



LJMU Research Online

Logie, RH, Trawley, S and Law, AS

Multitasking: multiple, domain-specific cognitive functions in a virtual environment

<http://researchonline.ljmu.ac.uk/id/eprint/1165/>

Article

Citation (please note it is advisable to refer to the publisher's version if you intend to cite from this work)

Logie, RH, Trawley, S and Law, AS (2011) Multitasking: multiple, domain-specific cognitive functions in a virtual environment. MEMORY & COGNITION, 39 (8). pp. 1561-1574. ISSN 0090-502X

LJMU has developed [LJMU Research Online](#) for users to access the research output of the University more effectively. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LJMU Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain.

The version presented here may differ from the published version or from the version of the record. Please see the repository URL above for details on accessing the published version and note that access may require a subscription.

For more information please contact researchonline@ljmu.ac.uk

<http://researchonline.ljmu.ac.uk/>

**Multitasking: Multiple, Domain-Specific Cognitive Functions
in a Virtual Environment**

Journal:	<i>Memory & Cognition</i>
Manuscript ID:	MC-ORIG-10-197.R2
Manuscript Type:	Resubmitted Manuscript
Date Submitted by the Author:	n/a
Complete List of Authors:	Logie, Robert; University of Edinburgh, Centre for Cognitive Ageing and Cognitive Epidemiology Trawley, Steven; University of Edinburgh, Human Cognitive Neuroscience Law, Anna; Liverpool John Moores University, Psychology
Keywords:	recall, short term memory, spatial cognition, working memory, multitasking

Multitasking: Multiple, Domain-Specific Cognitive Functions in a Virtual Environment

Robert H. Logie, Steven Trawley
University of Edinburgh, UK

Anna Law
Liverpool John Moores University

Corresponding Author:

Robert H Logie

Centre for Cognitive Ageing and Cognitive Epidemiology

University of Edinburgh

7 George Square

Edinburgh

EH89JZ

Scotland

United Kingdom

e-mail: rlogie@staffmail.ed.ac.uk

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Abstract

Multitasking among three or more different tasks is a ubiquitous requirement of everyday cognition, yet rarely is addressed in research on healthy adults who have had no specific training in multitasking skills. Participants completed a set of diverse subtasks within a simulated shopping mall and office environment, the Edinburgh Virtual Errands Test (EVET). The aim was to investigate how different cognitive functions such as planning, retrospective and prospective memory, and visuo-spatial and verbal working memory contribute to everyday multitasking. Subtasks were chosen to be diverse, and predictions were derived from a statistical model of everyday multitasking impairments associated with frontal lobe lesions (Burgess et al., 2000). Multiple regression indicated significant independent contributions from measures of retrospective memory, visuo-spatial working memory and on-line planning, but not from independent measures of prospective memory or verbal working memory. Structural Equation Modelling showed that the best fit to the data arose from three underlying constructs, with Memory and Planning having a weak link, but with both having a strong directional pathway to an Intent construct that reflected implementation of intentions. Participants who followed their pre-prepared plan achieved higher scores than those who altered their plan during multitask performance. This was true regardless of whether the plan was efficient or poor. Results substantially develop and extend the Burgess et al. (2000) model to healthy adults, and yield new insight into the poorly understood area of everyday multitasking. Findings also point to the utility of using virtual environments for investigating this form of complex human cognition.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
Everyday multitasking among three or more different tasks pervades domestic and working life, yet rarely is the focus of empirical study or theory development. This form of multitasking refers to the management of several different on-going tasks to be completed within a limited time. Tasks often must be interleaved efficiently to maximise performance (Burgess, 2000), and there may be an optimum order for their completion, for example cooking a meal (e.g. Craik & Bialystok, 2006) or a time-limited shopping trip (e.g. Shallice & Burgess, 1991). Many occupations require multitasking, for example emergency medicine and medical decision making (Chisholm, Dornfeld, Nelson & Cordell, 2001; Law et al., 2005; van der Meulen et al., 2010), management (Seshadri & Shapira, 2001), or navigation (e.g. Garden, Cornoldi & Logie, 2002; Law, Logie & Pearson, 2006; Spiers & Maguire, 2006). We report here an investigation of how healthy young adults achieve everyday multitasking in a simulated environment. We start with a statistical model of everyday multitasking derived from studies of individuals with frontal lobe damage (Burgess, Veitch, de Lacy Costello & Shallice, 2000). We assess the generality of this model, and develop it for healthy young adults using a new paradigm designed to address this ubiquitous requirement of everyday cognition.

33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
Most current theories of cognition address specific components of the cognitive system such as visual or auditory perception, focused or sustained attention, prospective memory, verbal or visuo-spatial working memory, episodic and semantic memory, task switching and cognitive bottlenecks. These topics are experimentally tractable and draw on well developed paradigms, allowing major advances in the understanding of each. Everyday multitasking requires co-ordinated and strategic deployment of several different cognitive functions, presenting a major challenge for paradigms and theories that focus on specific cognitive phenomena, and ensuring its rarity in studies of human cognition (Neisser, 1978). So, the research we describe is not concerned with understanding a particular cognitive function, or with cognitive bottlenecks between conflicting tasks (e.g. Borst, Taatgen & Rijn, 2010b). Rather, it addresses how the cognitive system avoids cognitive bottlenecks to perform multiple tasks (e.g. Craik & Bialystok, 2006). This requires new paradigms and development of theories concerned with how various parts of the cognitive system operate in concert rather than in isolation.

Studies of multitasking typically focus on deficits following brain injury (e.g., Alderman, Burgess, Knight & Henman, 2003; Miotto & Morris, 1998; Shallice & Burgess, 1991), on the

1
2
3 neuroanatomical correlates in healthy participants (Borst, Taatgen, Stocco & van Rijn, 2010a;
4 Burgess, Dumontheil & Gilbert, 2007; Spiers & Maguire, 2006), or on highly trained experts
5 such as in the military or aviation (e.g. Loukopoulos, Desmukes & Barshi, 2009; Wickens,
6 2008). Everyday multitasking in untrained adults is different from the microstructure of rapid
7 switching between laboratory tasks (e.g. Koch, Gade, Schuch & Philipp, 2010; Monsell,
8 2003), or from concurrent dual-task demands (e.g. Borst et al., 2010a; 2010b; Logie,
9 Cocchini, Della Sala & Baddeley, 2004a). Instead, it involves several subtasks that have
10 different requirements, and participants decide themselves how to schedule subtask attempts.
11 Other studies have assessed driving skills (e.g. Levy & Pashler, 2008; Strayer, Drews &
12 Crouch, 2006), or planning and implementation of subgoals such as in the Tower of London
13 task or simulated work (e.g. Hambrick, Oswald, Darowski, Rench & Brou, 2010; Phillips,
14 Gilhooly, Logie, Della Sala & Wynn, 2003; Ward & Allport, 1997). However, these tasks do
15 not address the broader demands of everyday multitasking, and despite its ubiquitous
16 everyday requirement, there remains limited theoretical or empirical insight into how
17 multitasking is achieved by healthy adults.
18
19
20
21
22
23
24
25
26
27
28
29
30

