Journal of the International Neuropsychological Society (2009), **15**, 112–120.

doi:10.1017/S1355617708090152

Walking and talking therapy: Improving cognitive–motor dual-tasking in neurological illness

JONATHAN J. EVANS,[1,2] EVE GREENFIELD,[2] BARBARA A. WILSON,[2,3] AND ANDREW BATEMAN[2]

[1]Section of Psychological Medicine, Faculty of Medicine, University of Glasgow, Glasgow, United Kingdom
[2]Oliver Zangwill Centre for Neuropsychological Rehabilitation, Ely, United Kingdom
[3]MRC Cognition and Brain Sciences Unit, Cambridge, United Kingdom

(RECEIVED April 10, 2008; FINAL REVISION September 30, 2008; ACCEPTED September 30, 2008)

Abstract

Using randomized control trial methodology, we evaluated the effectiveness of a 5-week cognitive–motor dual-tasking training program developed to improve performance of a group of people with dual-tasking difficulties arising from acquired brain injury. Training involved twice-daily practice on exercises involving walking being combined with tasks which increased in cognitive demand over the course of the intervention. A treatment group ($n = 10$) was compared with a control group ($n = 9$). The primary outcome measure was a task requiring participants to walk and carry out a spoken sentence verification task simultaneously. Secondary outcome measures were measures of dual-tasking involving either two motor tasks or two cognitive tasks. A questionnaire measure relating to activities of daily living requiring dual-tasking was also completed. Compliance with the training program was good. We found evidence of improvement in performance on the primary outcome measure, but little evidence of generalization to other measures. There was some evidence that participants believed that their dual-tasking performance in everyday life was improved after the intervention. The study was limited in terms of sample size, was not blinded and did not control fully for therapist contact time, but has produced valuable data relating to effect sizes associated with this form of intervention.
(*JINS*, 2009, *15*, 112–120.)

Keywords: Attention, Brain injury, Rehabilitation, Cognition, Walking, Randomized controlled trial

INTRODUCTION

Cerebrovascular disease, traumatic brain injury, and other neurological illnesses frequently result in impairments in attention and concentration (Brooks & McKinlay, 1987; Manly & Mattingly, 2004). Such deficits cause major problems in everyday life and limit return to work, social, and leisure activities. Robertson et al. (1997) demonstrated that the presence of attention impairments in the acute phase after stroke was a significant predictor of longer term functional recovery—poor attention compromises new learning and hence reduces patients' ability to adapt to disabilities.

One of the three main attention systems of the brain, the selective attention system, supports dual-tasking, or "doing two things at once" (Posner & Peterson, 1990). People who have suffered brain injury frequently complain that doing two things at once is difficult or even impossible since their injury. This problem affects situations involving two cognitive tasks (e.g., monitoring a cooking pot while listening to the news) or a cognitive task and a motor task (e.g., walking and holding a conversation). Haggard and Cockburn (1998) demonstrated that, compared with healthy controls, people with head injury show a significantly greater decrement in performance when simultaneously engaged in a motor and a cognitive task when contrasted with performance on each individual task alone. Haggard et al. (2000) found a significant correlation between dual-task decrement and scores on the Barthel activities of daily living scale (Mahoney and Barthel, 1965), highlighting the importance of this problem for everyday life. Cognitive–motor dual-task impairments have also been found in people with Alzheimer's disease (Camicioli et al., 1997, 2006; Cocchini et al., 2004; Logie et al., 2004) and Parkinson's disease

Correspondence and reprint requests to: Jonathan Evans, University of Glasgow, Faculty of Medicine, Section of Psychological Medicine, Gartnavel Royal Hospital, 1055 Great Western Road, Glasgow, G12 0XH UK. E-mail: jje2k@clinmed.gla.ac.uk

(Rochester et al., 2004; Yogev et al., 2005). Lundin-Olsson et al. (1997) found that the simple measure of whether or not elderly persons (with dementia, depression, or stroke) stopped walking when a conversation was started predicted the likelihood of falling during a subsequent 6-month period.

With regard to treatment, there is preliminary evidence that attention training may improve performance on tests of attention and may show generalization to everyday behavior (Sohlberg & Mateer, 2001; Sturm et al., 2002). However, to the best of our knowledge, no studies have been undertaken where the focus of the intervention is on cognitive–motor dual-tasking. A debate within the field of cognitive rehabilitation concerns whether interventions should target impaired underlying cognitive processes or specific functional skills. In other words, if we want to improve walking and talking, can we train divided attention, using say computerized attention training tasks, and expect generalization to functional tasks such as walking and talking? Alternatively, must we specifically train walking and talking with no expectation of generalization to other tasks?

We examined whether the ability to walk and carry out additional cognitive tasks (akin to walking and talking) could be improved by a training program designed to exercise this ability. We developed a low-cost program to ensure that it could be applied in a typical rehabilitation service. Furthermore, we investigated whether there was any generalization to other tasks that require dual-tasking, but which do not involve walking. This latter issue addresses the question of whether the training program improves divided attention or dual-tasking more generally or whether any improvement (if it occurs) only occurs for tasks similar to those trained. The hypotheses we tested were as follows.

