Rehabilitation of executive functioning in patients with frontal lobe brain damage with goal management training

Brian Levine[1,2,3]*, Tom A. Schweizer[4,5], Charlene O'Connor[1,2,6], Gary Turner[6,7], Susan Gillingham[1], Donald T. Stuss[1,2,3], Tom Manly[8] and Ian H. Robertson[9]

[1] *Rotman Research Institute, Baycrest Centre, Toronto, ON, Canada*
[2] *Department of Psychology, University of Toronto, Toronto, ON, Canada*
[3] *Department of Medicine (Neurology), University of Toronto, Toronto, ON, Canada*
[4] *Department of Medicine (Neurosurgery), University of Toronto, Toronto, ON, Canada*
[5] *Keenan Research Centre of the Li Ka Shing Knowledge Institute at St. Michael's Hospital, Toronto, ON, Canada*
[6] *Department of Occupational and Rehabilitation Sciences, University of Toronto, Toronto, ON, Canada*
[7] *Heart and Stroke Foundation Centre for Stroke Recovery, Sunnybrook Health Sciences Centre, Toronto, ON, Canada*
[8] *Medical Research Council Cognition and Brain Sciences Unit, Cambridge, England*
[9] *Department of Psychology and Neurosciences, Trinity College, Dublin, Ireland*

Executive functioning deficits due to brain disease affecting frontal lobe functions cause significant real-life disability, yet solid evidence in support of executive functioning interventions is lacking. Goal Management Training (GMT), an executive functioning intervention that draws upon theories concerning goal processing and sustained attention, has received empirical support in studies of patients with traumatic brain injury, normal aging, and case studies. GMT promotes a mindful approach to complex real-life tasks that pose problems for patients with executive functioning deficits, with a main goal of periodically stopping ongoing behavior to monitor and adjust goals. In this controlled trial, an expanded version of GMT was compared to an alternative intervention, Brain Health Workshop that was matched to GMT on non-specific characteristics that can affect intervention outcome. Participants included 19 individuals in the chronic phase of recovery from brain disease (predominantly stroke) affecting frontal lobe function. Outcome data indicated specific effects of GMT on the Sustained Attention to Response Task as well as the Tower Test, a visuospatial problem-solving measure that reflected far transfer of training effects. There were no significant effects on self-report questionnaires, likely owing to the complexity of these measures in this heterogeneous patient sample. Overall, these data support the efficacy of GMT in the rehabilitation of executive functioning deficits.

Keywords: rehabilitation, executive functioning, frontal lobe, stroke, traumatic brain injury mindfulness

INTRODUCTION

Deficits in frontal and executive function are among the commonest causes of disability following brain injury. This disability affects planning, strategy application, self-regulation, inhibition, goal-directed behavior, initiation, and insight (Tranel et al., 1994; Stuss and Levine, 2002). Although primarily associated with frontal lobe damage, these deficits can occur following a host of other conditions affecting interconnected brain function through axonal or white matter changes, such as traumatic brain injury (Levine et al., 1998), multiple sclerosis (McDonald and Ron, 1999), ischemic white matter disease (Swartz et al., 2008), aging (Raz, 2000), the dementias (Neary et al., 1998), and psychiatric conditions (Cohen and Servan-Schreiber, 1992; Mayberg, 1997).

In spite of the societal costs associated with this disability, there are no widely accepted standardized interventions targeting these capacities. In a recent review of the literature, we identified 55 studies on rehabilitation of executive or frontal lobe function (Levine et al., 2008). Only 16% of these met criteria for class I evidence

sufficient to guide treatment (see also Cicerone et al., 2000, 2005; Rohling et al., 2009). Research in this area is limited by heterogeneity of patient samples, lack of control groups, lack of theoretically based intervention protocols, and limited outcome assessment (see also Levine and Downey-Lamb, 2002).

Goal management training (GMT; Robertson, 1996; Levine et al., 2000b, 2007) is based on theory of sustained or vigilant attention (Robertson and Garavan, 2000). Ongoing activation of the right frontal-thalamic-parietal sustained attention system is required to actively, endogenously maintain higher order goal states in working memory. When the sustained attention system is compromised, habits, or environmental conditions may oppose and displace higher order goals, resulting in cue-dependent or distracted behavior that is a hallmark of patients with attentional and executive deficits. For example, if one has to deviate from a habitual route to post a letter, the goal of posting the letter must be actively maintained in working memory en route in order to prevent habit (i.e., taking the usual route) from displacing the goal.

The brain's sustained attention system is distributed, mediated not only by the nodal right lateralized dorsolateral prefrontal, posterior parietal, and thalamic regions (Posner and Petersen, 1990) but also by their interconnections with ascending bottom-up systems (Corbetta and Shulman, 2002), accounting for the ubiquity of attentional deficits manifesting as dysexecutive syndrome or impaired goal management among patients with various forms of brain disease, although amodal goal management deficits are expected to be greatest in patients with frontal lobe brain damage. In this framework, sustained attention is viewed as crucial to supporting various processes collectively referred to as executive or frontal lobe function (Stuss and Levine, 2002; Stuss and Alexander, 2007), even though these may be dissociated under specific experimental conditions.

The primary objective of GMT of is to train patients to stop ongoing behavior in order to define goal hierarchies (Duncan et al., 1996) and monitor performance. This is achieved through instructional material, interactive tasks, discussion of patients' real-life deficits, and homework assignments (see Materials and Methods). Exogenous, content-free cues (i.e., random tones) can serve as a prosthetic in patients with sustained attention deficits (Manly et al., 2002; Fish et al., 2007) through the interruption of automatic responding and facilitation of the resumption of supervisory attentional control (Norman and Shallice, 1986). When random tones are presented during performance of the Sustained Attention to Response Task (SART; Robertson et al., 1997, see Materials and Methods) the classic frontoparietal neural signature of sustained attention is disengaged (O'Connor et al., in press), suggesting that the tones can serve as an external prosthetic for what are normally endogenous attentional processes. GMT incorporates this principle by introducing external prompts during the performance of analogs to experimental tasks, such as SART-like tasks and complex tasks, such as simulated real-life tasks (Levine et al., 2007), then promoting internalization of these prompts through self-cueing. Mindfulness meditation (Kabat-Zinn, 1990) is incorporated toward developing a skill of repetitively bringing one's mind to the present to monitor ongoing behavior, goal states, and the correspondence between them. The intervention also includes real-life examples provided by the trainer and the patient to illustrate goal attainment failures and successes, promoting awareness of alterations in goal states, and in session practice on complex tasks that mimic real-life tasks that are problematic for patients with executive deficits (e.g., planning a party).

