Neuropsychology
2004, Vol. 18, No. 4, 621–628

0894-4105/04/$12.00 DOI: 10.1037/0894-4105.18.4.621

A Meta-Analytic Review of Verbal Fluency Performance in Patients With Traumatic Brain Injury

Julie D. Henry and John R. Crawford
King's College, University of Aberdeen

A meta-analysis of 30 studies with 1,269 participants was conducted to investigate the sensitivity of tests of verbal fluency to the presence of traumatic brain injury (TBI). As has been found for patients with focal frontal lobe injuries (but not for patients with focal temporal lobe lesions), TBI patients were comparably impaired on tests of phonemic and semantic fluency. The phonemic fluency deficit could not be accounted for by patients' level of premorbid or current verbal IQ and was also substantially (although not significantly) in excess of the deficit on a measure of psychomotor speed. Phonemic fluency was also significantly more sensitive to the presence of TBI than was the Wisconsin Card Sorting Test (R. K. Heaton, 1981).

Traumatic brain injury (TBI) is the most common cause of brain damage and can be associated with either closed or open head injuries (CHI and OHI, respectively). However, for most patients (over 90%) the skull is not penetrated, and such injuries are usually typified by focal damage superimposed on more diffuse white matter and brain stem damage. Although it has long been recognized that short-term cognitive deficits are associated with CHI, it is increasingly accepted that even in cases of mild CHI, at least some patients have lasting neuropsychological problems. Indeed, Stuss et al. (1985) reported that CHI patients who had attained a good recovery according to a clinical outcome measure (the Glasgow Outcome Scale) were impaired on a number of cognitive measures relative to control participants matched for age, education, and general intelligence. Thus, understanding the neuropsychological sequelae of TBI may have important long-term implications for the development of effective rehabilitative techniques.

Executive processes are of particular importance for rehabilitation because they are presumed to underlie many of the complex behaviors necessary for both functional independence and social integration. Measures of executive functioning and verbal memory, for instance, have been found to be stronger predictors of functional outcome than have measures of functional motor ability (Hanks, Rapport, Millis, & Deshpande, 1999). Although TBI is often associated with deficits on many measures of cognitive functioning, it has been suggested that in addition to episodic memory being impaired, executive functioning is particularly impaired (Stuss & Gow, 1992). Neuropathological evidence is consistent with this possibility, as focal contusions are most frequently found in frontal and temporal regions (see Levin & Kraus, 1994; Levin, Williams, Eisenberg, High, & Guinto, 1992). Because there is a great deal of evidence that executive processes rely heavily on the intact functions of frontal structures, the presence of frontal damage would therefore suggest that deficits in this aspect of cognition should be especially marked.

In addition, TBI is usually characterized by diffuse white matter pathology that is believed to be particularly associated with disturbances of the executive control system (Stuss & Gow, 1992). Indeed, Beers, Goldstein, and Katz (1994) have argued that "diffuse brain injury, regardless of severity, appears to cause a reduction in the speed, efficiency, and integration of mental processes" (p. 316). Because executive functioning is thought to be responsible not for basic cognitive processes but for the complex systems that integrate these capacities, it will presumably be disproportionately affected by diffuse injury.

Stuss and Gow (1992) have drawn parallels between focal frontal injuries and TBI and have argued that executive dysfunction is the most prominent disturbance associated with both. They have pointed out that there are many similarities between the two disorders, including deficits in anticipation and attentional processes. Moreover, many studies have reported that patients with TBI were impaired on tests designed to capture executive dysfunction, such as phonemic fluency (Cooke & Kausler, 1995; McDowell, Whyte, & D'Esposito, 1997), semantic fluency (Lannoo, Colardyn, De Deyne, et al., 1998; Raskin & Rearick, 1996), the Wisconsin Card Sorting Test (WCST; Cockburn, 1995; Stuss et al., 1985), and the Stroop interference test (Lannoo, Colardyn, De Deyne, et al., 1998; McDowell et al., 1997). However, as noted previously, patients with TBI may be impaired on virtually all measures of cognitive function, and this includes tests presumed to make only minimal demands on executive processes. If TBI is to be validly regarded as a disorder characterized by disproportionate executive decline, it is necessary to demonstrate that deficits on executive measures are greater than deficits on tasks presumed to be relatively insensitive to frontal damage (see Crawford, Blackmore, Lamb, & Simpson, 2000; Laws, 1999; Miller, 1984).

In particular, episodic memory deficits are also often associated with head injury. However, it has been suggested that the episodic memory deficit in TBI may reflect executive dysfunction (Raskin & Rearick, 1996). Working memory invokes strategic, supervisory

Julie D. Henry and John R. Crawford, Department of Psychology, King's College, University of Aberdeen, Aberdeen, Scotland.

Additional materials are on the Web at http://dx.doi.org/10.1037/0894-4105.18.4.621.supp.

Correspondence concerning this article should be addressed to Julie D. Henry, Department of Psychology, King's College, University of Aberdeen, Aberdeen AB24 3HN, Scotland. E-mail: j.d.henry@abdn.ac.uk. After November 1, 2004, correspondence should be addressed to Julie D. Henry, School of Psychology, University of New South Wales, New South Wales, Australia. E-mail: julie.henry@unsw.edu.au

control systems, and Baddeley (1990) has drawn parallels between the "supervisory attentional system" that is presumed to be responsible for higher order cognitive functions and the "central executive" system of working memory. Using fMRI, McAllister et al. (1999) found that although the performance of patients with mild TBI on a working memory task was similar to that of healthy control participants, as processing load altered, their respective activation patterns differed. These researchers concluded that memory difficulties may reflect abnormalities in the activation or modulation of working memory processing resources. However, although there is strong evidence that executive deficits account for memory problems in other neurological conditions (see Crawford et al., 2000), it is unlikely that executive dysfunction can account for all of the episodic memory dysfunction associated with TBI, as deficits for the latter have been found to exceed those associated with executive functioning (Levin, Amparo, et al., 1987).