31
32 One theoretical approach to multitasking (Burgess et al., 2000) was derived from studies of
33 individuals with frontal lobe damage, who often show impairments in everyday multitasking
34 but intact performance on tests of attention, memory and executive functions. This suggests
35 multitasking is not solely dependent on the latter functions. Burgess et al. (2000) used the
36 'Greenwich test', in which participants switch between three manual sub-tasks (sorting beads,
37 sorting tangled lines on paper, and constructing plastic Meccano). Their statistical model
38 (Figure 1) identified important, and largely independent roles for retrospective memory and
39 intentionality (prospective memory). Planning did not reliably contribute to the model.
40 However, planning deficits only appeared for patients with lesions in the right dorsolateral
41 prefrontal cortex. Damage to anterior regions, Brodmann areas 8, 9 and 10, did not affect
42 planning, but did affect task switching and task rule adherence. For these reasons, Burgess et
43 al. included planning in their structural equation model, but they had only one measure of
44 planning, so loading for the 'Plan' construct was set at '1.0'. Their model suggested that the
45 'Memory' construct drives separate constructs for 'Plan' and 'Intent'.
46
47
48
49
50
51
52
53
54
55
56
57

58 Figure 1 about here
59
60

1
2
3 Burgess, Simons, Coates and Channon (2005) suggested that planning is multifaceted and
4 supported by a range of cognitive abilities. Phillips et al. (2003) noted that many healthy
5 participants plan on-line during complex tasks rather than following a plan formed in advance.
6
7 Therefore it would also be important to identify the individual cognitive functions that
8 contribute to advance and on-line planning. Burgess et al. argued that forming and
9 implementing a plan draws upon cognitive systems responsible for remembering instructions
10 and rules for the multitasking paradigm. Therefore, planning and intentionality were
11 represented downstream from retrospective memory in their model.
12
13
14
15
16
17
18

19 The roles for advance and on-line planning may be greater for subtasks with an optimum
20 order, such as preparing a meal. Craik and Bialystok (2006) addressed this in a cognitive
21 aging study using simulated breakfast making. Participants repeatedly set a simulated table
22 by moving cutlery and plates on the computer screen, switching to alternate screens for
23 starting and stopping preparation of foods with different cooking times. However, the
24 reliance on prospective memory for starting and stopping foods, and similarity of sub-tasks,
25 make it less well suited for multitasking assessment, which we assume to involve a range of
26 different cognitive functions, only one of which is prospective memory.
27
28
29
30
31
32
33

34
35 The Multiple Errands Test (MET) (Shallice and Burgess, 1991; Alderman et al., 2003), has
36 an optimum task order. Individuals with frontal lobe damage and healthy controls planned
37 and attempted several tasks in a real shopping mall, following as efficient a route as possible.
38 Although the MET is close to real life, and requires little or no initial practice, there are
39 obvious drawbacks in real-life settings (e.g. Bailey, Henry, Rendell, Phillips & Kliegel, 2010).
40 It is time consuming, requires transport for participants, and consent from local businesses.
41 The lack of experimental control can compromise participant safety and data reliability, and
42 tasks cannot easily be adapted for other clinical or research settings.
43
44
45
46
47
48
49

50
51 Simulated real-life tasks have been used for assessing brain damaged patients, including
52 planning on a map (e.g. Burgess et al. 2005), moving furniture in virtual buildings (Morris,
53 Kotitsa & Bramham, 2005), or selecting stores in a video of a high street (e.g. Knight, Titov
54 & Crawford, 2006). Knight and Titov (2009) and others (e.g. Burgess et al., 2005; Bailey et
55 al., 2010) note the need for ecologically valid and scientifically robust tests of multitasking.
56 McGeorge et al. (2001) created a virtual version of the MET, the Virtual Errands Test (VET)
57 that retained many advantages of the real environment while achieving experimental control.
58
59
60

1
2
3 The errands were tasks such as collect a book or meet a colleague. The VET was as sensitive
4 as the real environment to executive dysfunction in brain damaged patients (see also Rand,
5 Basha-Abu Rukan, Weiss & Katz, 2009). Thus, virtual environments may offer sufficient
6 face and content validity for assessment of multitasking in a tractable and controlled setting.
7 However, these studies have focused on impairments associated with brain damage (Knight
8 & Titov, 2009). Law et al. (2006) modified the VET to be challenging for healthy adults.
9 However, the graphics were unrealistic, and performance assessment involved video
10 recording test sessions with subsequent manual scoring. The Edinburgh Virtual Errands Test
11 (EVET), used in the current study, was developed from a widely available and inexpensive
12 commercial games platform that permits non-profit development of virtual environments. It is
13 well suited to creating an environment for multitasking with realistic graphics and a smooth
14 interface as well as automatic recording of multiple performance measures.
15
16
17
18
19
20
21
22
23
24

25
26 In sum, little is known about how healthy young adults behave in everyday multitasking
27 situations, and what cognitive variables affect their performance. We addressed the cognition
28 that underlies this important activity, using sets of multiple errands performed by healthy
29 young adults in a realistic simulation of everyday multitasking. EVET involves a higher
30 memory load than did Burgess et al. (2000), because participants have to memorise an errand
31 list. EVET requires a substantial degree of pre-planning, because inefficient ordering of the
32 errands may result in time expiring before all tasks are completed. It requires navigation
33 around a virtual environment, so participants may draw on the resources of visuo-spatial
34 working memory (Logie, 1995). Therefore we expected that the Burgess et al. model might
35 require modification to account for multitasking performance in healthy participants.
36
37
38
39
40
41
42
43
44

45
46 One additional important aspect of cognition is working memory, and König et al. (2005; see
47 also Hambrick et al., 2010) found that working memory was a more important predictor of
48 multitasking performance than fluid intelligence or attention. Baddeley and Logie (1999;
49 Logie, in press) argued that working memory comprises multiple, domain-specific resources
50 that are deployed selectively by participants according to task demands. This contrasts with
51 the view that working memory is primarily a control system for focusing limited capacity
52 domain-general attention on currently activated contents of episodic or semantic memory (e.g.
53 Cowan, 2005). Therefore, we used independent measures of verbal and spatial working
54 memory, and of verbal free recall to explore whether a single or a multiple factor model
55 could best account for our multitasking data.
56
57
58
59
60

1
2
3
4
5 Participants completed an additional test of planning within EVET, and a battery of cognitive
6 assessments to investigate relationships between individual differences in multitasking and
7 separate measures of retrospective and prospective memory, on-line planning, and spatial and
8 verbal working memory. From the Burgess et al. (2000) model and from our own previous
9 work (e.g. Garden et al., 2002; Law et al., 2006; Logie, Baddeley, Mane, Donchin & Sheptak,
10 1989), we expected that retrospective memory and prospective memory, but also on-line
11 planning (rather than pre-planning), spatial working memory and verbal working memory
12 would make independent contributions to shared variance with EVET performance.
13
14
15
16
17
18
19
20

21 In summary, our goals were:

- 22
23
24 (a) To add to understanding of everyday multitasking in young healthy adults.
25
26
27
28 (b) To test the generality in healthy young adults of a statistical model originally developed to
29 account for multitasking impairments associated with frontal lobe damage.
30
31
32
33 (c) To test the hypothesis that multitasking among several diverse subtasks draws on a range
34 of several different cognitive functions, not only a single, general purpose attentional capacity.
35
36
37
38
39 (d) To develop a new methodology to address the above given that most current paradigms in
40 experimental cognitive psychology would be ill suited to these goals.
41
42
43
44
45
46
47
48

49 Method

50 *Participants*

51 A total of 165 students at the University of Edinburgh (102 women, 63 men), mean age 19.59
52 (sd 2.43, range 16-32) participated.
53
54
55

56 *Tests and Procedure*

57 Participants first completed the EVET procedure and then five individual difference measures,
58 administered in the order described below. Except for word recall, tasks were viewed on a 42
59
60