Primary Hypothesis

(1) Participants receiving the cognitive–motor dual-tasking (walking and talking) training program will improve performance on a measure of walking and talking more than participants in the control group.

Secondary Hypotheses

(2) Participants in the training group will also show greater improvement of performance on other dual-task combinations compared with those in the control condition.
(3) Participants in the training group will report a greater reduction in everyday dual-tasking difficulties compared with those in the control group.

METHOD

This study was a randomized controlled trial of a 5-week cognitive–motor dual-tasking training program.

Participants

We identified participants with impairments in cognitive–motor dual-tasking arising from neurological injury (e.g., head injury) or illness (e.g., stroke). The inclusion criteria were that participants should: (a) be between the ages of 18 and 65, (b) have been independently diagnosed as having suffered a brain injury or other neurological illness; (c) show evidence of performance at least one standard deviation below the mean on one or more of a set of tests designed to assess cognitive motor dual-tasking (the Divided Attention and Dual Tasking test battery; Evans et al., 2004); (d) self-report difficulties with dual-tasking in everyday life. Exclusion criteria included (1) presence of severe language comprehension deficit (sufficiently severe to mean that the participant could not complete baseline assessment tests); (2) presence of psychiatric disorder of sufficient severity to make it unlikely that the participant can complete cognitive training exercises regularly; (3) presence of physical disability of sufficient severity to mean that the participant could not complete a 2-min walking task.

Ethics approval for this study was granted by the Peterborough and Fenland NHS research ethics committee.

We aimed to recruit 10 patients to each group. We had no previous studies of cognitive–motor dual-tasking to estimate possible treatment effect size. However, Sturm et al. (2002) found a large effect size on a computerized measure of divided attention with 6 participants following an intensive computerized training program. We recognized that the present study might be underpowered, but given the early stage of development of this intervention we believed this was appropriate and would allow an estimate of effect size to be derived on which to base any future power calculations.

Baseline Assessment

A baseline assessment involved administration of:

(1) An estimate of premorbid intellectual level was obtained from clinical records of either the Spot the Word Test (from the Speed and Capacity of Language Processing Test; Baddeley et al., 1992) or the National Adult Reading Test (NART).

(2) Divided Attention and Dual-Tasking Battery (Evans et al., 2004; Wilson et al., 2006): There are currently no published standardized assessment tools for examining cognitive–motor dual-tasking. We, therefore, used a battery of tasks currently in development by our research group. Normative data were available from a sample of 200 healthy controls which enabled scaled scores to be derived for each of our participants. The battery has been shown to be sensitive to the effects of brain injury (Evans et al., 2004; Greenfield et al., 2004) and Alzheimer's disease (Greenfield et al., 2006). The battery contains four tasks, two motor (walking and clicking a hand-held counter) and two cognitive tasks (sentence verification and tone counting). Each task lasts 2 min and are done singly and then combined in pairs.

Walking

The first of the motor tasks is walking up and down a 5-meter track. Participants are asked to walk at a brisk pace for

2 min and the experimenter notes how many complete lengths are covered in 2 min.

Clicking

The second motor task is clicking a hand-held mechanical counter with the thumb of the dominant hand. Participants are asked to click as quickly as they can for 2 min.

Sentences

The first cognitive task is the sentence verification task which was adapted from the Speed of Comprehension Task (part of the Speed and Capacity of Language Processing Test, Baddeley et al., 1992). Participants listen to sentences and say whether each one is true or false. A true sentence would be *Birds have wings* or *Dogs have four legs* while false sentences are based on a mismatch of subject and predicate from true sentences, for example, *Birds have four legs*, *Dogs have wings*. Approximately half the sentences are true and half are false. The 2-min recording of 50 sentences is presented *via* a compact disc player.

Tone counting

The second cognitive task is tone counting (developed from Wilkins et al., 1987 and Robertson et al., 1994). Participants are asked to count a string of tones, presented *via* a compact disc. The number of tones in a string varies from 8 to 13 and the tones are delivered at unpredictable intervals. At the end of each 2-min string, a voice on the compact disc asks, "How many?" and participants give an oral response.

Each single task is combined with the other three tasks creating six dual-tasks, one motor–motor (walking and clicking), one cognitive–cognitive (sentences and tones), and four cognitive–motor (sentences and walking, tones and walking, sentences and clicking, tones and clicking). The four single tasks yield a baseline measure against which to measure the effects of these dual-tasks (the dual-task decrement). The tasks are ordered so as to minimize any effects of physical fatigue so (i.e., no successive walking or clicking tasks). Before the test is administered patients briefly practice each of the four single and six dual-tasks.

We hypothesized that the dual-task most likely to show an improvement would be "Sentences and Walking" as this most closely resembled the training program tasks. This was, therefore, our primary outcome measure. The other task combinations in the battery were used to investigate whether any improvement in performance generalized to other dual-task combinations and were considered secondary outcome measures.

