/10.1093/brain/awp284>
- Scott, G. G., Keitel, A., Becirspahic, M., Yao, B., & Sereno, S. C. (2019). The Glasgow Norms: Ratings of 5,500 words on nine scales. *Behavior Research Methods*, 51, 1258–1270. doi:<https://doi.org/10.3758/s13428-018-1099-3>
- Sikora, K., Roelofs, A., Hermans, D., & Knoors, H. (2016). Executive control in spoken noun-phrase production: Contributions of updating, inhibiting, and shifting. *Quarterly Journal of Experimental Psychology*, 69, 1719–1740. doi:<https://doi.org/10.1080/17470218.2015.1093007>
- Simmons, W. K., Reddish, M., Bellgowan, P., & Martin, A. (2010). The selectivity and functional connectivity of the anterior temporal lobes. *Cerebral Cortex*, 20, 813–825. doi:<https://doi.org/10.1093/cercor/bhp149>
- Slevc, L. R., Martin, R. C., Hamilton, A. C., & Joanisse, M. F. (2011). Speech perception, rapid temporal processing, and the left hemisphere: A case study of unilateral pure word deafness. *Neuropsychologia*, 49, 216–230.
- Snyder, H. R., Banich, M. T., & Munakata, Y. (2011). Choosing our words: Retrieval and selection processes recruit shared neural substrates in left ventrolateral prefrontal cortex. *Journal of Cognitive Neuroscience*, 23, 3470–3482.
- Sternberg, S. (1998). Discovering mental processing stages: The method of additive factors. In D. Scarborough & S. Sternberg (Eds.), *Invitation to cognitive science, Vol. 4: Methods, models, and conceptual issues* (pp. 703–863). MIT Press.
- Stuss, D. T., & Alexander, M. P. (2007). Is there a dysexecutive syndrome? *Philosophical Transactions of the Royal Society B*, 362, 901–915. doi:<https://doi.org/10.1098/rstb.2007.2096>
- Tamber-Rosenau, B. J., Esterman, M., Chiu, Y. C., & Yantis, S. (2011). Cortical mechanisms of cognitive control for shifting attention in vision and working memory. *Journal of Cognitive Neuroscience*, 23, 2905–2919.
- Thompson, H. E., Almaghyuli, A., Noonan, K. A., Ohr, B., Lambon Ralph, M. A., & Jefferies, E. (2018). The contribution of executive control to semantic cognition: Convergent evidence from semantic aphasia and executive dysfunction. *Journal of Neuropsychology*, 12, 1–29.
- 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.
- Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). Mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, 45, 1–67. <http://www.jstatsoft.org/v45/i03/> doi:<https://doi.org/10.18637/jss.v045.i03>
- Varney, N. R. (1984). Phonemic imperception in aphasia. *Brain and Language*, 21, 85–94.
- 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. doi:[https://doi.org/10.1016/S0896-6273\(01\)00359-2](https://doi.org/10.1016/S0896-6273(01)00359-2)
- Walker, G. M., Schwartz, M. F., Kimberg, D. Y., Faseyitan, O., Brecher, A., Dell, G. S., & Coslett, H. B. (2011). Support for anterior temporal involvement in semantic error production in aphasia: New evidence from VLSM. *Brain and Language*, 117, 110–122.
- Warrington, E. K. (1975). The selective impairment of semantic memory. *Quarterly Journal of Experimental Psychology*, 27, 635–657. doi:<https://doi.org/10.1080/14640747508400525>
- Warrington, E. K., & Cipolotti, L. (1996). Word comprehension: The distinction between refractory and storage impairments. *Brain*, 119, 611–625. doi:<https://doi.org/10.1093/brain/119.2.611>
- Warrington, E. K., & McCarthy, R. (1983). Category specific access dysphasia. *Brain*, 106, 859–878. doi:<https://doi.org/10.1093/brain/106.4.859>
- Warrington, E. K., & McCarthy, R. A. (1987). Categories of knowledge: Further fractionations and an attempted integration. *Brain*, 110, 1273–1296. doi:<https://doi.org/10.1093/brain/110.5.1273>
- Warrington, E. K., & Shallice, T. (1979). Semantic access dyslexia. *Brain*, 102, 43–63. doi:<https://doi.org/10.1093/brain/102.1.43>
- Warrington, E. K., & Shallice, T. (1984). Category specific semantic impairments. *Brain*, 107, 829–854. doi:<https://doi.org/10.1093/brain/107.3.829>
- Wechsler, D. (1981). Wechsler Memory Scale—Revised. New York, NY: Psychological Corp.
- Whitney, C., Kirk, M., O'Sullivan, J., Lambon Ralph, M. A., & Jefferies, E. (2011). The neural organization of semantic control: TMS evidence for a distributed network in left inferior frontal and posterior middle temporal gyrus. *Cerebral Cortex*, 21, 1066–1075. doi:<https://doi.org/10.1093/cercor/bhq180>
- Whitwell, J. L., Avula, R., Senjem, M. L., Kantarci, K., Weigand, S. D., Samikoglu, A., . . . Jack, C. R. (2010). Gray and white matter water diffusion in syndromic variants of frontotemporal dementia. *Neurology*, 74, 1279–1287.
- Winer, B. J. (1971). Statistical principles in experimental design (2nd ed.). New York, NY: McGraw-Hill.
- Wollams, A. M., Cooper-Pye, E., Hodges, J. R., & Patterson, K. (2008). Anomia: A doubly typical signature of semantic dementia. *Neuropsychologia*, 46, 2503–2514. doi:<https://doi.org/10.1016/j.neuropsychologia.2008.04.005>
- Zhao, Q., Guo, Q., Li, F., Zhou, Y., Wang, B., Hong, Z. (2013). The shape trail test: Application of a new variant of the trail-making test. *PLoS ONE*, 8, e57333 doi:<https://doi.org/10.1371/journal.pone.0057333>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.