1
2
3 cm color monitor, on a Dell XPS PC with Intel Core Quad 2.33Ghz processor and 1GB ATI
4 Graphics Card.
5
6
7

8
9 *The Edinburgh Virtual Errands Task* (EVET) environment was created using the Hammer
10 environment editor supplied as a software development kit with the computer game Half Life
11 2™ to construct a 3D model of a four storey building with five rooms along the left and right
12 of each floor around a central stairwell, with two sets of stairs (one left, one right) and a
13 central elevator. Figure 2 shows a screen shot of the concourse on the ground floor (floor
14 zero).
15
16
17
18
19

20
21 Figure 2 about here
22
23

24
25 Participants explored the virtual environment using the standard keyboard - w, s, a, d - for
26 forward, backward, and lateral movement, and the mouse to look in any direction. Key “e”
27 was used for actions such as picking up objects or opening doors. Participant movement
28 within the virtual building was automatically recorded at 10Hz, represented as a series of
29 XYZ coordinates, with actions recorded and time stamped. Participants were to complete
30 eight errands within eight minutes (Table 1). Three of the errands were two-stage, requiring
31 object collection and drop off. The remaining five required one action. Two tasks had time
32 constraints (e.g. turn off cinema at 5:30). Sorting folders was the only open ended EVET task.
33 It could be performed at any time for as long as the participant chose during the eight minute
34 test period. Participants were informed that folder sorting was no more important than other
35 tasks, but they should try to sort as many as possible. Half of the participants were given
36 errand list A and started on the ground floor between the stairwells. The other half were given
37 errand list B and started in the equivalent position on the top floor. Errands were given in a
38 non-optimal order for completion, and participants were asked to plan the optimal order
39 before commencing EVET.
40
41
42
43
44
45
46
47
48
49
50
51

52
53 Table 1 about here
54
55

56
57 First, each participant read EVET instructions detailing the task, building layout and rules.
58 The experimenter described the layout of the building prior to the practice session. Building
59 rules were to use left stairs for travelling down and right stairs for travelling up, to avoid
60 entering non-task related rooms and to avoid picking up non-task related objects. Participants

1
2
3 practiced using the controls to move around the building to complete five practice errands;
4 object collection and delivery, button pressing, unlocking stairwell door with key-code, and
5 folder sorting. This also allowed participants to familiarise themselves with the building.
6
7
8 None of the practice errands were used in the main testing session. The practice session took
9
10 approximately five minutes.
11

12
13
14 Next, participants studied their allocated errand list for two minutes followed by free recall,
15 then five minutes of further study and a test of cued recall. The total number of errands
16 correct from free and cued recall comprised the errand list memory score, equivalent to a
17 measure classified as 'Learn' by Burgess et al. (2000). Finally, participants were provided
18 with a schematic building map and a copy of the errand list, and were asked to indicate the
19 order in which they planned to perform each errand to achieve maximum efficiency in task
20 completion. They were informed that they could change their plan during the actual test.
21
22 Upon completion of their plan, the task list was removed along with their written plan, and
23 they were asked again to verbally recall the errand list and building rules with any mistakes
24 corrected. This process was repeated until 100% recall of the list. This last procedure was to
25 minimise the chances that participants would fail to complete errands in EVET because they
26 could not recall all of the errands.
27
28
29
30
31
32
33
34
35

36
37 Participants next performed the EVET test for 8 minutes (neither task list nor plan were
38 present during the test). On completion, they were asked to recall what errands they had
39 attempted, or failed to complete, and any building rules they had broken (see Table 2). This
40 assessed whether participants could recall all of the actions they performed in EVET and is
41 equivalent to the measure labelled 'Recount' by Burgess et al (2000). It was followed by free
42 recall of the complete errand list regardless of whether or not all of the errands had been
43 completed. Participants were then cued about any errands that they had omitted in this post-
44 test free recall. A point was awarded for each errand correctly recalled by both methods to
45 generate a measure equivalent to the "remember" variable in Burgess et al (2000). Finally, all
46 participants were presented with the alternative errand list (set B if they had used set A and
47 vice versa) with a fresh sheet showing the layout of each floor, and were asked to generate
48 another plan for the order of errands in this alternative list. This was used as a second
49 measure of planning within the context of EVET, but without performing EVET a second
50 time. "EVET travel time" indicated the total amount of time each participant spent travelling
51 in the EVET building, but excluded time spent inside rooms. This measure was intended to
52
53
54
55
56
57
58
59
60

1
2
3 index efficiency of navigation through the building. Finally, the EVET score indicated errand
4 completion efficiency. Points were awarded for errand completion and were removed for
5 breaking the building rules (Table 2).
6
7
8
9

10 *The Word Recall Task* required participants to recall orally, in any order, twelve words read
11 out by the experimenter, one word per second following the standardised procedure in
12 Capitani, Della Sala, Logie and Spinnler (1992). This was repeated for five different lists (60
13 maximum). This was used as an independent measure of retrospective memory.
14
15
16
17

18
19 *Working Memory Verbal Span* was based on working memory sentence span from Baddeley,
20 Logie, Nimmo-Smith and Brereton (1985) and Duff and Logie (2001). Participants verified
21 each of a set of sentences as they appeared consecutively on a computer monitor, and were to
22 memorize the last word of each sentence. At the end of the set of sentences they were asked
23 to orally recall, in order, the sentence-final words. The task began with sequences of two
24 sentences, with each set size repeated three times, after which the set size increased by one.
25 All participants continued, regardless of performance, until the maximum set size (seven
26 sentences) was completed. All sentences were presented for three seconds, preceded by a
27 fixation cross for one second. Total correct recall of the sentence-final words was calculated
28 as a proportion of maximum possible (81).
29
30
31
32
33
34
35
36
37

38
39 *Working Memory Spatial Span* was based on a task from Shah and Miyake, 1996).
40 Participants verified whether block capital letters that appeared consecutively were shown in
41 their normal configuration or as a mirror image. Each letter was shown in a different
42 orientation within a circular area. Participants were instructed to memorize the orientation of
43 each letter for recall at the end of each set. The task began with sequences of two letters, with
44 each set size repeated three times, after which the set size increased by one. All participants
45 continued, regardless of performance, until the maximum set size (five letters) was completed.
46 All letters were presented for 3 seconds, and preceded by a fixation cross for one second.
47 Total correct recall was calculated as a proportion of maximum recall score (70).
48
49
50
51
52
53
54
55

56 *The Travelling Salesman Task (TSP)* involved presenting nine colored target shapes along the
57 bottom of the screen, with the first shape labelled "Start/End" and the rest "Target Locations".
58 The nine target shapes for each trial were also placed in the main section of the screen at
59 random locations in a 5 x 5 array, and different colored shapes were placed as distracters in
60

1
2
3 the remaining 16 locations. Participants were asked to click on each of the target locations in
4 the order that formed the shortest possible path connecting the start shape with the locations
5 of each of the targets in the array, finally returning to the start location. As each target was
6 clicked in the array it was marked to show which targets in the array had been visited, and it
7 disappeared from the target list at the bottom of the screen to indicate which targets had yet to
8 be visited in the array. When all nine target locations had been visited and the participant
9 returned to the start point, the next array was presented, using a different selection of nine
10 target shapes from the full set of 25, and with the targets shown in new random locations in
11 the array. There were two practice trials, both with nine different targets, but the first showed
12 only an array of the nine targets in random locations, and no distracters. The second showed
13 nine targets along with 16 distracters within the full 5x5 array. Optimum distance for each
14 array was calculated using an algorithm for travelling salesman problems (Kirk, 2007). Each
15 participant completed a total of ten arrays. There was only one possible optimum solution for
16 each combination of array and target set, and the average proportion of distance longer than
17 the optimum was calculated for each array completed, with the average proportions across the
18 ten arrays comprising the score.
19
20
21
22
23
24
25
26
27
28
29
30
31
32