Performance on tasks done under dual-task conditions was scored on the basis of decrement from single task to dual-task conditions (i.e., the extent to which performance is poorer under dual-task conditions compared with single task conditions), with age, IQ and gender matched scaled scores being derived from a normative sample of 200 healthy controls, data which were collected as part of an ongoing standardization study. Mean scaled score for healthy controls was 10, with a standard deviation of 3. Thus, the lower the scaled score, the greater the level of decrement from single to dual-task conditions. We used individual task scores (e.g., scaled score for decrement in walking from single to dual-task, scaled score for decrement in sentences from single to dual-task) and also looked at combined scores (e.g., sum of scaled scores for walking and sentences when done together in the dual-task condition).

(3) The Memory Span & Tracking Task of Baddeley et al. (1997): This pen and paper task provided an additional measure of dual-tasking ability which enabled us to further examine any generalization of the training intervention effect. The test also involves doing two tasks (crossing boxes on an A4-sized sheet of paper and doing a digit span task) under single and dual-task conditions, with the measure being the level of decrement from single to dual-task conditions. The score derived was the percentage of single task performance obtained under dual-task conditions, averaged for the box-crossing and digit span tasks.

(4) The Telephone Search with Counting subtest of the Test of Everyday Attention (Robertson et al., 1994). The test also involves doing two tasks (searching a telephone directory and counting tones) under single and dual-task conditions, with the measure being the level of decrement from single to dual-task conditions.

(5) Dual-tasking Questionnaire (Appendix 1). Information relating to everyday difficulties with dual-tasking was obtained *via* a 10-item questionnaire compiled for this study asking for ratings of frequency (using a 5-point, 0–4 scale) of difficulties with everyday tasks involving dual-tasking. The score derived was an average rating per question (i.e., total score for questionnaire/10).

Training Intervention

Participants were randomly allocated to treatment or no treatment groups. Twenty sealed envelopes were used, each containing a note specifying the group to which the participant would be allocated (10 control, 10 treatment). The envelopes were held by an administrator distant from the research worker who was recruiting participants and running the trial. As each participant was recruited the administrator selected an envelope to establish to which group that participant would be allocated.

Treatment program

The treatment program took place over a 5-week period. This length of program was selected taking into account previous studies of attention training, for example, Sturm et al. (2002) which suggested that large effect sizes could be obtained from programs of this length. The program focused on combined walking/cognitive task performance and involving gradually introducing additional cognitive demands beginning with low demand stimuli and increasing to higher demand tasks

Training involved weekly sessions with the researcher, the patient carrying out exercises between sessions. Sessions with the therapist (approximately 30 min) were used to discuss the previous week's training and to introduce the exercises for the forthcoming week. Participants were asked to do two practice sessions per day, 5 days per week, for 5 weeks. Practice sessions involved two 2-min walking sessions with a recovery period (3 to 10 min) between them. At the end of every session the subject rated their dual-task performance on the work sheet, answering the following questions: "Did you manage to dual-task? (5-point scale ranging from "almost none of the time" to "almost all of the time"); "Where was your attention most of the time?" (5-point scale ranging from "equally divided" to "all over the place"); "How difficult did you find this task? (Subject marked on a line ranging from "very difficult" to "very easy"). When participants showed no evidence of mobility problems, two or three items (e.g., chairs, potted plants) were used as obstacles to prevent the walking task becoming too easy. Over the 5-week period, the demands of the secondary cognitive task were gradually increased. Relative "demand" of the secondary cognitive tasks was based on the judgment of the project team. In week 1, the task was listening to instrumental music. In week 2, the task was listening to vocal music. In week 3, it was listening to a recording of talk based radio (news, current affairs, plays) then answering recorded questions. In week 4, the task was a verbal fluency task (letter and category tasks). In week 5, the task was to answer autobiographical questions, on different topics each session, for example, family, school, hobbies, and so on. The third and fifth week had a prerecorded tape that gave instructions, played the recorded news item or asked the autobiographical questions. In the other three sessions, the subject used a countdown timer to indicate the end of the 2 min.

When introducing the exercises, the therapist instructed the participants to "try to consciously dual-task, that is to keep listening to the music (or other task depending on week of intervention) while walking". A metaphor was used involving thinking of the two tasks as like the wheels of a wagon running along railway lines—both have to go at the same speed or the wagon will derail. The importance of giving equal attention to both tasks was emphasized. It was also anticipated that the participants' awareness of their attention (i.e., how well they were attending and whether or not they were able to divide their attention) would be enhanced by completing the daily task ratings. The main intervention was, therefore, seen as involving exercising divided attention skills, though there was an additional emphasis on enhancing meta-cognitive awareness of divided-attention.

Control Group

Control group participants were assessed on the same baseline tests as the treatment group, which were repeated at the 6-week point. In between, participants continued with any treatment as usual. To control for some of the potential non-specific treatment affects arising from therapist contact in the treatment group, the control group received a weekly phone call from the research therapist (to review general progress, enquire about difficulties or successes in dual-tasking situations). This tended to last from 5 to 10 min. They were also asked to enter, on 5 days a week for 5 weeks, a brief diary entry regarding examples of dual-tasking problems that had arisen, to simulate the awareness-raising aspect that the training condition may have produced.