Moreover, the possibility remains that executive functioning is not differentially impaired per se. Mathias and Coats (1999) found that, with the exception of phonemic fluency, when education was controlled for, no deficit was found on any cognitive measure, including the WCST. Cockburn (1995) also reported that whereas phonemic fluency and the WCST differentiated patients from control participants, the Tower of London test did not. Other studies have failed to find significant deficits on any executive measure (Levin, Lippold, et al., 1987; McAllister et al., 1999). Thus, although patients with TBI often perform poorly on executive measures, exceptions have been reported, and it remains unclear whether executive functioning—or certain aspects of executive functioning—are differentially impaired.

In an attempt to resolve whether TBI patients present with differential executive deficits, researchers have studied verbal fluency performance extensively. Measures of verbal fluency require time-restricted generation of multiple response alternatives that are based on phonemic or semantic criteria (phonemic and semantic fluency, respectively). Verbal fluency tests and, in particular, phonemic fluency tests are commonly used to assess executive dysfunction, as they require efficient organization of verbal retrieval and recall as well as self-monitoring aspects of cognition, effortful self-initiation, and inhibition of responses when appropriate. However, although some studies have failed to find significant deficits on either type of measure (Beers et al., 1994; Wertz, Dronkers, & Shubitowski, 1986), others have reported impaired performance on both types (Lannoo, Colardyn, De Deyne, et al., 1998; Raskin & Rearick, 1996).

The comparability of phonemic and semantic fluency deficits in relation to TBI also remains unclear. Although the standard scores for semantic and phonemic fluency measures did not differ significantly for patients with mild to moderate CHI, Goldstein et al. (1996) found that there was a tendency for semantic fluency to be more impaired than phonemic fluency. However, in a series of previous investigations, they had found that "young patients recall significantly more words that belong to categories such as clothing and fruits than words that do not share categorical relationships" (Goldstein et al., 1996, p. 148).

Comparison of the relative magnitude of deficits on phonemic and semantic fluency is particularly important, because this may be used to draw inferences about the prominence of executive dysfunction and semantic memory dysfunction, respectively, and, when in the context of performance on other cognitive measures, whether these deficits qualify as differential deficits. As noted previously, frontal structures are particularly associated with executive functioning (Stuss & Benson, 1986). In a previous meta-analytic review of the verbal fluency performance of patients with focal cortical lesions (Henry & Crawford, 2004), we found that frontal injuries were associated with equivalent phonemic and semantic fluency deficits (rs = .52 and .54, respectively), suggesting that the two types of fluency place comparable demands on executive processes. However, semantic fluency was more impaired following focal temporal damage (r = .61), and the deficit was significantly larger than the corresponding phonemic fluency deficit (r = .44). Because there is a great deal of evidence that temporal structures are the neural substrates particularly responsible for semantic memory, this finding was presumed to reflect the greater reliance of semantic fluency on the integrity of semantic memory.

However, it is difficult to draw clear conclusions about the relative prominence of phonemic and semantic fluency deficits from the existing TBI literature because of inconsistencies between individual studies. Apparent discrepancies may reflect the influence of substantive factors, such as severity of injury, whereas Binder, Robling, and Larrabee (1997) have argued that it is important to distinguish between clinical and prospective studies, particularly in cases of mild TBI. They defined clinical studies as those that assess symptomatic patients with TBI who are referred for assessment because of cognitive complaints; in prospective studies participants are selected simply on the basis of having had a head injury. Discrepancies may also reflect differences in the duration of time between injury and assessment. When patients are assessed shortly after injury, diffuse brain swelling may obscure the specific effects of focal lesions (Levin et al., 1992).

To our knowledge, the current study is the first to apply meta-analytic techniques to compare performance on phonemic and semantic fluency in TBI. One of the most important advantages of this methodology is that corrections can be implemented for sampling error. Thus, it was possible to assess whether discrepancies between studies reflect the influence of substantive factors or artifactual variance. In addition, with the use of meta-analysis, an effect's generalizability can be subjected to a level of scrutiny not possible in a single study and examined with a level of objectivity and methodological consistency that is difficult to achieve in nonquantitative reviews (Stanley, 2001). Thus, meta-analysis enables highly reliable, robust effect size estimates to be calculated for the corresponding parameters of interest.

Our first aim was to derive effect size estimates for phonemic and semantic fluency for patients with TBI relative to healthy controls to provide reliable estimates of these parameters and their associated confidence intervals (CIs). Our second aim was to compare the magnitude of phonemic versus semantic fluency deficits to allow a quantitative assessment of the relative prominence of executive and semantic memory dysfunction in TBI.

Another issue we wish to examine relates to whether verbal fluency deficits in TBI qualify as differential deficits, and thus, our third aim was to derive mean effects for other cognitive measures and compare these with the corresponding phonemic fluency deficits (insufficient studies assessed semantic fluency to permit these comparisons).