33 *The Breakfast Task* was used as a measure of prospective memory ability. It was devised by
34 Craik and Bialystok (2006) who kindly provided a copy of the programme. In this task
35 participants had to swap between setting plates and cutlery on a virtual table on the computer
36 screen, and starting and stopping the virtual cooking of eggs, coffee, pancakes, sausages and
37 toast, each with different cooking times ranging from 120 seconds to 330 seconds. The goal
38 was for all foods to complete cooking at the same time. The five foods were each shown on a
39 separate screen from each other and from the table being set. To start cooking, the participant
40 clicked on an icon for the food with the longest cooking time (330 seconds). This took them
41 to a screen for that food where they clicked to start a timer displaying the progression of
42 cooking time for that food. They then had to return to setting the virtual table until it was time
43 to start the food with the next longest cooking time (240 seconds). This continued until the
44 time when all the foods should be ready and the participant had to visit each screen to stop
45 cooking each food. Participants practiced with two foods. Craik and Bialystok reported a
46 range of different measures of cooking performance, which, in our data were highly
47 correlated. Our chosen outcome measure was the average deviation between the actual start
48 time for each food and the time that each food should have started. Given that the task
49 primarily involved prospective memory for starting each of the foods at the correct time
50
51
52
53
54
55
56
57
58
59
60

1
2
3 while engaged in another task (table setting), it was taken as an independent measure of
4 prospective memory ability. We chose this task rather than established laboratory measures
5 of prospective memory, because it was a simulation of everyday prospective memory, and
6 our overall goal was to investigate simulations of everyday complex cognition in multitasking.
7
8
9

10 11 12 Results

13
14
15
16 Twelve participants were unable to complete either the breakfast or travelling salesman task
17 due to technical problems. Their data were excluded, leaving a final total of 153 participants
18 (95 women and 58 men) for subsequent analysis.
19
20
21

22
23 The EVET is a completely novel task and a novel paradigm. There are therefore no clear
24 precedents for generating scores to be used as dependent measures. We describe below the
25 rationale for, and the procedure followed to generate scores. Our aim was to generate
26 indicator variables that were as close as possible to those considered by Burgess et al. (2000),
27 with additional variables that were intended to assess the contribution to multitasking of
28 additional cognitive functions that were not considered in the earlier study.
29
30
31
32
33

34 35 *EVET Score*

36
37 The overall *EVET Score* comprised a weighted score procedure devised to emphasise task
38 efficiency. Following Burgess et al. (2000), the general principle was to award points for
39 tasks completed, and remove them for rule breaks. Errands that could be completed at any
40 time were awarded one point for each successful action (maximum two points for two-part
41 errands), yielding a maximum action score of 8. Finding the door code and unlocking the
42 stair door counted as one action for this purpose. The time-restricted (cinema and meeting)
43 and open ended (folder sort) tasks were weighted on a five point scale (0, 1, 2, 3 and 4),
44 yielding a possible maximum bonus of 12 points, which were added to the action score for a
45 potential overall maximum of 20 points. There were no obvious a-priori criteria for allocating
46 bonus points. Rather than generate arbitrary criteria, we based the allocation on inspection of
47 the frequency distribution of raw scores on each of these measures (Table 2). The rationale
48 was to generate a distribution of scores that fairly reflected the distribution of actual
49 performance across participants on each task. We acknowledge that the precise cut-off
50 criteria for allocating scores to adjacent bonus categories remain somewhat arbitrary, but
51 given that the same criteria were applied across all participants, we have no grounds for
52
53
54
55
56
57
58
59
60

1
2
3 suspecting that this procedure resulted in systematic bias in the overall EVET score, or
4 distorted its use as a measure of individual differences.
5
6
7

8
9 A similar scoring procedure was used for breaking EVET rules, except minus points were
10 allocated for going up the down stairs and vice versa, and for going into rooms or picking up
11 objects that were not part of the errand list. We allocated minus points (0, -1,-2,-3,-4) on the
12 basis of inspection of the frequency distributions of each error type (Table 2), and to fairly
13 reflect the error performance across participants. For example, over 80% of participants broke
14 the stair rule once, and so this attracted zero penalty. Across the three error types, the
15 maximum possible penalty score was -12, which was combined with the action+bonus score
16 for successful performance of errands in EVET. A negative score was possible if participants
17 failed to completed most of the errands and incurred a large penalty score, although this never
18 occurred within our participant sample.
19
20
21
22
23
24
25
26
27

28 Table 2 about here
29
30
31

32 *EVET learn score ("Learn")*. This is a sum of the free and cued recall scores that participants
33 took prior to starting the EVET test, based on number of points for each errand. For example,
34 "Pickup computer in G4 and take to T7" contained three elements (the item, collection point,
35 and delivery point), so was worth 3 points. For free recall 23 points were available and for
36 cued recall 14 points (1 point was removed from each errand due to the cue) making a
37 possible maximum of 37 points.
38
39
40
41
42
43

44 *EVET Plan Efficiency: ("Plan")*. The score for plan efficiency was different from that used
45 for the Greenwich test (Burgess et al., 2000). There are more possible permutations of order
46 for EVET sub-tasks, and, unlike the Greenwich test, inefficient ordering may result in failure
47 to complete all errands within the time limit. Assessment of efficiency of individual EVET
48 plans involved comparison with an optimum plan for each set of errands. There were no clear
49 a-priori criteria for identifying the optimum plan. Therefore we examined the rank order in
50 which errands were actually completed by each of the five highest scoring participants
51 (scoring 19 or 20) for each errand list. The rank order correlations between these five
52 individuals ranged from 0.791 to 0.955. We therefore used the average rank order as the
53 optimum plan for each list against which to compare the efficiency of the order in which
54 errands were completed by individual participants. For each errand order for each participant,
55
56
57
58
59
60

1
2
3 one point was awarded for each errand that was in the same serial position as in the optimum
4 plan. If there was a mismatch of serial position, then a point was awarded if a pair of errands
5 were in the same sequential order as on the optimum plan, even if, as a pair of errands, they
6 were completed earlier or later in the sequence. This ensured that there was credit given for
7 partial use of the planned errand sequence. The maximum possible overlap score was 11.
8
9 Each individual overlap score was divided by 11 to derive a plan efficiency score. We
10 defined “EVET pre-test plan” as the efficiency score for the plan participants made before
11 they attempted the EVET, and defined “EVET post-test plan” as the efficiency score of the
12 plan that participants made at the end of the EVET procedure for the alternative errand set.

13
14
15
16
17
18
19
20
21 *EVET follow score (“Follow”).* This was designed to measure the correspondence between
22 pretest plan and the order in which participants actually completed the errands. It was based
23 on allocating one point for each errand that was completed in the same sequential position as
24 planned, and a point if a pair of errands were in the same sequential order but not in the same
25 overall sequential position as on the pre-test plan, following an analogous scoring procedure
26 to that for the plan efficiency score. Each individual score was divided by the total possible
27 points that could have been gained from the number of tasks actually completed by each
28 individual participant to index how closely participants followed their initial plan.