Goal Management Training has been evaluated in patients with traumatic brain injury (Levine et al., 2000b; Fish et al., 2007), normal aging (Levine et al., 2007; van Hooren et al., 2007), and in case studies of patients with focal cerebellar damage (Schweizer et al., 2008) and encephalitis (Levine et al., 2000b). Significant effects in support of GMT were found for performance on laboratory analogs of real-life tasks requiring complex attentional skills (Levine et al., 2000b, 2007) and surveys of real-life attentional and executive deficits (Levine et al., 2007; van Hooren et al., 2007). Yet there are no studies in which a fully expanded GMT has been systematically compared to a viable alternative treatment. The first study of GMT, although employing a randomized control trial, consisted of a brief (1 h) version of the protocol (Levine et al., 2000b) as a rehabilitation probe or proof-of-principle in patients with traumatic brain injury (see also Fish et al., 2007 for a similar

application of GMT). Studies of GMT in aging (Levine et al., 2007; van Hooren et al., 2007) used longer protocol, but lacked an active control group, and, in the case of Levine et al. (2007), combined GMT with other interventions.

In this study, patients with brain injury (mostly stroke) were assigned to a 14-h, 7 week version of GMT or a control group consisting of standard-of-care treatment. Outcome, assessed immediately after training and at 4 months post-training, was assessed with a battery of experimental and clinical neuropsychological tests and outcome questionnaires. Each measure was selected for its capacity to assess specific processes supported by sustained attention and for its position along a continuum of ecological validity, spanning from laboratory reaction time tests to questionnaires sensitive to real-life outcome. The SART (Robertson et al., 1997), a computerized test involving speeded responses to lengthy target series, was considered a measure of "near transfer" in that GMT itself (but not the control condition) employs SART-like tasks. Different types of errors on the SART can reflect either attentional drift or resolution of response competition (O'Connell et al., 2009), both of which are related to failures of sustained attention. As the SART provides reaction time data, we also examined the effect of GMT on variability of response time, which is sensitive to frontal brain damage (Stuss et al., 2003) and associated with indices of integrated brain function in healthy individuals (McIntosh et al., 2008). While the SART directly assessed sustained attention processes targeted by GMT, the remaining tests, while more removed from pure sustained attention, assess behaviors that should nonetheless be supported by sustained attention, and in turn affected by GMT even though they were not directly trained. In this sense, they are considered measures of "far transfer." These included two desktop tests, the Tower Test from the Delis–Kaplan executive function system (D–KEFS; Delis et al., 2001), a test of visuospatial planning, and the Hotel Test (Manly et al., 2002), which mimics a real-life multitasking situation. Questionnaires (Broadbent et al., 1982; Burgess et al., 1996) focused on self-reported executive deficits in real-life.

MATERIALS AND METHODS
PARTICIPANTS
Twenty adults aged 23–70 were recruited from metropolitan hospitals in the Toronto area. Criteria for inclusion were a focal brain injury stabilized in the chronic phase of recovery (minimum 6 months post-injury) and executive functioning problems, such as impaired planning, decision making, self-regulation as determined by clinician referrals and intake interview. We did not restrict participation to those with impaired performance on traditional tests of frontal lobe function as these tests are of limited utility in the assessment of real-life executive deficits of interest in this study (Shallice and Burgess, 1991; Levine et al., 1998; Stuss and Levine, 2002). Criteria for exclusion, based on review of records, included impaired basic linguistic, mnemonic, motor, or perceptual function that would interfere with the patient's ability to engage with the training. Additionally, patients with Axis I psychiatric disorders were excluded. The study was approved by the Baycrest Research Ethics Board. All patients gave informed consent for participation.

The first 15 patients were assigned to GMT or control groups (hereafter referred to as brain health workshop; BHW). GMT and BHW were each delivered in groups of 2–4 participants. Twelve

of the 15 participants were randomly assigned; the remaining two participants were assigned to the BHW group because their schedules did not permit assignment to the GMT group. One patient randomly assigned to GMT failed to complete the training. We thus had N's of 6 and 8 in the GMT and BHW groups, respectively. We ran an additional GMT group that contained five participants (three stroke, one TBI, one tumor), resulting in total N's of 11 and 8 for the GMT and BHW groups, respectively. Although the last five participants were not randomly assigned (i.e., they were all in a GMT group), ancillary analyses (see Results) indicated that the treatment effects described below cannot be accounted for by this group.

Participants in the GMT and BHW groups were matched for age, gender, education, and neuropsychological test performance (all p's for group comparisons on these variables >0.3; see **Table 1**). The GMT group contained six patients with stroke, four patients with traumatic brain injury, and one patient with tumor. The BHW group contained five patients with stroke, and one patient each with traumatic brain injury, epilepsy (status post right temporal lobe resection), and heterotopia. Lesion information was documented with clinical imaging. In most cases, lesions were located in the frontal lobes or frontal systems (basal ganglia, cerebellum, thalamus), or diffusely in the cases of traumatic brain injury.

INTERVENTION

Both GMT and BHW were administered in seven 2 h sessions (see **Table 2**) following a script with Powerpoint slides and participant workbooks. Session 1 was devoted to defining the concept of absentmindedness and raising awareness of absentminded errors in daily life. In Session 2, absentminded errors were operationalized as inappropriate expressions of habit (i.e., the "automatic pilot"). This is when the "STOP" concept was introduced. In Session 3, working memory was defined as the mind's "mental blackboard" for on-line maintenance of goals. Given the sensitivity of working memory to distraction, frequent checking (i.e., stopping) is required to ensure that the contents of working memory is matched to appropriate goal states. Mindfulness-based meditation (Kabat-Zinn, 1990; Segal et al., 2002) was introduced to enhance awareness toward current behavior, feelings, and goal

states. Session 4 emphasized stating the goal as a method to activate goal representations following stopping and bringing one's mind to the present. Session 5 dealt with decision making in the context of competing goals and the use of to-do lists. In Session 6, participants practiced splitting unwieldy goals into subgoals. The final session involved checking, or stopping to interrupt ongoing behavior for the purposes of output monitoring.