Of particular importance was to compare the deficit for current verbal intelligence (VIQ; Wechsler, 1955, 1981) with the corresponding phonemic fluency deficit, because phonemic fluency imposes substantial demands on verbal abilities and was originally developed as a measure of VIQ. Premorbid IQ, as estimated by the National Adult Reading Test (Nelson, 1982) and the Reading subtest of the Wide Range Achievement Test (Jastak & Wilkinson, 1984), was also included to address the possibility that if a phonemic fluency deficit is present, it reflects the fact that TBI patients have not been successfully matched to the control participants for premorbid ability. It is also important to note that TBI is often associated with generalized slowing that may affect virtually all measures of cognition but, in particular, tasks that impose substantial demands on cognitive speed. Therefore, the mean effect size for the Trail Making Test—Part A (TMT–A; Reitan, 1990), a widely used measure of psychomotor speed, was quantified to assess whether deficits on phonemic fluency reflect generalized slowing rather than executive dysfunction.

To permit a comparison of the magnitude of deficits on tests of episodic memory relative to phonemic fluency, we recorded the total and delayed scores for Verbal Learning and Delayed Recall, respectively, from the Rey Auditory Verbal Learning Test (Lezak, 1995) and the California Verbal Learning Test (Delis, Kramer, Kaplan, & Ober, 1987), with Delayed Recall from the Selective Reminding Test (Buschke & Fuld, 1974) also permitted to contribute to this latter construct. For executive functioning, performance on WCST–Categories Completed and WCST–Perseverative Errors (WCST–CC and WCST–PE, respectively; Heaton, 1981) was recorded.

Method

Sample of Studies

A manual search of most issues of *Neuropsychology; The Journal of the International Neuropsychological Society; The Clinical Neuropsychologist*; *Neuropsychiatry, Neuropsychology and Behavioural Neurology; Journal of Neuropsychiatry and Clinical Neurosciences; Neuropsychologia;* and the *Journal of Clinical and Experimental Neuropsychology* was conducted. A search involving the *Web of Science*, *PsycLit* [CD-ROM], and *Science Direct* databases was also undertaken, with the following terms as search parameters: *letter fluency, FAS, semantic fluency, category fluency, controlled oral word association, COWA(T), word fluency, verbal fluency, oral fluency, phonemic fluency, executive test,* and *frontal test.* The search was completed in October 2002.

The inclusion criteria for the studies were as follows: (a) The patient group had to consist entirely of adults with TBI, (b) the healthy control group had to be free of neurological or psychiatric disease, and (c) a measure of phonemic or semantic fluency had to be used. Effect size estimates for current VIQ, premorbid IQ, the TMT–A, Verbal Learning, Delayed Recall, WCST–CC, and WCST–PE were derived from studies that also reported verbal fluency results. The studies must also have (d) presented precise statistics convertible to effect size *r,* (e) been published in a journal, and (f) been written in English.

Statistical Analysis

Meta-analysis is a rigorous, quantitative alternative to the traditional review process because it involves statistical integration of results. The basis of this methodology is the *effect size,* a standardized statistic that quantifies the magnitude of an effect. In the present study, the effect size *r* was used, which corresponds to the degree of correlation between group membership (i.e., presence or absence of TBI) and performance on the cognitive measure of interest. For each construct, effects were pooled to derive an estimate of the mean, with each effect weighted for sample size to correct for sampling error. To cumulate effect sizes, we selected the random effects meta-analytic model in preference to the more commonly used fixed effects model because it yields more generalizable parameter estimates. The National Research Council (1992) has argued that the fixed effects model should be the exception rather than the rule, as it may lead to inappropriately strong conclusions. Thus, although the random effects model is more technically demanding than the fixed effects model, we considered it important to use the random effects model in the present work.

It was also important to test whether the difference in the magnitude of mean effects between, for instance, phonemic and semantic fluency was statistically significant. However, there is no agreed-upon method for statistically comparing mean effects when using the random effects meta-analytic model. A particular difficulty is whether the degrees of freedom (*df*s) in such analyses should be based on the number of participants (*N*) or on the number of studies (*K*). As is discussed, in the present study paired *t* tests were conducted with the more conservative number of studies as the degree of freedom.

Because it has been argued that there may be substantive differences among patients with TBI, such as the severity of injury and the duration of time since injury, for each statistical comparison only studies that assessed both variables of interest were included. For example, although 33 TBI groups were assessed on phonemic fluency and nine groups on semantic fluency, only six TBI groups were assessed on both phonemic and semantic fluency. Thus, when conducting inferential statistics to compare phonemic and semantic fluency, we permitted only these six groups to contribute to the analyses. This effectively controlled for substantive differences between studies when comparing two different cognitive measures, as exactly the same participants were being compared on each of these measures. Because the same participants were compared on each measure, paired *t* tests were used for these statistical comparisons.

Mean effects were also calculated for each of the nonfluency variables identified (current VIQ, premorbid IQ, the TMT–A, Verbal Learning, Delayed Recall, WCST–CC, and WCST–PE) and compared with phonemic fluency. Again, to ensure that TBI severity was controlled for, only studies that assessed both phonemic fluency and the particular nonfluency variable of interest were included in each comparison. Thus, for each comparison we recalculated the mean effect for phonemic fluency using only those studies that assessed both the nonfluency measure of interest as well as phonemic fluency.

To interpret how important a particular effect was in practical terms, we adopted Cohen's (1977) guidelines. These suggest that a correlation of .1 should be regarded as representing a small effect, .3 as medium, and .5 as large. In addition, squares of the effect size multiplied by 100 were also calculated, as these latter quantities represent the percentage of the variance accounted for (PVAF) by group membership (i.e., having had a TBI vs. being a member of the healthy adult population) on a measure of interest. For inferential statistics, we made comparisons using the PVAF by group membership on each of the measures of interest because the difference between effect sizes is nonlinear as the effect size increases, and thus, PVAF is the more appropriate index when comparing variables.