29
30
31
32
33
34
35
36
37 *Recall EVET actions (“Recount”).* After completion of EVET, a point was awarded for each
38 errand or type of rule break that participants recalled actually carrying out; equivalent to the
39 Burgess et al. (2000) 'Recount' variable. Only the number of rule break types was recorded.
40 For example, if the participant had gone down the 'up' stair 5 times, they were asked to recall
41 if this had happened on at least one occasion. The maximum possible score was 26.
42
43
44
45
46

47
48 *EVET remember score (“Remember”).* Here, the scoring procedure for the measure “Learn”
49 was repeated but was based on free and cued recall of the errand list after completing the
50 EVET.
51
52

53
54
55 The six measures above were designed to be comparable to those considered by Burgess et al.
56 (2000): Score, Learn, Plan, Plan follow, Recount, Remember. We included additional
57 measures that we hypothesised would address the broader theoretical constructs that might be
58 incorporated within or added to the Burgess model (EVET travel time, verbal and spatial
59 working memory span, travelling salesman task, EVET post test plan, word recall, and
60

1
2
3 breakfast task). The two planning measures (EVET pre-test plan and EVET post-test plan),
4 were intended to allow free estimation of the error variance of the planning construct
5 indicators in our structural equation model. This was not possible for Burgess et al. (2000)
6 who included only one indicator of planning.
7
8
9

10
11
12 Scores were converted to percentages for analysis. Descriptive statistics and intercorrelations
13 are shown in Table 3.
14
15
16

17
18 Table 3 about here
19
20

21 Analyses

22
23 First we examined which of the five independent measures contributed unique variance to
24 prediction of EVET score, using multiple regression with backwards stepwise elimination
25 (Table 4). As expected from the Burgess et al. (2000) model, retrospective memory (word
26 recall) was a significant predictor of EVET score. However, unlike the earlier model, our
27 measure of prospective memory (breakfast task) did not reliably share variance with EVET,
28 nor did it correlate with any other variables. Like the breakfast task, EVET involves unique
29 items for each of the errands, and so it is not straightforward to calculate split half reliability
30 across equivalent test items. However, Table 3 shows significant correlations between
31 several different measures of EVET - Travel Time, Learn, Recount, Remember and Plan
32 Follow, indicating a reasonable level of internal consistency. EVET Score also correlates
33 with established independent measures of word recall, verbal and spatial working memory.
34
35
36
37
38
39
40
41
42

43
44 Our independent measure of planning ability (TSP) was a significant predictor, and so unlike
45 Burgess et al., (2000), we included planning in the model on statistical grounds rather than
46 inferring its contribution indirectly from associations with brain lesion sites. Unlike our other
47 independent measures, the TSP was a novel task, that allowed calculation of internal
48 reliability. This revealed a Chronbach Alpha of 0.737. Spatial working memory span was also
49 a reliable predictor, indicating that it made a contribution that was independent of word recall,
50 and of planning.
51
52
53
54
55
56

57
58 Verbal working memory did not make an independent contribution in this initial analysis.
59 Although Table 3 shows that this measure correlated with verbal free recall, when verbal
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

working memory was forced into the regression equation before verbal free recall, the verbal working memory measure did not share significant variance with EVET (Table 4 notes).

Insert Table 4 about here

To further examine the role of planning in EVET performance, we explored the partial correlations for the efficiency measures of the two pre-test plans relative to the optimum plan, the plan follow score, and the EVET score. Controlling for plan following, a significant correlation was found between pre-test plan efficiency and EVET performance (partial $r = 0.16$, $p=0.04$), but a stronger correlation was found between EVET performance and the efficiency with which participants *followed* their plan (partial $r = .32$, $p<.01$), controlling for pre-test plan efficiency. Including the interaction between pre-test plan efficiency and plan following added no significant contribution to the prediction of EVET score, so these appear to be separate main effects.

A graphical demonstration of the importance of plan-following can be seen in Figure 3 that illustrates the length of time participants spent at each specific set of x,y,z co-ordinates in the 3-D space. Based on the plan following measure, participants were split into upper ($N = 38$) and lower quartiles ($N = 37$) and their movements across all four floors during the eight minute EVET test were characterized using kernel density estimation. In the figure, a peak indicates that participants remained in that particular position for a period of time, whereas the absence of a peak indicates movement through the virtual building. It is clear from Figure 3A that participants who did not follow their plan spent more time travelling along the building corridors (lower peaks). Figure 3B, in contrast, shows that participants who adhered to their plan had a strong tendency to focus their time in the folder sorting room (high peak on lower right side) with corridor time and movement kept to a minimum.

Insert Figure 3 about here

An exploratory factor analysis confirmed our theoretical expectations in identifying the three factors that were included in the Burgess et al. (2000) model, labelled here as 'Memory', 'Plan' and 'Intent'. It was clear that the breakfast task had no relationship with any of the factors so could not be used as an indicator. The EVET-learn measure was not easily identified with a single factor (first and second factor loadings, .39 and .34 respectively) so was also excluded.

1
2
3
4
5 Structural equation modelling (SEM) with maximum likelihood estimation was carried out
6 using EQS 6.1 (Bentler, 2004) to assess the fit of the Burgess et al. model to the EVET data.
7 The following model fit indices were used: χ^2 (Bollen, 1989), which tests the hypothesis that
8 an unconstrained model fits the covariance/variance matrix better than the proposed model
9 (non-significant values indicate good fit); Bentler's (1990) comparative fit index (CFI),
10 which compares the proposed model with the null model that assumes all variables are
11 uncorrelated (values between .90 and .95 are acceptable); root mean square error of
12 approximation (RMSEA) is a measure of the closeness of fit, with values below .08
13 indicating reasonable model fit and below .05 a good fit.
14
15
16
17
18
19
20
21
22

23 The Burgess et al. structural framework (Figure 1), positioned 'Memory' upstream from both
24 'Plan' and 'Intent'. Fit indices indicated a reasonable fit to our data: χ^2 ($df = 42, N = 153$) =
25 74.09, $p = < 0.01$, CFI = .87 RMSEA = .07 (.04 - .10). Modification indices were used to
26 evaluate possible model changes towards improved model fit. On the basis of these
27 modification index recommendations and theoretical relevance, a path between 'Plan' and
28 'Intent' was included, which was not possible for Burgess et al. (2000), given that they had
29 only one index for planning. This second structural model (Figure 4), therefore assumed that
30 the 'Intent' factor was predicted by both 'Memory' and 'Plan' factors, with a weak link
31 between 'Memory' and 'Plan'. Fit indices suggested a better fit for this second model: χ^2 ($df =$
32 $41, N = 153$) = 64.91, $p = < 0.01$, CFI = .90 RMSEA = .06 (.03 - .09). A chi-square difference
33 test, ($\chi^2_{\text{Model 1}} - \chi^2_{\text{Model 2}}, df_{\text{Model 1}} - df_{\text{Model 2}}$), showed a significant difference between
34 the models, χ^2 ($df = 1$) = 9.18, $p = 0.01$. This suggested that the second model offered a
35 significantly different and better fit, hence supporting the important relationship between
36 'Plan' and 'Intent' latent variables, and between 'Memory' and 'Intent', but with 'Memory' and
37 'Planning' only weakly related.
38
39
40
41
42
43
44
45
46
47
48
49
50

51 Figure 4 about here
52
53
54

55 DISCUSSION

56 Our primary goal was to investigate the cognitive factors that contribute to a realistic
57 simulation of everyday multitasking. This was intended to address a major lacuna in
58 understanding of complex cognition in which a several cognitive functions are required to
59 operate in a co-ordinated fashion, rather than focus on each function in isolation. We drew on
60