RESULTS

Participants

Forty-four participants volunteered to discuss participation. Eight were excluded on the basis of meeting one or more of the exclusion criteria (e.g., physical health prevented participation) and five declined following provision of information about requirements for participation. Ten participants underwent the baseline assessment but did not meet inclusion criteria and so did not proceed to the study. One participant commenced the study and was allocated to the treatment condition, but withdrew before the training program was completed as he was unable to keep up with practice sessions for reasons unrelated to the study. One participant also commenced the training phase, but was unable to complete any of the training practice sessions independently. As she was nevertheless keen to contribute to the study she completed the second baseline assessment and her data was used in an intention-to-treat analysis. Nineteen participants met both inclusion and exclusion criteria, were randomly allocated to treatment or no treatment condition and completed the study as allocated. Therefore, 90.5% of participants initially allocated to groups completed the trial as allocated. Demographic details of participants are presented in Table 1. There were no

Table 1. Demographic details of participants

Group	Mean age (*SD* and range)	Gender	Etiology	Mean months since injury (*SD* and range)	Estimated premorbid intellectual ability
Treatment ($n = 10$)	44.4 (8.51; 32–60)	9M; 1F	7 TBI 2 CVA 1 Tumor	59.0 (54.34; 6–180)	3 above 6 average 1 below
Control ($n = 9$)	45.11 (9.73; 28–58)	8M; 1F	5 TBI 4 CVA	115.6 (107.6; 12–280)	3 above 2 average 4 below

significant differences in age ($t=0.169$; $p=.868$), estimated premorbid IQ (χ^2 (2)=3.758; $p=.153$) or time since injury ($t=1.411$; $p=.185$), though the control group were more variable in terms of time since injury than the treatment group.

Compliance with Training Program

Participants recorded whether they completed each practice session. Results showed that there was a 91% compliance (456/500 sessions completed). Of the 44 sessions missed, 25 were within week 3, with some participants reporting that they found the task involving remembering news items too challenging and hence avoided it.

Compliance with Control Group Diary Keeping Task

Participants in the control condition were asked to keep a diary of dual-tasking activities. However, compliance with this task was poor, with few diaries returned and little evidence of recording of events.

Primary and Secondary Outcome Measures

Data were analyzed using a series of 2 × 2 analyses of variance (ANOVAs), with one within subject variable (testing occasion) and one between subject variable (treatment/no treatment group) examining performance on primary and secondary outcome measures. Table 2 summarizes these data. Results are presented for all of the dual-task combinations involving walking as one of the dual-task components.

Significance (p) values were not adjusted for multiple comparisons because it was acknowledged that this was a preliminary study aimed at establishing effect sizes in relation to the measures used and because specific hypotheses were made with regard to those measures that were considered most likely to show improvement.

A primary purpose of this study was to determine effect sizes for the intervention. We report partial eta squared (η_p^2) as the measure of effect size. Partial eta squared is a measure of the proportion of variance (PV) in the dependent variable that is attributed to the effects examined in the ANOVA (i.e., main effects and interaction) in the study sample. In a treatment study, the critical effect is the interaction effect as this represents the extent to which the treatment condition has a greater effect on the dependent variable than the control condition. In the present study we, therefore, report η_p^2 values for group × time interaction effect. With regard to the effect sizes, Cohen (1988) suggested that for measures of PV explained, small, medium and large effects correspond to the values of .010, .059, and .138.

Table 2 shows that for the Sentences and Walking dual-task condition there were indications that the training program improved performance, with evidence of a large effect size, particularly when the Walking with Sentences and the Sentences with Walking data were summed. There was no indication of a treatment effect on any of the other task combinations involving walking. For the tones and walking task there was in fact a significant interaction, but this was because the control group improved while the treatment group did not. This may be because on this measure the treatment group were in fact within a normal range at baseline (scaled score 9.20) meaning there may not have been much room for improvement. By contrast the control group were poorer at baseline (scaled score 6.89), but improved to be within the normal range (scaled score 10.88). It is not clear what accounts for the baseline difference. The improvement of the control group over time may have reflected a practice effect which was less apparent in the treatment group because of their better initial performance. There were no significant group × time interaction effects for single task walking or single task sentences.

On the Dual-tasking Questionnaire there was a trend toward the treatment group showing a reduction in self-reported difficulties, while the control group showed no change. The effect size for this (partial eta squared=.162) was in fact large using Cohen's (1988) conventions, but because of the small sample size only achieved $p=.087$. When each group was examined separately, the control group showed a test–retest correlation of $r=.690$ ($p=.04$), and no evidence of a difference between test occasions ($t=-.327$; df (8); $p=.752$) suggesting the questionnaire has reasonably good reliability. For the treatment group, there was a similar correlation between test occasions (before and after the intervention) of $r=.741$ ($p=.014$), but there was a significant difference between test occasions ($t=2.641$; df (9); $p=.027$).