Results

Participant Characteristics

Thirty studies published between 1986 and 2002, in which there were a total of 667 TBI patients and 602 control participants, met the inclusion criteria. Supplemental information, including a complete list of the articles included in this meta-analysis, can be found on the Web at http://dx.doi.org/10.1037/0894-4105.18.4.621.supp. Patients and control participants were closely matched for age

(M = 36.3 years, SD = 13.19, vs. M = 36.4 years, SD = 13.07, respectively) and education (M = 12.5 years, SD = 1.77, vs. M = 12.9 years, SD = 1.74, respectively), but a higher percentage of the patient group was male (67.2% vs. 59.8%, respectively).

Verbal Fluency Performance of Patients With TBI

Table 1 presents estimates of the mean effects for phonemic and semantic fluency and their variability and practical importance. For the mean effects, a positive sign indicates that patients performed more poorly than control participants. To estimate the degree of heterogeneity of the effects contributing to each mean, we also estimated the homogeneity statistic Q and the random effects variance (σ^2_θ), as well as the standard deviation of random effects and the 95% CIs within which random effects could be expected to fall. The homogeneity statistic quantifies within-group heterogeneity (i.e., the degree to which the studies contributing to each respective mean can be regarded as homogeneous). If the homogeneity statistic associated with a mean effect is significant, this suggests that there are substantive differences between the studies contributing to that particular mean. In contrast, a nonsignificant estimate of the homogeneity statistic suggests that once sampling error has been removed, no substantive differences between the studies contributing to the respective mean in question remain (i.e., the null hypothesis of homogeneity of effects cannot be rejected). Both mean effects were significantly different from zero ($p < .01$) and moderate to large in magnitude (rs = .48 and .45 for phonemic and semantic fluency, respectively). For phonemic but not semantic fluency, the homogeneity statistic was significant. To avoid any potential confusion, we note that these effect sizes differ from those presented in the abstract and elsewhere, as the latter effect sizes are based on studies that included both phonemic and semantic fluency, whereas the former effects are based on any studies that included phonemic or semantic fluency.

Phonemic Fluency Deficits Relative to Other Cognitive Deficits

Table 2 presents estimates of the mean effects and their variability and practical importance for semantic fluency, premorbid IQ, current VIQ, TMT–A, WCST–CC, WCST–PE, Verbal Learning, and Delayed Recall; these mean effects are those used for the inferential statistical tests and were thus calculated with data only from those studies that included the particular measure of interest in addition to phonemic fluency. Thus, it can be seen that for each measure, the mean effect for phonemic fluency was recalculated to ensure that comparisons were fair. It should be noted that two outliers were identified in the present study (one for WCST–CC and one for WCST–PE). Although excluding these outliers reduced the heterogeneity observed for the two respective means in question, their exclusion did not alter the pattern of results, and they were therefore retained. It can be seen that for all except three of the mean effects in Table 2, the homogeneity statistic was significant.

In terms of the PVAF by group membership, phonemic and semantic fluency did not differ significantly (rs = .42 vs. .46, respectively), $t(6) = 1.27$, $p = .26$. Turning to the comparison of phonemic fluency with nonfluency measures (as noted earlier, insufficient studies assessed semantic fluency to permit these comparisons), we note that the effect size for phonemic fluency was significantly in excess of both premorbid IQ and current VIQ, $t(7) = 4.20$, $p < .01$, and $t(10) = 3.47$, $p < .01$, respectively; however, although the effect size was substantially in excess of the deficit for the TMT–A (PVAF by group membership = 17.14% vs. 10.43%, respectively), this latter comparison did not attain statistical significance, $t(7) = 1.85$, $p = .11$.

In terms of the PVAF by group membership, the phonemic fluency deficit was significantly larger than the deficits for both the WCST–CC and WCST–PE, $t(8) = 5.05$, $p < .01$, and $t(10) = 5.21$, $p < .01$, respectively, indicating that of these measures, phonemic fluency was the most sensitive to the presence of TBI. Finally, phonemic fluency was more impaired than episodic memory as measured by Verbal Learning and Delayed Recall, although not significantly, $t(8) = 2.06$, $p = .07$, and $t(10) = 0.55$, $p = .59$, respectively.

The PVAF (as well as the 95% CIs of the PVAF) for phonemic and semantic fluency is illustrated in Figure 1, alongside the corresponding values for patients with focal frontal versus focal temporal cortical lesions (focal lesion data are taken from Henry & Crawford, 2004). For the focal lesion analyses, the number of studies ranged from 5 to 32, and the number of participants from 60 to 472. Both patients with TBI and patients with focal frontal lesions were substantially and comparably impaired on phonemic and semantic fluency (i.e., for both groups, the relative prominence of deficits on these two measures was comparable).

Assessing the Potential Presence of Publication Bias

A number of validity threats have been identified that may lead to imprecise conclusions in both nonquantitative and meta-analytic reviews. Of particular concern is the file drawer problem, which refers to the fact that significant results are more likely to be

Table 1
Verbal Fluency Performance of Patients With TBI Relative to Healthy Control Participants

					95% CI of mean							95% CI of mean effects	
Fluency type	M	K	n[a]	SE	Lower	Upper	Z	PVAF	Q	σ^2_θ	SD	Lower	Upper
Phonemic	.48	33	613	.030	.42	.53	16.0**	22.6	64.2**	.014	.117	.25	.70
Semantic	.45	9	171	.047	.36	.55	9.6**	20.5	8.9	.002	.046	.36	.54

Note. TBI = traumatic brain injury; CI = confidence interval; PVAF = percentage of the variance accounted for.
[a] Number of participants in the patient group.
** $p < .001$.