Each session contained narrative examples to illustrate concepts. Patients' own real-life examples were solicited and used whenever possible. Analogs of experimental and real-life tasks were used to induce experiential learning of the concepts within the sessions. Homework assignments between sessions included monitoring of absentminded slips and successes (i.e., recovery from or avoidance of slips), and daily mindfulness practice. The same trainer (either Brian Levine or Charlene O'Connor) conducted both GMT and BHW. Recordings of sessions were periodically reviewed to ensure consistency of administration.

Brain Health Workshop was matched to GMT for session length and contact with the facilitator. It comprised educational materials and lifestyle interventions employed in rehabilitation centers (see **Table 2**). Educational sessions addressed brain function and dysfunction, neuroplasticity, and cognitive psychological and neuropsychological principles concerning memory, attention, language, perception, and motor skills. The lifestyle sessions included information on stress reduction, sleep hygiene, energy management, exercise, communication, and nutrition. Within session and homework activities included reading assignments, testing of acquired knowledge from the educational materials, mental games (e.g., word finding, mazes), and logs of communication and sleep habits.

OUTCOME MEASURES

The efficacy of treatment was assessed via a battery of standardized and experimental tests of executive functioning and attention and questionnaires that ranged from "near transfer" measures methodologically close to the tasks incorporated into GMT to "far transfer" measures that, while theoretically supported by GMT, were not directly addressed in the intervention (see **Table 2**). Outcome assessments were conducted immediately pre- and post-intervention, and at 4 month follow-up by an examiner blind to group membership. The examiner was from a different lab than that from which the groups were assigned and run, and she had no access to group membership information.

The SART (Robertson et al., 1997) is a go/no-go task in which single-digit numbers are presented randomly every 1150 ms. Participants are instructed to respond to all numbers (go stimuli) except for a single no-go number, appearing 11% of the time. Dependent variables include errors of commission (responding to no-go stimuli), errors of omission (not responding to go stimuli), reaction time, and reaction time variability (coefficient of variation). Errors of commission and omission can be dissociated both behaviorally and at the brain level (ERP), with the former reflecting resolution of response competition and the latter reflecting attentional drift (O'Connell et al., 2009).

The D–KEFs Tower Test (Delis et al., 2001) is a measure of visuospatial planning that is modeled on the Tower of London test (Shallice, 1982), in which disks must be placed on dowels to

Table 1 | Demographic and neuropsychological test data.

	GMT (N = 11)	BHW (N = 8)
Age	48.91 (12.83)	49.25 (13.85)
Sex	8M, 3F	6M, 2F
Education	17.33 (2.96)	15.63 (3.54)
NART FSIQ estimate	110.71 (5.23)	107.52 (6.86)
Digits forward	10.36 (2.29)	9.75 (1.04)
Digits backward	5.55 (2.50)	6.38 (1.60)
Digit symbol	65.09 (13.25)	58.63 (23.04)
Trails A time (s)	37.73 (11.72)	33.38 (10.93)
Trails B time (s)	85.27 (39.66)	115.25 (99.67)
Phonemic word list generation	36.18 (10.25)	36.25 (9.44)
WCST categories	7.10 (2.77)	8.14 (4.06)
WCST errors	29.40 (19.46)	33.57 (25.39)
WCST set loss	1.30 (1.49)	1.29 (1.80)

Table 2 | Description of goal management training (GMT) and brain health workshop (BHW) interventions.

GMT session	Objectives	BHW session	Objectives
1. Slip-ups	Overall introduction Defining goals, absentmindedness, action slips Raising awareness of consequences of action slips	1. Introduction	General introduction Basic brain anatomy and cognition Etiology of brain damage and assessment modalities
2. Stop the Automatic Pilot	Defining automatic pilot (habit vs. control distinction) How automatic pilot can lead to errors Stopping the automatic pilot	2. Neuroplasticity	Hebbian learning Importance of keeping brain active Functional assessment of brain activity
3. The mental blackboard and present-mindedness	Defining the mental blackboard (working memory) Using "STOP!" to check the mental blackboard Mindfulness exercise to promote present-mindedness	3. Memory I	Review: Brain jeopardy Importance of memory Types of memory, memory processes
4. Slate your goal	Being sidetracked from your goal Stating goals to activate working memory representation "STOP!" (present-mindedness)-STATE cycle	4. Memory II	Memory and the brain How memory breaks down Functional implications of memory loss
5. Making decisions	Examples of competing goals Understanding emotional reaction to competing goals, including indecision To-Do Lists in the "STOP"-STATE cycle	5. Attention and executive functions	Defining executive functioning and attention How executive functions and attention break down Group problem-solving exercises
6. Splitting tasks into subtasks	Defining overwhelming goals that require splitting Organizing goal hierarchies "STOP!"-STATE-SPLIT cycle	6. Lifestyle and neuroplasticity I	Influence of lifestyle on neuroplasticity and recovery Stress and brain function Sleep and brain function
7. Checking (STOP!)	Recognizing errors in "STOP!"-STATE-SPLIT cycle Using "STOP!" to monitor output Review	7. Lifestyle and neuroplasticity II	Nutrition and brain function Physical exercise and brain function Review and brain jeopardy

match models of increasing complexity, with "rules" constraining the movement of the disks (i.e., it is forbidden to place a larger disk on top of a smaller disk). For the purposes of this study, the standardized test was split into two forms by alternating items. We examined the total achievement score and the number of rule violations (expressed as a proportion of items completed).

The Hotel Task (Manly et al., 2002), modeled on the Six Element Test (Shallice and Burgess, 1991), is a desktop model of a real-life multitasking situation in which the participant plays the role of a hotel manager with five different tasks. In order to maximize performance and complete the tasks effectively, participants must distribute their time equally across tasks within the 15-min allotment. The dependent variables are the number of tasks completed and the deviation time, defined as the number of seconds spent above or below the optimal time of 3 min per task.