Analysis of the additional cognitive–motor task combinations (clicking and sentences, clicking and tones) showed no time, group or group × time interaction effects and are omitted. The results for the cognitive–cognitive combination (sentences and tones) are presented as they show that despite significant room for improvement, no significant generalization of the treatment effect was seen. This was also the case for the memory span and tracking task and the TEA dual-task.

With regard to power, for the primary outcome measures, the power for the Sentence (& Walking) interaction term was .473, for the Walking (& Sentences) interaction term was .513 and for the combined Walking and Sentences score power was .781. This latter score is in fact approaching what one would consider an acceptable level of power. For the questionnaire measure, statistical power was .402.

An intention-to-treat (ITT) analysis was completed including the two participants from whom baseline data was available and were initially allocated to the treatment group, but who did not complete the training. For one participant data was available from a second baseline assessment (this participant was unable to complete the training but was willing to undergo a second assessment as per the control group condition) so data from that assessment was included in the ITT analysis. For the other participant only initial baseline data was available and so data from the baseline assessment was carried forward. The ITT was conducted on the primary outcome measures along with the Divided Attention Questionnaire. Results are presented in Table 3 and show that large effect sizes were still evident for the primary outcome measures, although the effect size for the

Table 2. Results from a series of analyses of variance conducted on primary outcome measures (labeled[1]) and secondary outcome measures (labeled[2]), comparing pre- (T1) and posttraining (T2) scores for the Treatment and Control groups

	Treatment Group		Control Group		F values 1. Time 2. Group 3. Group × Time (η_p^2)
Measure	T1 Mean (*SD*)	T2 Mean (*SD*)	T1 Mean (*SD*)	T2 Mean (*SD*)	
[1]Sentences (& Walking)	5.30 (2.45)	9.20 (4.18)	5.88 (3.85)	6.11 (2.84)	5.06** 0.97 4.03* (.192)
[1]Walking (& Sentences)	8.60 (2.27)	10.70 (1.70)	9.56 (1.59)	9.22 (2.10)	2.35 0.147 4.46* (.208)
[1]Sentences & Walking Combined	13.90 (2.80)	19.90 (4.88)	15.44 (4.21)	15.33 (3.91)	7.83** 0.99 8.43** (.331)
[2]Tones (& Walking)	9.20 (3.39)	9.50 (2.99)	6.89 (4.88)	10.88 (1.05)	6.76** 0.04 5.00** (.160)
[2]Walking (& Tones)	8.00 (2.58)	9.60 (1.89)	8.44 (3.32)	9.22 (2.44)	3.93* 0.00 0.47 (.027)
[2]Tones & Walking combined	17.2 (4.75)	19.10 (3.67)	15.33 (6.10)	20.11 (2.42)	9.49** 1.76 1.76 (.094)
[2]Clicking (& Walking)	12.50 (2.13)	11.0 (2.61)	10.13 (2.85)	8.88 (3.83)	3.54* 0.03 0.03 (.002)
[2]Walking (& Clicking)	9.63 (1.92)	11.5 (2.20)	10.38 (3.16)	9.75 (2.56)	0.67 0.26 2.68 (.161)
[2]Clicking and Walking Combined.	22.12 (3.04)	22.5 (4.17)	20.50 (5.29)	18.63 (4.93)	0.53 1.95 1.19 (.078)
[2]Sentences (& Tones)	4.60 (3.27)	6.20 (4.83)	5.33 (3.46)	5.44 (3.32)	1.43 0.00 1.08 (.060)
[2]Tones (& Sentences)	6.40 (4.27)	8.90 (2.68)	8.78 (4.66)	9.00 (4.21)	1.50 0.72 1.05 (.058)
[2]Tones & Sentences combined	11.00 (5.62)	15.10 (6.33)	14.11 (4.83)	14.44 (5.43)	2.41 0.51 1.74 (.093)
[2]Memory Span & Tracking Task^	93.0 (7.76)	94.01 (6.73)	84.68 (9.91)	90.44 (7.83)	2.88 3.58* 1.42 (.078)
[2]TEA Telephone Search while Counting	7.2 (3.39)	7.7 (3.59)	7.89 (3.33)	7.78 (4.23)	0.06 0.07 0.15 (.009)
[2]Dual Tasking Questionnaire†	2.09 (0.68)	1.71 (0.56)	1.99 (0.80)	2.06 (0.78)	1.61 0.17 3.29* (.162)

Note. All scores are scaled decrement scores unless otherwise noted. All DivA scores are age, gender, and IQ scaled. ^Scores are performance during dual task condition as a percentage of single task performance, averaged for digit span and the box crossing tracking task. †Scores are mean scores for average questionnaire response (scale ranges from 0 to 4, with 4 indicating very frequent problems). Partial eta squared (η_p^2) is reported as measure of interaction effect size.
*$p < .10$.
**$p < .05$.

divided attention questionnaire had reduced to medium-large and was no longer significant even at the $p = .10$ level.