Table 2
Performance on Tests of Phonemic Fluency Relative to Other Cognitive Measures in TBI

Measure	*M*	*K*	n^a	*SE*	95% CI of mean Lower	95% CI of mean Upper	*Z*	PVAF	*Q*	σ^2_θ	*SD*	95% CI of mean effects Lower	95% CI of mean effects Upper	Phonemic fluency M^b
Semantic fluency	.46	6	117	.066	.33	.59	7.0*	21.4	8.1	.010	.100	.27	.66	.42 (*K* = 6)
Premorbid IQ	.25	8	127	.059	.13	.36	4.2*	6.0	3.9	.000	—	—	—	.53 (*K* = 8)
VIQ	.28	11	211	.065	.15	.40	4.2*	7.6	25.7*	.027	.165	−.05	.60	.45 (*K* = 11)
TMT–A	.32	8	138	.087	.15	.49	3.7*	10.4	18.2*	.036	.192	−.03	.72	.41 (*K* = 8)
WCST–CC	.32	9	183	.081	.16	.48	4.0*	10.4	22.8*	.037	.193	−.06	.70	.54 (*K* = 9)
WCST–PE	.30	11	197	.063	.17	.42	4.7*	8.7	19.3*	.020	.143	.02	.58	.53 (*K* = 11)
Verbal Learning	.32	9	211	.060	.20	.43	5.3*	10.0	13.8	.013	.115	.09	.54	.42 (*K* = 9)
Delayed Recall	.48	11	188	.083	.31	.64	5.7*	22.5	57.6*	.059	.242	.00	.95	.50 (*K* = 11)

Note. Dashes indicate that the random effects variance is zero. TBI = traumatic brain injury; CI = confidence interval; PVAF = percentage of the variance accounted for; VIQ = verbal intelligence; TMT–A = Trail Making Test—Part A; WCST–CC = Wisconsin Card Sorting Test—Categories Completed; WCST–PE = Wisconsin Card Sorting Test—Perseverative Errors.
[a] Number of participants in the patient group. [b] For each variable of interest, the mean effect for phonemic fluency was recalculated with data from only those studies that assessed both the variable and phonemic fluency. For example, eight studies assessed both phonemic fluency and premorbid IQ; in addition to calculating a mean effect for premorbid IQ on the basis of these eight studies, the mean effect for phonemic fluency was also recalculated on the basis of only these eight studies (r = .53). Thus, when assessing whether the effect size for phonemic fluency is in excess of that for premorbid IQ, exactly the same participants were tested on each measure, effectively controlling for any substantive differences between studies, such as in level of TBI severity. This is a more rigorous method of comparing performance on different measures at the level of meta-analysis.
* $p < .05$.

published than nonsignificant results. To assess whether this bias posed a threat, we constructed funnel plot diagrams in which sample size was plotted against the corresponding study-level effect. If statistically nonsignificant results were discriminated against, there would have been a relative absence of studies with small sample sizes that reported weak effects. For none of the variables was there evidence of this bias operating.

Discussion

Quantifying Verbal Fluency Deficits in TBI

The presence of TBI was associated with comparable deficits on tests of phonemic and semantic fluency, suggesting that for this patient group the verbal fluency deficit primarily reflects executive dysfunction. This is because, as discussed previously in this article, focal frontal lesions were associated with equivalent deficits on the two types of fluency (Henry & Crawford, 2004), suggesting that the measures implicate executive processes to a comparable degree. However, because semantic fluency was more sensitive to temporal lesions, semantic fluency is relatively more dependent on the integrity of the semantic system. The finding of comparable deficits on these measures is therefore consistent with other evidence that the verbal fluency deficit in TBI predominantly reflects a breakdown in executive control functions.

It is noteworthy that phonemic fluency has been presumed to be relatively more dependent on executive processes than has semantic fluency (Perret, 1974), and this would have predicted greater impairment on phonemic relative to semantic fluency for patients with frontal injuries. It is therefore ironic that although it has been suggested that there are similarities in the cognitive deficits associated with patients with TBI and those with frontal injuries (Stuss & Gow, 1992), this may have been expected to have been demonstrated by greater impairment on fluency as determined by phonemic relative to semantic criteria. In fact, the present results indicate that the pattern of performance for TBI patients does resemble that found following focal frontal lesions, but this is because these patients exhibit comparable deficits on these measures.

However, it is important to point out that in interpreting data of this sort, there is the possibility of identity fallacies. Thus, although we suggest that a similar pattern in phonemic and semantic fluency effect sizes for patients with TBI and patients with frontal lobe

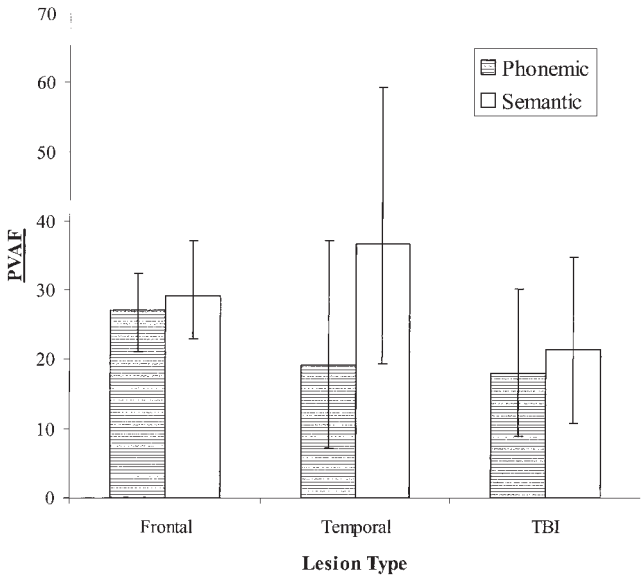

Figure 1. Percentage of the variance accounted for (PVAF) in performance on phonemic versus semantic fluency by the presence of traumatic brain injury (TBI) or focal frontal or focal temporal cortical lesions. Error bars denote the 95% confidence intervals.