1
2
3 previous research on cognitive impairments that are sequellae of frontal lobe damage, and
4 developed a novel paradigm using a virtual environment (EVET) and novel behavioral
5 measures to investigate the general cognitive principles of multitasking in the healthy young
6 adult brain.
7
8
9

10
11
12 The only existing statistical model of everyday multitasking (Burgess et al., 2000) was used
13 as an initial theoretical framework. However, that framework was based on a set of table-top
14 tasks carried out by individuals with frontal lobe damage and older controls, with a limited
15 range of measures and with only one measure of planning. This model was broadly useful in
16 explaining our data from simulated multiple-errand multitasking with young healthy adults,
17 but required modification. We identified separate constructs for memory, pre-planning, and
18 plan implementation. This suggests that memory and pre-planning may involve different
19 cognitive functions, both of which are required for this form of multitasking in young,
20 healthy adults. This separation between memory and planning also is consistent with the
21 finding by Shallice and Burgess (1991) and by Burgess et al. (2000) that planning can be
22 selectively impaired in individuals with frontal lobe damage, while leaving other cognitive
23 functions, including memory, intact. Further, our results are consistent with the Burgess et al.
24 (2005) suggestion that planning is not a unitary construct, and we identified separate
25 constructs for pre-planning and for on-line planning during task performance. However, these
26 latter two factors were not uncorrelated. Participants who formed an efficient plan also tended
27 to achieve a higher EVET score. In addition, participants who closely followed their plans
28 during EVET performance also tended to achieve higher scores than those who modified
29 their plans on-line. This was true regardless of whether their initial plan was poor or was
30 close to the optimum. Given that errors driven by a change in plan would arise after a plan
31 change, this suggests that changing a plan on-line in an attempt to make it more efficient at
32 the point of the change might have disrupted the plan for performance of errands yet to be
33 completed.
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

51
52
53 Our final model comprised three factors: 'Memory', 'Plan' and 'Intent', as did the Burgess et al.
54 (2000) model. However, for the latter, the 'Plan' construct was added on neuroanatomical
55 grounds, and 'Memory' was thought to drive 'Plan' and 'Intent' as separate constructs. Also, all
56 of the tasks in the Greewich test were visible, and could be performed in any order. In the
57 current study with the Edinburgh Virtual Errands Test, the 'Plan' construct was added on
58 statistical grounds, and a better fit was obtained when 'Memory' did not drive the 'Plan'
59
60

1
2
3 construct, but when both drove the 'Intent' construct for which there were several indicator
4 variables. Moreover, participants were required to memorise a list of errands and create a
5 plan of the order for completion, intended to simulate the everyday requirement of scheduling
6 a list of tasks for completion. These operations might be expected to require resources from
7 memory and planning, so the separate sets of indicator variables for 'Memory' and for 'Plan'
8 suggest that these two constructs reflect different cognitive functions that act together, but
9 have little mutual dependence. EVET required participants to maintain their internal list of
10 delayed intentions (plan), and then realise those intentions within the virtual environment.
11 Presumably this would require frequent consultation of the plan. Clearly, high scoring
12 participants followed their plans rather than deviating from them on-line. This argument is
13 consistent with the Model in Figure 4, which illustrates strong links between the INTENT
14 construct and the indicator variables 'EVET remember', 'EVET plan follow' and 'EVET
15 recount'. 'EVET remember' refers to the ability to remember, after EVET completion, the
16 errand list, whereas a high score on 'EVET plan follow' requires participants to keep track of
17 which errands had been completed, and to remember to perform the remaining errands in the
18 planned order. So, a combination of good memory for the errands, coupled with consultation
19 and updating of the current representation of the position on the errand list both appear to
20 contribute to the 'INTENT' construct. A high score on 'EVET recount', suggests that
21 participants had an accurate memory for actions they had actually performed during EVET,
22 so they could recall what errands had actually been completed and which had not. This is
23 consistent with the idea that they could update a representation of which errands had been
24 completed from their planned sequence, and which had not as they were performing EVET.
25 We would then expect that participants who effectively consulted and followed their plans
26 would have shorter EVET travel times with fewer deviations from the planned route, hence
27 the negative loading for the 'EVET travel time', and would achieve more errands as a result
28 (EVET score). These considerations suggest a factor that reflects plan following should draw
29 on resources of memory and preplanning. This is apparent in the final model that includes a
30 path between 'Plan' and 'Intent'. This was possible because we included two estimators for
31 planning.

32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57 Also intriguing was that the Travelling Salesman Problem loaded on 'Intent' rather than 'Plan'.
58 Our model therefore suggests it is an indicator variable for plan implementation. A detailed
59 analysis of the TSP task is beyond the scope of this paper. However, previous studies of the
60 Tower of London task, and of chess have shown that most planning for complex tasks is on-

1
2
3 line rather than in advance (e.g. Phillips et al., 2001; Saariluoma, 1995). Yet, we found that
4 sticking with an original plan resulted in better scores than modifying plans during EVET
5 performance. It is possible that participants created a planned order for the errands, but then
6 did not pre-plan the actual moves between errands. For example when moving around the
7 virtual building, they might have found shorter, or more efficient routes between the locations
8 they had to visit for each errand, so the errand order remained as planned, but the movements
9 between errands were modified on-line. This would also be consistent with EVET travel time
10 loading on the INTENT variable.
11
12
13
14
15
16
17

18
19 The model we propose does not exclude the possibility that other models, based on
20 alternative theoretical rationales, might offer a better fit. However, the results support for our
21 suggestion that multitasking requires the operation of different cognitive functions acting in
22 concert. One possible account, driven by our overall theoretical rationale is that different
23 cognitive functions are required to support performance, and they tend to correlate because
24 they act together to achieve a common goal, not because they reflect the operation of a single
25 construct such as a general attentional resource. This adds to our confidence that use of
26 complex paradigms such as EVET, together with a theoretical rationale that assumes the co-
27 ordinated operation of multiple cognitive functions to achieve task goals, can generate robust
28 and interpretable data to address the complex cognition that supports everyday multitasking.
29
30
31
32
33
34
35
36
37

38
39 The multiple regression analysis suggested that overall EVET score was predicted by
40 independent measures of retrospective memory, of visuo-spatial planning (travelling
41 salesman problem), and of spatial working memory. It may be that a participant's ability to
42 manipulate spatial information is a key factor for this particular type of multitasking. This
43 may be less important in the Greenwich test, where sub-tasks are in full view of the
44 participants. No additional unique variance in EVET score was explained by an independent
45 measure of prospective memory (simulated breakfast making) or verbal working memory.
46 The lack of a contribution from verbal working memory appeared to be due, at least in part to
47 co-linearity with the measure of retrospective memory, suggesting that these two measures
48 reflect a common, domain-specific cognitive function, that is different from visuo-spatial
49 working memory and planning, and consistent with the multicomponent framework for
50 working memory (Baddeley & Logie, 1999; Logie, in press).
51
52
53
54
55
56
57
58
59
60