DISCUSSION

The results of this study suggest that a cognitive–motor training program involving walking in conjunction with a gradually increasing level of additional cognitive demand may lead to improvements in walking and talking, but there was no substantial evidence that this improvement generalizes to other cognitive–motor, cognitive–cognitive, or motor–motor task combinations. There was, therefore, support for the first of the study hypotheses, but no support for the second hypothesis. There was some indication that participants rated their performance in everyday situations as improved providing partial support for the third hypothesis, though the unblinded nature of the study design and the possibility of response bias mean that these latter subjective data should be treated with caution.

Five of the participants initially approached declined to take part after being given information about requirements for participation suggesting that a proportion of patients might not find the training program acceptable. However, for those that did participate, there was very good compliance suggesting that the intervention could realistically be implemented with this population. For those in the control condition, keeping a diary of activities requiring dual-tasking proved difficult which suggests this is not necessarily an effective method of raising awareness of dual-tasking difficulties.

One potential explanation for improvement in the training group is that through extended walking practice, walking itself became more automated, allowing more cognitive resource to be allocated to the sentences task. If this were true, one might have expected single-task walking to be improved, but this was not the case. Furthermore, the amount of practice specifically involving walking was extremely small compared with the amount participants might walk anyway as part of their normal day so it seems unlikely that this would be an explanation.

Although the main content of the training involved simply exercising dual-tasking abilities, we speculate that carrying out these exercises and daily ratings combined with the instructions to maintain attention equally to both tasks may have increased participants' awareness of how they were allocating attention during the tasks and prompted more conscious control of attention. However, as we did not test this hypothesis this remains only speculation at this stage.

Our results are consistent with those of Sturm et al. (2002), who found that domain-specific attention training was effective in improving performance in the domain trained, but that improvements did not generalize to untrained domains. In their study, six patients received divided attention training and showed improvement on tests of divided attention, but not tests of vigilance, selective attention, or alertness. Although there was no generalization across domain, Sturm et al. (2002) argued that there was generalization *within-domain*, to the extent that the tasks used for training were different to those used for the measurement of outcome. It is not clear what level of similarity is crucial to the effectiveness of the training.

In our study effect sizes for the primary outcome measure and the questionnaire measure were, in Cohen's (1988) terms, large. Cohen (1988) defined a large effect size as d > 0.8 and in the present study the effect size for the Sentences with

Table 3. Results from an intention-to-treat analysis conducted on the primary outcome measures and the Divided Attention Questionnaire, comparing pre- (T1) and posttraining (T2) scores for the Treatment and Control groups

Measure	Treatment Group T1 Mean (*SD*)	Treatment Group T2 Mean (*SD*)	Control Group T1 Mean (*SD*)	Control Group T2 Mean (*SD*)	F values: 1. Time / 2. Group / 3. Group × Time (η_p^2)
Sentences (& Walking)	5.50 (2.28)	9.25 (3.93)	5.88 (3.85)	6.11 (2.84)	5.32** 1.39 4.20* (.181)
Walking (& Sentences)	8.92 (2.27)	10.92 (2.07)	9.56 (1.59)	9.22 (2.10)	2.42 0.53 4.75** (.200)
Sentences & Walking Combined	14.42 (2.84)	20.17 (5.04)	15.44 (4.21)	15.33 (3.91)	7.75** 1.64 8.37** (.306)
Dual Tasking Questionnaire†	1.89 (0.79)	1.61 (0.56)	1.99 (0.80)	2.06 (0.78)	0.80 0.84 2.11 (.100)

Note. All scores are scaled decrement scores unless otherwise noted. ^Scores are performance during dual task condition as a percentage of single task performance, averaged for digit span and the box crossing tracking task. †Scores are mean scores for average questionnaire response (scale ranges from 0 to 4, with 4 indicating very frequent problems). Partial eta squared (η_p^2) is reported as measure of interaction effect size.
*$p < .10$.
**$p < .05$.

Walking combined measure was the equivalent of d = 1.4. Zakzanis (2001) makes the point that the interpretation of effect sizes should depend upon the context of the field of enquiry. He suggested that a Cohen's d > 3 might be considered a heuristic marker of a large effect in neuropsychological research when the emphasis is on distinguishing one group from another (e.g., whether a test distinguishes a healthy group from a group with dementia), but also notes that in treatment studies a more modest effect size may still mean that a substantial number of people benefitted from the intervention. Another way of thinking about clinical significance is the extent to which participants move from an impaired range closer to normal levels of performance, and in the present study in the case of the scores on the Sentences task while Walking, it is noteworthy that the control group were 1.37 *SD*s below a normative mean at baseline and 1.30 *SD*s below at the end of the study. By contrast, the treatment group were 1.57 *SD*s below the mean at baseline and 0.27 *SD*s below the mean posttreatment. Thus, the treatment group moved from being outside the normal range for a healthy normative sample to close to the mean for that sample.

There were several limitations to the study. The modest sample size meant that the study was underpowered such that despite some large effect sizes for some of the outcome measures, only one treatment result was significant at $p < .05$. Nevertheless, an important purpose of the study was to examine effect sizes for possible future studies. On several measures (e.g., measures of clicking and walking) the level of performance was well within a normal range for many participants and thus for some it might be considered that there was no real room for improvement.