damage provides evidence that TBI patients are like patients with frontal lobe damage and, therefore, have greater difficulties with executive functioning than with processes that are reliant on semantic memory, other interpretations of these data are possible. That is, it may be that TBI patients show equivalent impairments in both phonemic and semantic fluency for reasons that are unrelated to executive dysfunction. However, we favor the interpretation that fluency deficits in TBI reflect a breakdown in executive control processes because, as discussed previously, there are many similarities between the two disorders (see Stuss & Gow, 1992), and there is a great deal of evidence that neuropathologically, the frontal lobes are particularly vulnerable in TBI (see Levin & Kraus, 1994; Levin et al., 1992). In addition, as is discussed, failure to match patients and control participants on premorbid IQ, current VIQ, or psychomotor speed is not sufficient to account for the magnitude of the deficit observed.

Phonemic Fluency Deficits Relative to Other Cognitive Deficits

The PVAF by group membership for phonemic and semantic fluency did not differ significantly in the studies that assessed both. The phonemic fluency deficit could not be accounted for by level of premorbid IQ and was also significantly in excess of the deficit for current VIQ. However, although the phonemic fluency deficit was substantially in excess of the deficit for the TMT–A, this comparison did not attain statistical significance. Thus, for phonemic fluency the deficit cannot be explained by a failure to match patients and control participants on premorbid IQ or on current level of VIQ but it may, to at least a certain extent, reflect a reduction in psychomotor speed.

However, group membership accounted for substantially more variance in phonemic fluency performance relative to the TMT–A (PVAF = 17.14% vs. 10.43%, respectively). As noted previously, comparisons based on the number of studies as opposed to the number of participants can be regarded as conservative. Although only eight studies contributed to the comparisons of TMT–A versus phonemic fluency and were thus associated with relatively low statistical power when the number of studies was used as the degree of freedom, these studies included a total of 258 participants. Thus, although the present results suggest that reduced psychomotor speed may contribute to the deficit observed on measures of phonemic fluency, it cannot be regarded as sufficient by itself to account for the magnitude of the deficit observed.

In terms of the PVAF by group membership, the phonemic fluency deficit was also significantly larger than the deficits for both the WCST–CC and WCST–PE, indicating that of these measures, phonemic fluency is the most sensitive to TBI. Again, this mirrors the pattern of deficits associated with patients with focal frontal lesions; in a previous study (Henry & Crawford, 2004) we found phonemic fluency to be more sensitive to frontal damage, as indicated by a larger effect size, than either the WCST–CC or WCST–PE. Questions have been raised about the sensitivity of the WCST as an executive measure (Mountain & Snow, 1993), yet it has been used extensively as a measure of this construct. The present results, as well as our previous meta-analytic review (Henry & Crawford, 2004), suggest that phonemic fluency may be a more sensitive marker of executive dysfunction.

Finally, relative to measures of episodic memory, phonemic fluency was more impaired, although not significantly so. Thus, although episodic memory deficits are widely regarded as the most prominent deficit associated with TBI, the present results indicate that executive deficits are of at least comparable magnitude. As noted earlier, it has been suggested that the episodic memory deficit in TBI reflects executive dysfunction, and some evidence for this position has been presented (McAllister et al., 1999). However, although the present results may be regarded as consistent with this possibility, the direction of causality may be the opposite of that assumed. Specifically, it could be argued that the episodic memory deficit observed was responsible for, or at least contributed to, the poor performance on the tests of verbal fluency; for example, poor memory for the task instructions or for previous responses led to impaired performance. However, evidence against this possibility is provided by demonstrations that participants with dense amnesia can perform at average or above-average levels on tests of verbal fluency (see Dall'Ora, Della Sala, & Spinnler, 1989), indicating that such tests can be performed adequately despite the presence of severe episodic memory deficits. Thus, it appears more likely that the episodic memory deficits in TBI stem from deficits in executive control processes, as has been found for Huntington's disease (Crawford et al., 2000), and it would be of value to explore this possibility further in future studies.

Future Directions

For most variables assessed in the present study, the homogeneity statistic Q was significant, suggesting that substantive differences between studies remain. It is important to note that difficulties often occur even when localizing more circumscribed injuries (e.g., focal frontal or temporal lobe lesions), and thus, this heterogeneity may reflect the fact that TBI is by definition a more diffuse injury that may cover a multitude of brain regions and may or may not include frontal structures. Nevertheless, we suggest that the fluency deficits in TBI reflect executive dysfunction because, as discussed previously, there is a great deal of evidence that frontal regions of the brain are particularly susceptible to injury (see Levin & Kraus, 1994; Levin et al., 1992). However, because diffuse axonal injury is considered to be particularly related to executive impairment (Stuss & Gow, 1992), it may be that only severe TBI is consistently associated with disproportionate executive impairment. Indeed, in Ommaya and Gennarelli's (1974) centripetal model, mild CHI damage predominated on the surface of the brain. Thus, unless mild head injury is associated with a focal lesion that specifically implicates frontal structures, such patients may not exhibit executive dysfunction, and it may therefore be that patients with mild CHI exhibit greater heterogeneity than their more severely affected counterparts in this respect.