1
2
3 It is striking that performance on the Breakfast Task (Craik & Bialystok, 2006) was unrelated
4 to overall EVET performance. Craik and Bialystok demonstrated that it showed clear effects
5 of cognitive ageing and could detect an advantage for older bilingual compared with older
6 monolingual individuals. Participants have to monitor the progress of a number of foods that
7 cook at different rates, and stop them cooking at the appropriate time. This is very different
8 from a situation where a participant has a specific, planned order in which they are attempting
9 a memorised list of quite diverse tasks, and this could be a crucial difference between EVET
10 and the breakfast task. It may also be the case that young healthy participants correct their
11 prospective memory errors by returning to complete 'forgotten errands' within the eight
12 minute period. In this case, the EVET score might be insensitive to individual variation in
13 prospective memory performance within EVET. However, this is speculative, and the general
14 issue of prospective memory in multitasking using EVET clearly merits further investigation.
15 However, detailed consideration of this issue is beyond the scope of the current paper, and
16 has been addressed elsewhere (Trawley, Law & Logie, in press).
17
18
19
20
21
22
23
24
25
26
27
28
29

30 A possible criticism is that the novelty of EVET also comes with uncertainty as to its
31 reliability. This is an issue with any novel paradigm, particularly when it is more complex
32 than is common in studies of human cognition. However, if we are to understand how human
33 cognition deals with everyday complexity, then the experimental paradigms have to be
34 complex as well as robust. The results of the current study illustrate the utility of exploring
35 this form of complex cognition using multiple errands in a virtual environment, retaining a
36 high degree of experimental control together with a degree of realism. Some confidence in
37 the reliability of our paradigm comes from the degree of internal consistency between the
38 different measures of EVET performance, and from the robust relationships with more
39 established measures of retrospective memory and working memory. However, further
40 assessment of its reliability would be useful to address in future studies.
41
42
43
44
45
46
47
48
49
50

51 One possible limitation of the implementation of EVET in this study is that the results might
52 not generalise to other scenarios. However, with minor modifications to the procedure EVET
53 could be used to explore a wide range of research questions, for example when multiple
54 errands are carried out without pre-planning, or when plans are interrupted. We have
55 completed experiments of this kind that will be reported elsewhere. Unlike the original
56 multiple errands test (Shallice and Burgess, 1991), the same environment can be used in a
57
58
59
60

1
2
3 range of different laboratories, and the environment and the data extraction utilities are
4
5 available for free¹ for not-for-profit research on request from the authors.
6
7

8
9 The study has developed and demonstrated the utility of a novel methodology to study
10 everyday multitasking. It has added insight into how different cognitive functions act in
11 concert to achieve complex cognitive goals, rather than studying each function in isolation.
12 There are new findings reported regarding the relationship between pre-planning and plan
13 following, and the relationship between three constructs (Memory, Planning and Intent) in
14 multitasking by healthy young adults, thereby substantially developing research that has
15 focused previously on the cognitive impact of frontal lobe damage. The study has also
16 demonstrated that this form of complex cognition can successfully be addressed and thereby
17 yield insights into how human cognition can meet and manage multiple requirements of daily
18 living.
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56

57
58 ¹ Setting up and running the EVET environment requires researchers to purchase a software licence for the
59 widely available computer game Half-Life 2. The software development kit licence that comes with the game
60 allows not-for-profit creation of virtual environments. The data extraction utilities require a local licence for
MATLAB. A detailed technical description of the EVET and data extraction utilities is given in Trawley, Law,
Logie and Logie (under review).

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Acknowledgements

We are grateful to Matthew Logie for undertaking the programming to create the EVET environment. We also are grateful to Fergus Craik for permission to use the Craik and Bialystok (2006) breakfast task in our research and to their programmer Perry Tohn for providing a copy of the programme and for help with its implementation in our laboratory. We acknowledge Leverhulme Trust research grant number F/00 158/W which supported the research that we report here.

References

- 1
2
3
4
5
6
7 Alderman, N., Burgess, P.W., Knight, C., & Henman, C. (2003). Ecological validity of a simplified
8 version of the multiple errands shopping test. *Journal of the International*
9 *Neuropsychological Society*, 9, 31-44.
- 10
11
12 Baddeley, A.D. & Logie, R.H. (1999). Working memory: The multiple component model. In A.
13 Miyake & P. Shah (eds.) *Models of Working Memory*, pp28-61. New York: Cambridge
14 University Press.
- 15
16
17 Baddeley, A.D., Logie, R.H., Nimmo-Smith, I & Brereton, N. (1985) Components of fluent reading.
18 *Journal of Memory and Language* 24, 119-131.
- 19
20
21 Bailey, P.E, Henry, J.D., Rendell, P.G., Phillips, L.H., & Kliegel, M. (2010). Dismantling the “age-
22 prospective memory paradox”: The classic laboratory paradigm simulated in a naturalistic
23 setting. *Quarterly Journal of Experimental Psychology*.
- 24
25
26 Bollen, K.A. (1989). *Structural Equations with Latent Variables*. New York: John Wiley.
- 27
28 Bentler, P.M., (1990). Bentler, Fit indices, LaGrange multipliers, constraint changes, and incomplete
29 data in structural models, *Multivariate Behavioral Research* 25 (1990), pp. 163–172.
- 30
31 Bentler, P.M., (2004). EQS for Windows, 6.1 Encino, CA: Multivariate Software, Encino, CA..
- 32
33 Borst, J.P., Taatgen, N.A., Stocco, A. & van Rijn, H. (2010). The Neural Correlates of Problem
34 States: Testing fMRI Predictions of a Computational Model of Multitasking. *PLoS ONE* 5(9):
35 e12966. doi:10.1371/journal.pone.0012966
- 36
37
38 Borst, J.P., Taatgen, N.A. & van Rijn, H. (2010). The Problem State: A Cognitive Bottleneck in
39 Multitasking. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 36,
40 363–382.
- 41
42
43 Burgess, P.W. (2000). Real-world multitasking from a cognitive neuroscience perspective. *Attention*
44 *and Performance*, 18, 465–472.
- 45
46
47 Burgess, P.W., Alderman, N., Forbes, C., Costello, A., Coates, L. M-A., Dawson, D.R. (2006),
48 Anderson, N. D., Gilbert, S. J., Dumontheil, I., & Channon, S. The case for the development
49 and use of "ecologically valid" measures of executive functions in experimental and clinical
50 neuropsychology. *Journal of International Neuropsychological Society* 12, 194–209.
- 51
52
53 Burgess, P.W., Dumontheil, I., & Gilbert, S. (2007). The gateway hypothesis of rostral prefrontal
54 cortex (area 10) function. *Trends in Cognitive Science*, 11, 290-298.
- 55
56
57 Burgess, P.W., Simons, J.S., Coates, L.M.-A., & Channon, S. (2005). The search for specific
58 planning processes. In R. Morris & G. Ward (Eds.), *The Cognitive Psychology of Planning*
59 (pp. 199–227). Hove, U.K.: Psychology Press.
- 60