A problem with tests of dual-tasking would appear to be low test–retest reliability. Baddeley et al. (1997) reported that the Digit Span & Box Crossing test had relatively low test–retest reliability. Data recently collected as part of a separate study examining reliability of the tests in the Divided Attention & Dual-Tasking battery (Wilson et al., 2006) have indicated that although reliability of individual test raw scores is mostly reasonably high, the correlation coefficients for the dual-task decrement scores were lower than would be hoped for in relation to test reliability. Among healthy control participants, level of decrement from single to dual-task conditions tends to be very small, which means there is relatively little variability in the scores between participants, which inevitably reduces the level of correlation (Streiner & Norman, 1995, p. 168). In the case of the Divided Attention & Dual Tasking battery tests, for some tasks (e.g., walking) the raw single and dual-task scores are highly correlated and, combined with the fact that the addition of a secondary task (e.g., clicking, sentences, or tones) has a relatively modest impact on the primary task (e.g., walking) then the result is a small range of and variance in decrement scores. There are several implications of these results. First, in the context of low reliability, a significant effect of the training intervention compared with the control condition is a positive finding suggesting that the effect of the intervention is robust. However, it is also possible that any generalization effects (i.e., improvement on secondary measures) were masked by noise (error variance) associated with the dual-task measures. This highlights the need to develop instruments for measuring dual-tasking that have good reliability, something that is important for clinical assessment of dual-tasking ability, and measurement of the effects of interventions.

The researcher conducting outcome assessments in this study was not blind to treatment condition and this should be rectified in any future study. With regard to the Dual-Tasking Questionnaire, in an attempt to minimize the possibility of response bias, participants were left questionnaires to complete independently after the last visit of the researcher, but it is quite possible that those participants who undertook the training may have responded more positively as a result of wanting to please the researcher or to justify their engagement in a relatively demanding training program. For this reason the results of the subjective questionnaire should be treated with caution.

Nonspecific effects of therapist involvement in the treatment program were controlled by including weekly telephone contact between the therapist and the control participants. However, in reality the telephone calls to control participants typically lasted no more than 10 min whereas the treatment group participants had a 30-min visit from the researcher. It is possible, therefore, that nonintervention related factors contributed to the treatment effect and, had resources allowed, it would have been better to provide a weekly half-hour session with the therapist to equate with the treatment time and mode of delivery offered to the intervention group.

The relative demand of the training tasks was judged by the project team. However, at least one task (remembering news items) was considered, by some participants, to be more demanding than other tasks that came later in the program. It would be useful to formally establish the relative demand of the secondary tasks by examining their impact on the primary motor task.

In summary, this study has provided data relating to effect sizes for an intervention specifically targeting walking and talking. Results suggest that the intervention may lead to improvement, but that any improvement may be limited to this task and not generalize to other cognitive–motor task combinations.

ACKNOWLEDGMENTS

This study was funded by the Chief Scientist's Office of the Scottish Executive Grant CZG/2/207, "Improving Cognitive-Motor Dual Tasking in Neurological Illness." The Research Sponsor for the study was the Cambridgeshire NHS Research and Development Support Team. No financial conflicts of interest exist with respect to this manuscript. The information in this manuscript is an original work that is not currently under review elsewhere and has not previously been published in any form. The authors thank all of the participants who kindly took part in this study. We also thank the editor and three anonymous reviewers for their constructive comments on the manuscript.

REFERENCES

Baddeley, A.D., Della Sala, S., Gray, C., Papagno, C., & Spinnler, H. (1997). Testing central executive functioning with a pencil and paper test. In P. Rabbitt (Ed.), *Methodology of Frontal and Executive Function*. Hove: Psychology Press.

Baddeley, A.D., Emslie, H., & Nimmo-Smith, I. (1992). *The Speed and Capacity of Language Processing (SCOLP) Test*. Bury St Edmunds: Thames Valley Test Company.

Brooks, D.N. & McKinlay, W. (1987). Return to work within the first seven years of severe head injury. *Brain Injury*, *1*, 5–15.

Camicioli, R., Bouchard, T., & Licis, L. (2006). Dual-tasks and walking fast: Relationship to extra-pyramidal signs in advanced Alzheimer's disease. *Journal of the Neurological Sciences*, *248*, 205–209.

Camicioli, R., Howieson, D., Lehman, S., & Kaye, J. (1997). Talking while walking: The effect of a dual task in aging and Alzheimer's disease. *Neurology*, *48*, 955–958.

Cocchini, G., Della Sala, S., Logie, R.H., Pagani, L., & Spinnier, H. (2004). Dual task effects of walking when talking in Alzheimer's disease. *La Revue Neurologique*, *160*, 1, 74–80.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioural Sciences* (2nd Ed). Hillsdale, NJ: Erlbaum.

Evans, J.J., Greenfield, E., & Wilson, B.A. (2004). The development of an ecologically valid test of divided attention [Abstract]. *Brain Impairment*, *5*(Suppl.), 51–52.