Questionnaires were used to assess everyday function that is not typically captured by laboratory tests. The Dysexecutive

Questionnaire (DEX; Burgess et al., 1996) assesses cognitive changes associated with executive functioning in the areas of inhibition, positive and negative affect, memory, and intention. Although the DEX contains self- and other-rated forms, we were unable to obtain sufficient data from significant others to include these ratings in our analyses. The Cognitive Failures Questionnaire (CFQ; Broadbent et al., 1982) addresses daily life mental errors on the dimensions of distractibility, blunders, names, and memory. Finally, we administered a questionnaire addressing daily life absentmindedness and goal attainment that was purpose-built for this study.

RESULTS

Data were analyzed using a 2×3 mixed-design ANOVA that treated Group (GMT, BHW) as a between-subjects factor and Session (baseline, post-intervention, follow-up) as a within-subjects variable. Group $\times$ Session interactions (across the three sessions and separately for baseline vs. post-intervention and baseline vs.

Table 3 | Outcome data.

Test	Assessment	Group	
		GMT M(SD)	CTL M(SD)
SART		**(n = 10)**	**(n = 8)**
Commission errors	Baseline	24.80 (14.13)	30.25 (20.82)
	Post-intervention	21.20 (15.32)	29.88 (23.44)
	Follow-up	17.40 (10.84)	23.14 (21.62)
Omission errors	Baseline	30.55 (32.51)	20.25 (15.23)
	Post-intervention	11.73 (16.99)	18.21 (19.14)
	Follow-up	8.93 (14.17)	22.29 (10.67)
Reaction time	Baseline	467.07 (134.23)	421.46 (110.01)
	Post-intervention	440.20 (108.95)	405.67 (109.72)
	Follow-up	434.08 (92.21)	428.56 (126.50)
Coefficient	Baseline	0.29 (0.09)	0.29 (0.07)
of variation	Post-intervention	0.26 (0.05)	0.28 (0.11)
	Follow-up	0.25 (0.06)	0.28 (0.08)
D–KEFs tower		**(n = 11)**	**(n = 7)**
Total achievement	Baseline	13.55 (4.52)	15.63 (4.81)
score	Post-intervention	16.64 (3.70)	17.14 (4.81)
	Follow-up	19.27 (2.94)	17.71 (3.15)
Rule violation	Baseline	0.19 (0.15)	0.11 (0.19)
per item ratio	Post-intervention	0.00 (0.00)	0.05 (0.09)
	Follow-up	0.02 (0.04)	0.03 (0.05)
Hotel task		**(n = 11)**	**(n = 8)**
Number of tasks	Baseline	4.55 (0.82)	4.38 (1.06)
attempted	Post-intervention	4.64 (0.81)	5.0 (0.00)
	Follow-up	4.91 (0.30)	4.86 (0.38)
Deviations from	Baseline	446.91 (198.04)	473.50 (179.45)
optimal time	Post-intervention	699.55 (256.49)	470.88 (159.62)
	Follow-up	580.09 (234.55)	706.00 (305.66)
CFQ		**(n = 11)**	**(n = 8)**
Total score	Baseline	43.360 (23.14)	52.25 (13.50)
	Post-intervention	40.73 (19.13)	54.38 (7.03)
	Follow-up	40.18 (20.19)	44.29 (7.76)
DEX		**(n = 11)**	**(n = 8)**
Total score	Baseline	27.27 (19.27)	36.38 (10.88)
	Post-intervention	25.18 (15.97)	37.13 (13.46)
	Follow-up	20.82 (16.58)	32.57 (10.91)
GMT-Q		**(n = 10)**	**(n = 6)**
Total score	Baseline	136.30 (88,37)	207.86 (62.62)
	Post-intervention	99.55 (69.74)	200.00 (59.27)
	Follow-up	106.00 (80.67)	179.71 (76.26)

N's are provided as data were missing for certain measures.
Framed cells indicate significant effects in comparison to baseline for GMT group (p < 0.05).
There were no significant effects for the BHW group.

follow-up) were decomposed with simple effects analyses of session effects within groups and planned paired contrasts of baseline vs. post-intervention and baseline vs. follow-up differences. A similar decomposition approach was taken for main effects of test

session. That is, session effects were probed separately for each group. Results are reported in **Table 3** and **Figure 1**. As each measure reflected distinct processes at different levels of ecological validity, experiment-wise error was not corrected. Interpretation of the strength of experimental effects is provided with effect size statistics, including partial eta-squared (η_p^2- reflecting the percentage of variance accounted for by the effect) for ANOVA results and Cohen's d for follow-up univariate tests.