Indeed, when the number of studies is relatively small, the power to detect heterogeneity is low (Hedges & Olkin, 1985). For all of the mean effects associated with nonsignificant estimates of the homogeneity statistic, a relatively small number of groups contributed, so it may be premature to conclude that the studies contributing to each of these respective means measured a common underlying parameter (i.e., as is well known, one cannot prove the null hypothesis). Thus, all that can be said is that for some of these effects, the results are consistent with the null hypothesis of homogeneity of effects. Moreover, it is also impor-

tant to note that where significant heterogeneity was observed, it cannot be attributed simply to the presence of a few extreme outliers. In the present analyses only two outliers were identified, yet five mean effects were associated with significant heterogeneity. These results are therefore consistent with the possibility that there are a number of important moderating factors that result in the heterogeneity observed. It is unfortunately beyond the scope of the present study to address the specific influence of each of these potential moderators, because much of this variance is bundled up within rather than between studies. The homogeneity statistic *Q* quantifies the degree of heterogeneity between studies but cannot address the degree of heterogeneity within each of the studies contributing to a mean. However, if future primary research breaks down samples more fully, meta-analysis should be conducted to address which variables moderate performance on tests of verbal fluency.

In particular, it would be worth investigating whether the long-term effects of TBI differ qualitatively as well as quantitatively at different levels of severity. Although it is presumed that moderate and severe TBI is associated with lasting neurocognitive difficulties, Binder et al. (1997), using meta-analytic techniques to assess prospective studies of mild TBI at least 3 months posttrauma, concluded that typically, the effect of mild TBI on neuropsychological performance is undetectable. However, symptomatic patients with mild TBI for whom there is enduring dysfunction have been identified in a number of clinical studies (see, e.g., Raskin & Rearick, 1996). Moreover, Binder et al. (1997) did not present mean effect sizes for individual cognitive measures but collapsed effects across neuropsychological domains, with cognitive flexibility and abstraction the most executive of these. Creation of a composite executive measure is problematic because of ambiguities regarding what actually constitutes a valid measure of executive function. Moreover, there is evidence for at least three dissociable executive processes (Miyake, Friedman, Emerson, Witzki, & Howerter, 2000), each of which may be differentially affected by TBI. It would therefore be worth investigating whether patients with mild TBI exhibit enduring deficits on specific measures of this construct. Such an approach is particularly important given that the present results provide clear evidence that different measures of executive functioning vary in their sensitivity to the presence of TBI, with phonemic fluency significantly more sensitive to TBI than are either the WCST–CC or the WCST–PE. Patients with mild TBI may therefore exhibit enduring executive deficits, but these may be more specific than was possible to detect with Binder et al.'s (1997) methodology.

Summary and Conclusions

Relative to healthy control participants, TBI patients, like patients with focal frontal lobe injuries, were comparably impaired on measures of semantic and phonemic fluency. The phonemic fluency deficit could not be accounted for by level of premorbid IQ or current VIQ and was also substantially (although not significantly) in excess of the deficit on a measure of psychomotor speed. Thus, although for most patients the phonemic fluency deficit reflected problems with executive functioning and not a failure to match patients and control participants on premorbid IQ or current VIQ, poor performance on this measure may also to a certain extent be attributable to psychomotor slowing. Phonemic fluency was more sensitive to the presence of TBI than was the WCST, which has also been found to be true of patients with focal frontal lobe lesions.

References

Baddeley, A. D. (1990). *Human memory: Theory and practice*. Hove, England: Erlbaum.

Beers, S. R., Goldstein, G., & Katz, L. J. (1994). Neuropsychological differences between college students with learning disabilities and those with mild head injury. *Journal of Learning Disabilities, 27,* 315–324.

Binder, L. M., Robling, M. L., & Larrabee, G. J. (1997). A review of mild head trauma. Part I: Meta-analytic review of neuropsychological studies. *Journal of Clinical and Experimental Neuropsychology, 19,* 421–431.

Buschke, H., & Fuld, P. A. (1974). Evaluation of storage, retention, and retrieval in disordered memory and learning. *Neurology, 11,* 1019–1025.

Cockburn, J. (1995). Performance on the Tower of London test after severe head injury. *Journal of the International Neuropsychological Society, 1,* 537–544.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (Rev. ed.). New York: Academic Press.

Cooke, D. L., & Kausler, D. H. (1995). Content memory and temporal memory for actions in survivors of traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology, 17,* 90–99.

Crawford, J. R., Blackmore, L. M., Lamb, A. E., & Simpson, S. A. (2000). Is there a differential deficit in fronto-executive functioning in Huntington's disease? *Clinical Neuropsychological Assessment, 1,* 3–19.

Dall'Ora, P., Della Sala, S., & Spinnler, H. (1989). Autobiographical memory. Its impairments in amnesic syndromes. *Cortex, 25,* 197–217.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (1987). *California Verbal Learning Test: Adult Version*. New York: Psychological Corporation.

Goldstein, F. C., Levin, H. S., Roberts, V. J., Goldman, W. P., Kalechstein, A. S., Winslow, M., & Goldstein, S. J. (1996). Neuropsychological effects of closed head injury in older adults: A comparison with Alzheimer's disease. *Neuropsychology, 10,* 147–154.