Greenfield, E., Evans, J.J., Emslie, H., Manly, T., Wilson, B.A., & Hodges, J.R. (2006). Cognitive-motor dual tasking in Alzheimer's disease [Abstract]. *Journal of the International Neuropsychological Society; JINS*, *12*(Suppl. 2), 49.

Greenfield, E., Evans, J.J., and Wilson, B.A. (2004). Use of a divided attention test to assess emergence from post traumatic amnesia [Abstract]. *Brain Impairment*, *5*(Suppl.), 53.

Haggard, P. & Cockburn, J. (1998). Dividing attention between cognitive and motor tasks in neurological rehabilitation. *Neuropsychological Rehabilitation*, *8*, 155–170.

Haggard, P., Cockburn, J., Cock, J., Fordham, C., & Wade, D. (2000). Interference between gait and cognitive tasks in a rehabilitating neurological population. *Journal of Neurology, Neurosurgery, and Psychiatry*, *69*, 479–486.

Logie, R.H., Della Sala, S., Cocchini, G., & Baddeley, A.D. (2004). Is there a specific executive capacity for dual task coordination? Evidence from Alzheimer's disease. *Neuropsychology*, *18*, 504–513.

Lundin-Olsson, L., Nyberg, L., & Gustafson, Y. (1997). "Stops walking when talking" as a predictor of falls in elderly people. *The Lancet*, *349*, 617.

Mahoney, F.I. & Barthel, D. (1965). Functional evaluation: The Barthel Index. *Maryland State Medical Journal*, *14*, 56–61.

Manly, T. & Mattingly, J. (2004). Visuospatial and attentional disorders. In L. Goldstein & J. McNeil (Eds.), *Clinical Neuropsychology: A Practical Guide to Assessment and Management for Clinicians*. Chichester: Wiley.

Posner, M.I. & Peterson, S.E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, *13*, 25–42.

Robertson, I.H., Ward, A., Ridgeway, V., & Nimmo-Smith, I. (1994). *Test of Everyday Attention*. London: Pearson.

Robertson, I.H., Ridgeway, V., Greenfield, E., & Parr, A. (1997). Motor recovery after stroke depends upon intact sustained attention: A two year follow-up study. *Neuropsychology*, *11*, 290–295.

Rochester, L., Hetherington, V., Jones, D., Nieuwboer, A., Willems, A.M., Kwakkel, G., & Van Wegen, E. (2004). Attending to the task: Interference effects of functional tasks on walking in Parkinson's disease and the roles of cognition, depression, fatigue, and balance. *Archives of Physical Medicine and Rehabilitation*, *85*, 1578–1585.

Sohlberg, M.M. & Mateer, C.A. (2001). *Cognitive Rehabilitation: An Integrative Neuropsychological Approach*. New York: Guilford Press.

Streiner, D. & Norman, G. (1995). *Health Measurement Scales: A Practical Guide to their Development and Use*. Oxford: Oxford University Press.

Sturm, W., Finn, B., Cantagallo, A., Cremel, N., North, P., Passadori, A., Pizzamiglio, L., Rousseaux, P., Zimmerman, P., Deloche, G., & Leclercq, M. (2002). Computerised training of specific attention deficits in stroke and traumatic brain injured patients: A multicentre efficacy study. In M. Leclerk, & P. Zimmerman (Eds.), *Applied Neuropsychology of Attention*. Hove: Psychology Press.

Wilkins, A.J., Shallice, T., & McCarthy, R. (1987). Frontal lesions and sustained attention. *Neuropsychologia*, *25*, 359–365.

Wilson, B.A., Evans, J.J., Greenfield, E., Emslie, H., Manly, T., & Bateman, A. (2006). *Identifying and Understanding Divided Attention Deficits in People with Brain Injury*. London: Department of Health, National Research Register.

Yogev, G., Peretz, C., Springer, S., Simon, E.S., & Hausdorff, J.M. (2005). Dual tasking, gait rhythmicity, and Parkinson's disease: Which aspects of gait are attention demanding? *European Journal of Neuroscience*, *22*, 1248–1256.

Zakzanis, K.K. (2001). Statistics to tell the truth, the whole truth, and nothing but the truth: Formulae, illustrative numerical examples, and heuristic interpretation of effect size analyses for neuropsychological researchers. *Archives of Clinical Neuropsychology*, *16*, 653–667.

APPENDIX 1

Dual-Tasking Questionnaire

The following questions are about problems which everyone experiences from time to time, but some of which happen more often than others. We want to know how often these things have happened to you in the past *few weeks*. There are five options, ranging from *very often* to *never, or not applicable*. Please tick the appropriate box.

Do you have any of these difficulties?

1. Paying attention to more than one thing at once.
2. Needing to stop an activity to talk.
3. Being unaware of others speaking to you when doing another activity.
4. Following or taking part in a conversation where several people are speaking at once.
5. Walking deteriorating when you are talking or listening to someone.
6. Busy thinking your own thoughts, so not noticing what is going on around you.
7. Spilling a drink when carrying it.
8. Spilling a drink when carrying it and talking at the same time.
9. Bumping into people or dropping things if doing something else as well.
10. Difficulty eating and watching television or listening to the radio at the same time.