SUSTAINED ATTENTION TO RESPONSE TASK
There was a significant Group × Session interaction for SART omission errors, $F_{(2,30)} = 4.872$, $p < 0.02$, $\eta_p^2 = 0.245$, due to a reduction in errors across sessions for the GMT group, $F_{(1.1,10.1)} = 7.255$, $p < 0.03$, $\eta_p^2 = 0.446$, but not the control group (see **Table 3**; **Figure 1**, top left). The GMT group's change from baseline to post-intervention, $t_{(9)} = 2.668$, $p < 0.03$, $d = 1.12$, held at follow-up, $t_{(9)} = 2.796$, $p < 0.03$, $d = 1.23$. There was also a main effect of commission errors across all three sessions, $F_{(2,30)} = 3.509$, $p < 0.05$, $\eta_p^2 = 0.190$, but this was reliable only for the GMT group at baseline vs. follow-up, $t_{(9)} = 2.355$, $p < 0.05$, $d = 0.78$ (see **Table 3**). There was a marginally significant main effect of Session for coefficient of variation of SART response times at baseline vs. follow-up, $F_{(1,15)} = 4.279$, $p < .06$, $\eta_p^2 = 0.222$, that was accounted for by the GMT group, $t_{(9)} = 2.410$, $p < 0.04$, $d = 0.993$ (see **Table 3**; **Figure 1**, top right).
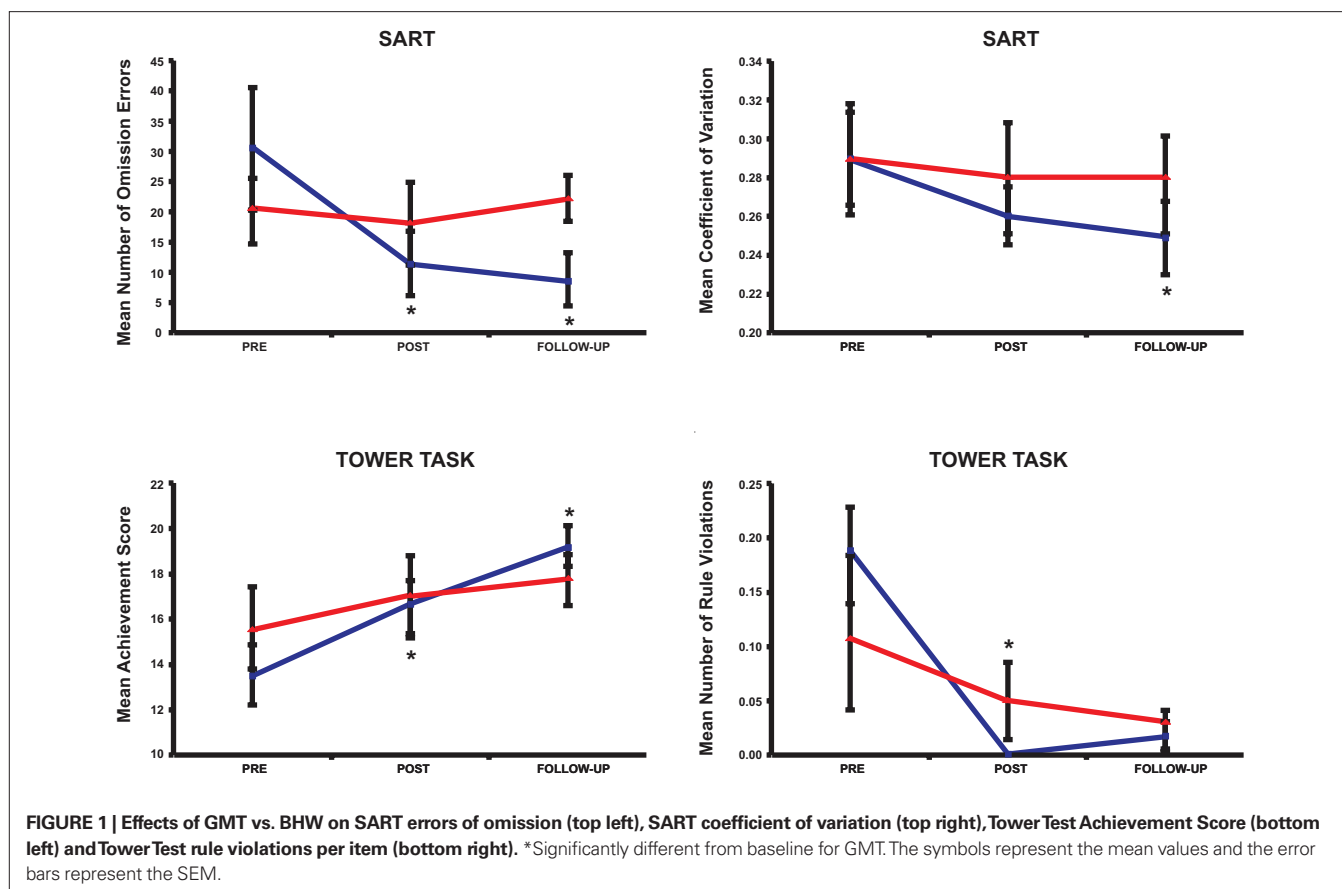
TOWER TEST
There were no reliable interactions on the D–KEFs Tower Test, although the Group × Session interaction for the achievement score approached significance at baseline vs. follow-up, $F_{(1,16)} = 3.942$, $p < 0.07$, $\eta_p^2 = 0.222$. Simple effects of test session were observed for the GMT group, $F_{(2,20)} = 11.355$, $p < 0.002$, $\eta_p^2 = 0.532$, holding at both post-intervention, $t_{(10)} = -2.437$, $p < 0.04$, $d = 0.746$, and follow-up, $t_{(10)} = -4.632$, $p < 0.002$, $d = 1.48$; there were no significant simple effects for the BHW group (see **Table 3**; **Figure 1**, bottom left). There was a main effect of test session for rule violations, $F_{(1.1,18.0)} = 13.15$, $p < 0.002$, $\eta_p^2 = 0.451$, that was significant for the GMT group, $F_{(1.1,11.3)} = 15.21$, $p < 0.003$, $\eta_p^2 = 0.603$ (holding at post-intervention, $t_{(10)} = 4.037$, $p < 0.003$, $d = 1.86$, and follow-up, $t_{(10)} = 3.903$, $p < 0.004$, $d = 1.54$) but not for the BHW group (see **Table 3**; **Figure 1**, bottom right).

HOTEL TEST
There were no effects on number of tasks attempted, likely owing to ceiling effects (see **Table 3**). There was a Group × Session interaction for total deviation time, $F_{(2,32)} = 4.260$, $p < 0.03$, $\eta_p^2 = 0.210$. Contrary to expectation, this was accounted for by an *increase* in deviation time for the GMT group at post-training, $t_{(10)} = -2.872$, $p < 0.02$, $d = 0.88$, meaning that the patients in GMT distributed their time less consistently across tasks at post-training as compared to baseline. This finding did not hold at follow-up.

QUESTIONNAIRES
There were no significant main effects or interactions on the questionnaires. Within-group exploratory analyses were also non-significant, although the change in the GMT group from baseline to post-intervention was marginally significant, $t_{(9)} = 2.090$, $p < 0.07$, $d = 0.698$ (see **Table 3**).

FIGURE 1 | Effects of GMT vs. BHW on SART errors of omission (top left), SART coefficient of variation (top right), Tower Test Achievement Score (bottom left) and Tower Test rule violations per item (bottom right). *Significantly different from baseline for GMT. The symbols represent the mean values and the error bars represent the SEM.

ADDITIONAL ANALYSES

As the last five GMT participants were not randomly assigned, we conducted ancillary analyses to determine if they were systematically different from the first six participants randomly assigned to GMT. There were no significant differences on the baseline neuropsychological tests (as listed in **Table 1**; all p's > 0.4, except for Trail Making, Part B, $t(9) = 2.115$, $p < 0.07$, where the first group was slower). Nor were there differences on any of the pre-intervention forms of the outcome measures (all p's > 0.5).