Hanks, R. A., Rapport, L. J., Millis, S. R., & Deshpande, S. A. (1999). Measures of executive functioning as predictors of functional ability and social integration in a rehabilitation sample. *Archives of Physical Medicine and Rehabilitation, 80,* 1030–1037.

Heaton, R. K. (1981). *Wisconsin Card Sorting Test (WCST)*. Odessa, FL: Psychological Assessment Resources.

Hedges, L., & Olkin, I. (1985). *Statistical methods for meta-analysis*. New York: Academic Press.

Henry, J. D., & Crawford, J. R. (2004). A meta-analytic review of verbal fluency performance following focal cortical lesions. *Neuropsychology, 18,* 284–295.

Jastak, S., & Wilkinson, G. S. (1984). *Wide Range Achievement Test—Revised*. Wilmington, DE: Jastak Assessment Systems.

Lannoo, E., Colardyn, F., De Deyne, C., Vandekerckhove, T., Jannes, C., & De Soete, G. (1998). Cerebral perfusion pressure and intracranial pressure in relation to neuropsychological outcome. *Intensive Care Medicine, 24,* 236–241.

Laws, K. R. (1999). A meta-analytic review of Wisconsin card sort studies in schizophrenia: General intellectual deficit in disguise? *Cognitive Neuropsychiatry, 4,* 1–35.

Levin, H. S., Amparo, E., Eisenberg, H. M., Williams, D. H., High, W. M., McArdle, C. B., & Weiner, R. L. (1987). Magnetic resonance imaging and computerized tomography in relation to the neurobehavioral sequelae of mild and moderate head injuries. *Journal of Neurosurgery, 66,* 706–713.

Levin, H., & Kraus, M. F. (1994). The frontal lobes and traumatic brain injury. *Journal of Neuropsychiatry and Clinical Neurosciences, 6,* 443–454.

Levin, H. S., Lippold, S. C., Goldman, A., Handel, S., High, W. M., Eisenberg, H. M., & Zelitt, D. (1987). Neurobehavioral functioning and magnetic resonance imaging findings in young boxers. *Journal of Neurosurgery, 67,* 657–667.

Levin, H. S., Williams, D. H., Eisenberg, H. M., High, W. M., & Guinto, F. C. (1992). Serial MRI and neurobehavioral findings after mild to moderate closed head injury. *Journal of Neurology, Neurosurgery & Psychiatry, 55,* 255–262.

Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). New York: Oxford University Press.

Mathias, J. L., & Coats, J. L. (1999). Emotional and cognitive sequelae to mild traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology, 21,* 200–215.

McAllister, T. W., Saykin, A. J., Flashman, L. A., Sparling, M. B., Johnson, S. C., Guerin, S. J., et al. (1999). Brain activation during working memory 1 month after mild traumatic brain injury: A functional MRI study. *Neurology, 53,* 1300–1308.

McDowell, S., Whyte, J., & D'Esposito, M. (1997). Working memory impairments in traumatic brain injury: Evidence from a dual task paradigm. *Neuropsychologia, 35,* 1341–1353.

Miller, E. (1984). Verbal fluency as a function of a measure of verbal intelligence and in relation to different types of cerebral pathology. *British Journal of Clinical Psychology, 23,* 53–57.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., & Howerter, A. (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.

Mountain, M. A., & Snow, W. G. (1993). Wisconsin Card Sorting Test as a measure of frontal pathology: A review. *The Clinical Neuropsychologist, 7,* 108–118.

National Research Council. (1992). *Combining information: Statistical issues and opportunities for research.* Washington, DC: National Academy Press.

Nelson, H. E. (1982). *National Adult Reading Test (NART): Test manual*. Windsor, England: NFER Nelson.

Ommaya, A. K., & Gennarelli, T. A. (1974). Cerebral concussion and traumatic unconsciousness. Correlation of experimental and clinical observations on blunt head injuries. *Brain, 97,* 633–654.

Perret, E. (1974). The left frontal lobe of man and the suppression of habitual responses in verbal categorical behaviour. *Neuropsychologia, 12,* 323–330.

Raskin, S. A., & Rearick, E. (1996). Verbal fluency in individuals with mild traumatic brain injury. *Neuropsychology, 10,* 416–422.

Reitan, W. (1990). *The Halstead–Reitan Neuropsychological Test Battery, theory and clinical interpretation*. Tucson, AZ: Neuropsychology Press.

Stanley, T. D. (2001). Wheat from chaff: Meta-analysis as quantitative literature review. *Journal of Economic Perspectives, 15,* 131–150.

Stuss, D. T., & Benson, D. F. (1986). *The frontal lobes*. New York: Raven Press.

Stuss, D. T., Ely, P., Hugenholtz, H., Richard, M. T., Larochelle, S., Poirier, C. A., & Bell, I. (1985). Subtle neuropsychological deficits in patients with good recovery after closed head injury. *Neurosurgery, 17,* 41–47.

Stuss, D. T., & Gow, C. A. (1992). "Frontal dysfunction" after traumatic brain injury. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology, 5,* 272–282.

Wechsler, D. (1955). *The Wechsler Adult Intelligence Scale: Manual*. New York: Psychological Corporation.

Wechsler, D. (1981). *The Wechsler Adult Intelligence Scale—Revised: Manual*. New York: Psychological Corporation.

Wertz, R. T., Dronkers, N. F., & Shubitowski, Y. (1986). Discriminant function analysis of performance by normals and left hemisphere, right hemisphere, and bilaterally brain damaged patients on a word fluency measure. *Clinical Aphasiology, 16,* 257–266.

Received June 23, 2003
Revision received October 6, 2003
Accepted October 6, 2003 ■