A 2 × 3 mixed-design ANOVA that treated Group (first six or last five GMT participants) as a between-subjects factor and Session (baseline, post-intervention, follow-up) as a within-subjects factor revealed no significant main effects or interactions involving Group for the SART (all p's > 0.5). For the Tower test, there was a non-significant Group × Session interaction for rule violations, $F_{(1.2, 10.6)} = 3.75$, $p < 0.08$, $\eta_p^2 = 0.294$, attributable to a more reliable effect at follow-up testing for the first six GMT participants, $t(5) = 3.993$, $p < 0.02$, $d = 1.98$) than for the last five GMT participants, $t(4) = 2.281$, $p < 0.09$, $d = 1.99$). There were no significant main effects of interactions involving Group for the Tower achievement score (p's > 0.4). For the Hotel Test, there were no significant effects on number of tasks attempted (p's > 0.3). For Hotel Test deviation time, there was a non-significant trend for a group effect, $F_{(1,9)} = 3.761$, $p < 0.09$ with higher scores for the last five participants as compared to the first six participants. There were no significant main effects or interactions involving Group for the questionnaires.

DISCUSSION

Goal Management Training is a theoretically derived intervention for executive functioning deficits intended to promote a mindful approach to problem-solving by raising awareness of attentional lapses and reinstating cognitive control when behavior is mismatched to the ongoing goal hierarchy (GMT; Robertson, 1996; Levine et al., 2000b, 2007). In this study of patients with stable brain lesions and self-reported executive deficits, GMT was associated with reduced attentional lapses, increased behavioral consistency, and improved problem-solving performance.

There is currently little high-quality evidence in support of neuropsychological interventions for executive functioning deficits (Levine et al., 2008). GMT differs from other training protocols that have shown transfer (e.g., Jaeggi et al., 2008; Persson and Reuter-Lorenz, 2008) in that it is a metacognitive intervention that combines education, narrative, task performance and feedback, and incorporation of participants' own personal task failures and successes, rather than simple repetitive practice on an automated task. It is therefore more intrinsically engaging and tailored to the individual than other methods.

Previous group studies with GMT have demonstrated positive effects on both analogs of real-life tasks and questionnaires, but these lacked an active control group (Levine et al., 2007; van Hooren et al., 2007) or used an abbreviated version of the protocol. In this study, participants were assigned to either GMT or the BHW, matched to GMT for contact with the trainer and other non-specific factors important to intervention outcomes.

Moreover, the BHW actively engaged participants with information, materials, and activities that reflect current standards of care. There was no attrition among BHW participants. Participants in both groups informally reported that they benefited from the training. Although this sample was small and heterogenous, effect size estimates indicated moderate, and in most cases, large training effects.

Goal Management Training emphasizes bolstering of sustained attention to maintain awareness of goal states and output monitoring. Accordingly, there were significant effects specific to the GMT group for errors on the SART. In particular, there was a reduction of errors of omission (i.e., not responding to go items) following GMT, but not BHW, that held at follow-up. These errors reflect attentional drift and loss of top-down control leading to mistimed responses (O'Connell et al., 2009). Errors of commission, reflecting resolution of response conflict as well as sustained attention, were also specifically reduced following the GMT, although this effect was significant only in the baseline–follow-up comparison. Instability or variability of responses is a signature effect of frontal brain damage (Stuss et al., 2003) and traumatic brain injury (Stuss et al., 1989, 1994) that is inversely correlated with "brain noise" or complexity of neural responses (McIntosh et al., 2008). GMT was associated with reduced response time variability on the SART. Although this difference reached significance only for the baseline–follow-up comparison, this finding suggests that increasing consistency of responding (or inversely, variability of neural responses) may be an important outcome target for executive functioning rehabilitation.

The positive findings for the SART are considered evidence of near transfer in that participants reliably demonstrated improved performance on a task that was directly addressed in GMT. The Tower test is a visuospatial problem-solving test that, while supported by processes trained in GMT, has little direct overlap with elements of GMT. Although less reliable than the SART, the effects on the Tower test indicated significant improvements (improved performance, reduced rule violations) that were not observed for the BHW group, supporting the efficacy of GMT on a multifactorial far transfer test.

Results on the Hotel test were contrary to expectation in that GMT participants showed greater deviations from the optimal per-task time allocation than did the BHW participants at post-intervention. This suggests that GMT may have been associated with a change in participants' approach to the task that was inconsistent with the original intent of the time deviation measure. Indeed, in both the Hotel task and in multitasking exercises that are part of the GMT intervention, participants are told that they must sample from all of the task elements, but they are not required to-do so equally. It is possible that they adopted asymmetric task completion as a strategy to focus on elements of the test where they felt most capable.

The ultimate far transfer of a behavioral intervention is to the daily life of the trainees. GMT is designed to promote generalization through intrinsic engagement and emphasis on patients' real-life issues. It is notoriously difficult to reliably measure daily life function in patients with executive functioning deficits due to brain disease. This lack of reliability is compounded in the instance of assessing intervention effects. That is, the ability to detect a change

in behavior is by necessity limited by the reliability of measurement at a single time point. Alternatives to assessing real-life behavior include laboratory analogs of real-life tasks (Shallice and Burgess, 1991; Levine et al., 1998), such as the Hotel task, and outcome questionnaires (Burgess et al., 1998; Levine et al., 2000a). We failed to find effects of GMT on three such questionnaires, the DEX, the CFQ, and a purpose-built questionnaire designed to assess goal management application. In previous studies, significant effects following GMT were observed on the DEX (Levine et al., 2007) and the CFQ (van Hooren et al., 2007). These studies employed healthy older participants who likely have higher insight into their deficits relative to patients. In the present study, we were unable to obtain significant other ratings that would likely be more sensitive to deficits and change than self-ratings, especially for participants with executive deficits due to brain disease (Burgess et al., 1998). It is also possible that participants' responses may have reflected a combination of increased insight into deficits after training, which would increase item endorsement and work against demonstration of training effects.

The data presented support the claims that prior positive findings with GMT were not due to non-specific intervention effects such as professional attention or group dynamics. Furthermore, the data indicate that GMT promotes far transfer to untrained tests of executive functioning, and that these changes are lasting. Nonetheless, limitations of this study qualify the findings and provide direction for future research. The trial was only partially randomized in that two participants were assigned to BHW due to scheduling constraints and the final five participants were assigned to GMT. It is unlikely that the significant findings could be accounted for by the incomplete randomization. There is no reason to expect that scheduling conflicts affecting group assignment for two participants would be systematically associated with outcomes. The assignment of the last five participants to GMT is a greater threat to interpretation. However, ancillary analyses indicated that this group was not different from the first six GMT participants on background demographic and neuropsychological test performance, nor did these two groups differ on baseline outcome measures. There were no significant main effects or interactions involving these two subgroups across outcome assessments for the main experimental measures. In other words, there was no evidence that the last five, non-randomized participants responded differentially to GMT. A marginally significant Group × Session interaction for rule violations on the Tower test suggested, if anything, an attenuated effect in the last five GMT participants, which would work against the significant treatment effects for this variable.

The main inclusion criterion for this study was self- and clinician-reported executive functioning deficits associated with a significant brain injury affecting the frontal lobes or their interconnected systems. In theory, executive functioning intervention should target psychological processes rather than etiology; an effective intervention should apply across etiologies, especially in patients with stable lesions, as in this study. Nonetheless, the heterogeneity of this sample, although representative of a true clinical population with executive impairment, may have influenced the results, especially considering the small sample size. Similarly, the incomplete counterbalancing of diagnosis across groups cannot

be ruled out as contributing to the effects. We are therefore currently conducting trials of GMT on more homogeneous patient samples. We did not have access to high-quality imaging data that would be helpful in characterizing patients and identifying lesion profiles associated with different treatment outcomes. We attempted to capture the executive functioning deficits of patients with frontal lobe damage using a battery of laboratory tests and outcome questionnaires. The ability to detect change specific to an intervention decreases as the dimensionality of outcome measures increase. Thus the strongest effects of GMT were observed on the SART, a specific measure of sustained attention. There were no effects on multifactorial questionnaires. There was an effect on an ecologically valid laboratory test, the Hotel test, but this was difficult to interpret as it did not conform to expectation. Future research will require more advanced outcome measures to assess generalization to real-life in larger, more homogenous, randomized samples.

REFERENCES

Broadbent, D. E., Cooper, P. F., FitzGerald, P., and Parkes, K. R. (1982). The cognitive failures questionnaire (CFQ) and its correlates. *Br. J. Clin. Psychol.* 21(Pt 1), 1–16.

Burgess, P. W., Alderman, N., Evans, J., Emslie H., and Wilson, B. A. (1998). The ecological validity of tests of executive function. *J. Int. Neuropsychol. Soc.* 4, 547–558.

Burgess, P. W., Alderman, N., Evans, J. J., Wilson, B. A., and Emslie, H. (1996). "The dysexecutive questionnaire," in *Behavioral Assessment of the Dysexecutive Syndrome*, eds B. A. Wilson, N. Alderman, P. W. Burgess, H. Emslie, and J. J. Evans (Bury St. Edmunds: Thames Valley Test Company).

Cicerone, K. D., Dahlberg, C., Kalmar, K., Langenbahn, D. M., Malec, J. F., Bergquist, T. F., Felicetti, T., Giacino, J. T., Harley, J. P., Harrington, D. E., Herzog, J., Kneipp, S., Laatsch, L., and Morse, P. A. (2000). Evidence-based cognitive rehabilitation: recommendations for clinical practice. *Arch. Phys. Med. Rehabil.* 81, 1596–1615.

Cicerone, K. D., Dahlberg, C., Malec, J. F., Langenbahn, D. M., Felicetti, T., Kneipp, S., Ellmo, W., Kalmar, K., Giacino, J. T., Harley, J. P., Laatsch, L., Morse, P. A., and Catanese, J. (2005). Evidence-based cognitive rehabilitation: updated review of the literature from 1998 through 2002. *Arch. Phys. Med. Rehabil.* 86, 1681–1692. [See comment].

Cohen, J. D., and Servan-Schreiber, D. (1992). Context, cortex, and dopamine: a connectionist approach to behavior and biology in schizophrenia. *Psychol. Bull.* 99, 45–77.

Corbetta, M., and Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215.

Delis, D. C., Kaplan, E., and Kramer, J. H. (2001). *Delis-Kaplan Executive Functioning System (D-KEFS).* San Antonio: The Psychological Corporation.

Duncan, J., Emslie, H., Williams, P., Johnson, R., and Freer, C. (1996). Intelligence and the frontal lobe: the organization of goal-directed behavior. *Cogn. Psychol.* 30, 257–303.

Fish, J., Evans, J. J., Nimmo, M., Martin, E., Kersel, D., Bateman, A., Wilson, B. A., and Manly, T. (2007). Rehabilitation of executive dysfunction following brain injury: "content-free" cueing improves everyday prospective memory performance. *Neuropsychologia* 45, 1318–1330.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., and Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proc. Natl. Acad. Sci. U.S.A.* 105, 6829–6833.

Kabat-Zinn, J. (1990). *Full Catastrophe Living.* New York: Dell Publishing.

Levine, B., Dawson, D., Boutet, I., Schwartz, M. L., and Stuss, D. T. (2000a). Assessment of strategic self-regulation in traumatic brain injury: its relationship to injury severity and psychosocial outcome. *Neuropsychology* 14, 491–500.

Levine, B., Robertson, I. H., Clare, L., Carter, G., Hong, J., Wilson, B. A., Duncan, J., and Stuss, D. T. (2000b). Rehabilitation of executive functioning: an experimental-clinical validation of goal management training. *J. Int. Neuropsychol. Soc.* 6, 299–312.

Levine, B., and Downey-Lamb, M. (2002). "Design and evaluation of intervention experiments," in *Neuropsychological Interventions: Emerging Treatment and Management Models for Neuropsychological Impairments*, ed. P. J. Eslinger (New York: Guilford Publications), 80–104.

Levine, B., Stuss, D. T., Milberg, W. P., Alexander, M. P., Schwartz, M., and Macdonald, R. (1998). The effects of focal and diffuse brain damage on strategy application: evidence from focal lesions, traumatic brain injury, and normal aging. *J. Int. Neuropsychol. Soc.* 4, 247–264.

Levine, B., Stuss, D. T., Winocur, G., Binns, M. A., Fahy, L., Mandic, M., Bridges, K., and Robertson, I. H. (2007). Cognitive rehabilitation in the elderly: effects on strategic behavior in relation to goal management. *J. Int. Neuropsychol. Soc.* 13, 143–152.

Levine, B., Turner, G. R., and Stuss, D. T. (2008). "Rehabilitation of frontal lobe functions," in *Cognitive Neurorehabilitation, Evidence and Applications*, eds D. T. Stuss, G. Winocur, and I. H. Robertson (Cambridge: Cambridge University Press), 464–486.

Manly, T., Hawkins, K., Evans, J., Woldt, K., and Robertson, I. H. (2002). Rehabilitation of executive function: facilitation of effective goal management on complex tasks using periodic auditory alerts. *Neuropsychologia* 40, 271–281.

Mayberg, H. S. (1997). Limbic-cortical dysregulation: a proposed model of depression. *J. Neuropsychiatry Clin. Neurosci.* 9, 471–481.

McDonald, W. I., and Ron, M. A. (1999). Multiple sclerosis: the disease and its manifestations. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 354, 1615–1622.

McIntosh, A. R., Kovacevic, N., and Itier, R. J. (2008). Increased brain signal variability accompanies lower behavioral variability in development. *PLoS Comput. Biol.* 4, e1000106. doi: 10.1371/journal.pcbi.1000106

Neary, D., Snowden, J. S., Gustafson, L., Passant, U., Stuss, D. T., Black, S., Freedman, M., Kertesz, A., Robert, P. H., Albert, M., Boone, K., Miller, B. L., Cummings, J., and Benson, D. F. (1998). Frontotemporal lobar degeneration: a consensus on clinical diagnostic criteria. *Neurology* 51, 1546–1554.

Norman, D. A., and Shallice, T. (1986). "Attention to action: willed and automatic control of behaviour," in *Consciousness and Self-regulation*, eds G. E. Schwartz and D. Shapiro (New York: Plenum Press), 1–18.

O'Connell, R. G., Dockree, P. M., Bellgrove, M. A., Turin, A., Ward, S., Foxe, J. J., and Robertson, I. H. (2009). Two types of action error: electrophysiological evidence for separable inhibitory and sustained attention neural mechanisms producing error on go/no-go tasks. *J. Cogn. Neurosci.* 21, 93–104.

O'Connor, C., Robertson, I. H., and Levine, B. (in press). The prosthetics of vigilant attention: random cuing cuts processing demands. *Neuropsychology.*

Persson, J., and Reuter-Lorenz, P. A. (2008). Gaining control: training executive function and far transfer of the ability to resolve interference. *Psychol. Sci.* 19, 881–888.

Posner, M. I., and Petersen, S. E. (1990). The attention system of the human brain. *Annu. Rev. Neurosci.* 13, 25–42.

Raz, N. (2000). "Aging of the brain and its impact on cognitive performance: integration of structural and functional findings," in *The handbook of Aging and Cognition*, 2nd Edn, eds F. I. M. Craik and T. A. Salthouse (Mahwah, NJ: Lawrence Erlbaum Associates, Inc.), 1–90.

Robertson, I. H. (1996). *Goal Management Training: A Clinical Manual.* Cambridge: PsyConsult.

Robertson, I. H., and Garavan, H. (2000). "Vigilant attention," in *The New Cognitive Neurosciences*, ed. M. Gazzaniga (Cambridge, MA: MIT Press), 563–578.

Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. T., and Yiend, J. (1997). "Oops!": performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia* 35, 747–758.

Rohling, M. L., Faust, M. E., Beverly, B., and Demakis, G. (2009). Effectiveness of cognitive rehabilitation following acquired brain injury: a meta-analytic re-examination of Cicerone et al.'s (2000, 2005) systematic reviews. *Neuropsychology* 23, 20–39.

Schweizer, T. A., Levine, B., Rewilak, D., O'Connor, C., Turner, G., Alexander, M. P., Cusimano, M., Manly, T., Robertson, I. H., and Stuss, D. T. (2008). Rehabilitation of executive functioning after focal damage to the cerebellum. *Neurorehabil. Neural. Repair* 22, 72–77.

Segal, Z. V., Williams, J. M. G., and Teasdale, J. D. (2002). *Mindfulness-based Cognitive Therapy for Depression: A New Approach to Preventing Relapse xiv.* New York, NY: Guilford Press, 351.

Shallice, T. (1982). Specific impairments in planning. *Philos. Trans. R. Soc. Lond., B* 298, 199–209.

Shallice, T., and Burgess, P. W. (1991). Deficits in strategy application following frontal lobe damage in man. *Brain* 114, 727–741.

Stuss, D. T., and Alexander, M. P. (2007). Is there a dysexecutive syndrome? *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 362, 901–915.

Stuss, D. T., and Levine, B. (2002). Adult clinical neuropsychology: lessons from studies of the frontal lobes. *Annu. Rev. Psychol.* 53, 401–433.

Stuss, D. T., Murphy, K. J., Binns, M. A., and Alexander, M. P. (2003). Staying on the job: the frontal lobes control individual performance variability. *Brain* 126(Pt 11), 2363–2380.

Stuss, D. T., Pogue, J., Buckle, L., and Bondar, J. (1994). Characterization of stability of performance in patients with traumatic brain injury: variability and consistency on reaction time tests. *Neuropsychology* 8, 316–324.

Stuss, D. T., Stethem, L. L., Hugenholtz, H., Picton, T., Pivik, J., and Richard, M. T. (1989). Reaction time after traumatic brain injury: fatigue, divided and focused attention, and consistency of performance. *J. Neurol. Neurosurg. Psychiatr.* 52, 742–748.

Swartz, R. H., Stuss, D. T., Gao, F., and Black, S. E. (2008). Independent cognitive effects of atrophy and diffuse subcortical and thalamico-cortical cerebrovascular disease in dementia. *Stroke* 39, 822–830.

Tranel, D., Anderson, S. W., and Benton, A. (1994). "Development of the concept of "executive function" and its relationship to the frontal lobes," in *Handbook of Neuropsychology*, eds F. Boller and J. Grafman (Amsterdam: Elsevier), 125–148.

van Hooren, S. A., Valentijn, S. A., Bosma, H., Ponds, R. W., van Boxtel, M. P., Levine, B., Robertson, I. H., and Jolles, J. (2007). Effect of a structured course involving goal management training in older adults: a randomised controlled trial. *Patient Educ. Couns.* 65, 205–213.