- 1
2
3 Burgess, P.W., Veitch, E., de Lacy Costello, A. & Shallice, T., (2000).The cognitive and
4 neuroanatomical correlates of multitasking. *Neuropsychologia*, 38, 848–863.
5
6 Capitani, E., Della Sala, S., Logie, R., & Spinnler, H. (1992). Recency, primacy and memory:
7 Reappraising and standardising the serial position curve. *Cortex*, 28, 315-342.
8
9 Chisholm, C.D., Dornfeld, A.M., Nelson, D.R., & Cordell, W.H. (2001). Work interrupted: A
10 comparison of workplace interruptions in emergency departments and primary care offices.
11 *Annals of Emergency Medicine*, 38, 146–151.
12
13 Cohen, J. & Cohen, P. (1983). *Applied Multiple Regression/Correlation Analyses for the Behavioral*
14 *Sciences*. Hillsdale, NJ, USA: Lawrence Erlbaum Associates.
15
16 Craik, F.I., & Bialystok, E. (2006). Planning and task management in older adults: cooking breakfast.
17 *Memory & Cognition*, 34, 1236-1249.
18
19 Duff, S.C. & Logie, R.H. (2001). Processing and storage in working memory span. *Quarterly*
20 *Journal of Experimental Psychology*, 54A, 31-48.
21
22 Garden, S., Cornoldi, C. & Logie, R.H. (2002).Visuo-spatial working memory in navigation. *Applied*
23 *Cognitive Psychology*, 16, 35-50.
24
25 Hambrick, D.Z., Oswald, F.L., Darowski, E.S., Rench, T.A. & Brou, R. (2010). Predictors of
26 Multitasking Performance in a Synthetic Work Paradigm. *Applied Cognitive Psychology* 24,
27 1149–1167.
28
29 Imbo, I. &LeFevre, J.-A. (2010). The role of phonological and visual working memory in complex
30 arithmetic for Chinese- and Canadian-educated adults. *Memory and Cognition*, 38 (2), 176-
31 185.
32
33 Kirk, J. (2007) Traveling Salesman Problem - Genetic Algorithm. Retrieved from Matlab File
34 Exchange website: [http://www.mathworks.com/matlabcentral/fileexchange/13680-traveling-](http://www.mathworks.com/matlabcentral/fileexchange/13680-traveling-salesman-problem-genetic-algorithm)
35 [salesman-problem-genetic-algorithm](http://www.mathworks.com/matlabcentral/fileexchange/13680-traveling-salesman-problem-genetic-algorithm).
36
37 Knight, R.G., & Titov, N. (2009). Use of virtual reality tasks to assess prospective memory:
38 Applicability and evidence. *Brain Impairment*, 10, 3-13.
39
40 Knight, R.G., Titov, N., & Crawford, M. (2006).The effects of distraction on prospective
41 remembering following Traumatic Brain Injury assessed in a simulated naturalistic
42 environment. *Journal of the International Neuropsychological Society*, 12, 8–16.
43
44 Koch, I., Gade, M., Schuch, S. & Philipp, A.M. (2010). The role of inhibition in task switching: A
45 review. *Psychonomic Bulletin and Review*, 17, 1-14.
46
47 König, C.J., Bühner, M. & Mürling, G. (2005). Working Memory, Fluid Intelligence, and Attention
48 Are Predictors of Multitasking Performance, but Polychronicity and Extraversion Are not.
49 *Human Performance*, 18, 243–266.
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 Law, A. S., Freer, Y., Hunter, J., Logie, R. H., McIntosh, N., & Quinn, J. (2005). A comparison of
4 graphical and textual presentations of time series data to support medical decision making in
5 the Neonatal Intensive Care Unit. *Journal of Clinical Monitoring and Computing*, *19*, 183-
6 194.
7
8
9
10 Law, A., Logie, R.H., & Pearson, D.G. (2006). The impact of secondary tasks on multitasking in a
11 virtual environment. *Acta Psychologica*, *122*, 27-44.
12
13
14 Levy, J., Pashler, H. (2008) Task prioritization in multitasking during driving: Opportunity to abort a
15 concurrent task does not insulate braking responses from dual-task slowing. *Applied*
16 *Cognitive Psychology*, *22*, 507-525.
17
18
19 Logie, R.H. (1995). *Visuo-Spatial Working Memory*. Hove, UK: Lawrence Erlbaum Associates.
20
21 Logie, R.H. and Baddeley, A.D. (in press). Working Memory: State of the Science. *Current*
22 *Directions in Psychological Science*.
23
24
25 Logie, R.H., Baddeley, A.D., Mane, A., Donchin, E. and Sheptak, R. (1989). Working memory and
26 the analysis of a complex skill by secondary task methodology. *Acta Psychologica*, *71*, 53-87.
27
28 Logie, R.H., Cocchini, G., Della Sala, S., & Baddeley, A.D. (2004a). Is there a specific executive
29 capacity for dual task co-ordination? Evidence from Alzheimer's Disease. *Neuropsychology*,
30 *18*, 504-513.
31
32
33 Logie, R.H., Gilhooly, K.J. & Wynn, V. (1994b). Counting on working memory in arithmetic
34 problem solving. *Memory and Cognition* *22*, 395-410.
35
36
37 Loukopoulos, L.D., Dismukes, K. & Barshi, I. (2009). *The Multitasking Myth: Handling Complexity*
38 *in Real-World Operations*. Farnham, UK: Ashgate.
39
40
41 McGeorge, P., Phillips, L.H., Crawford, J.R., Garden, S.E., Della Sala, S., & Milne, A.B. (2001).
42 Using virtual environments in the assessment of executive dysfunction. *Presence*, *10*, 375-383.
43
44 Miotto, E.C., & Morris, R.G. (1998). Virtual planning in patients with frontal lobe lesions. *Cortex*, *34*,
45 631-657.
46
47
48 Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, *7*, 134-140.
49
50 Morris, R., Kotitsa, M. & Bramham, J. (2005). Planning in patients with focal brain damage: From
51 simple to complex task performance. In R. Morris & G. Ward (Eds.), *The Cognitive*
52 *Psychology of Planning* (pp. 153-179). Hove, U.K.: Psychology Press.
53
54
55 Morris, R., & Ward, G. (2005). *The Cognitive Psychology of Planning*. Hove, UK: Psychology Press.
56
57 Neisser, U. (1978). Memory: What are the important questions? In M.M. Gruneberg, P.E. Morris and
58 R.N. Sykes (Eds.) *Practical Aspects of Memory*. London: Academic Press.
59
60
61 Phillips, L.H., Gilhooly, K.J., Logie, R.H., Della Sala, S. & Wynn, V. (2003). Age, working memory,
62 and the Tower of London task. *European Journal of Cognitive Psychology*, *15*, 291-312.

- 1
2
3 Phillips, L.H., Wynn, V.E., McPherson, S. & Gilhooly, K.J. (2001). Mental planning and the Tower
4 of London task. *Quarterly Journal of Experimental Psychology*, 54A, 579–597.
5
6
7 Rand, D., Basha-Abu Rukan, S., Weiss, P. L., & Katz, N. (2009). Validation of the virtual MET as
8 an assessment tool for executive functions. *Neuropsychological Rehabilitation*, 19, 583-602.
9
10 Saariluoma, P. (1995). *Chess Players' Thinking*. London: Routledge.
11
12 Seshadri, S. & Shapira, Z. (2001). Managerial Allocation of Time and Effort: The Effects of
13 Interruptions. *Management Science*, 47, 647-662.
14
15 Shah, P., Miyake, A. (1996) The Separability of Working Memory Resources for Spatial Thinking
16 and Language Processing: An Individual Differences Approach. *Journal of Experimental*
17 *Psychology: General*, 125, 4-27.
18
19 Shallice, T., & Burgess, P. (1991). Deficits in strategy application following frontal lobe damage in
20 man. *Brain*, 114, 727-741.
21
22 Spiers, H.J. & Maguire, E.A. (2006). Thoughts, behaviour, and brain dynamics during navigation in
23 the real world. *Neuroimage*, 31, 1826-1840.
24
25 Strayer, D.L., Drews, F.A., & Crouch, D.J. (2006). Comparing the cellphone driver and the drunk
26 driver. *Human Factors*, 48, 381-391.
27
28 Subrahmanyam, K., & Greenfield, P.M. (1994). Effect of video game practice on spatial skills in
29 girls and boys. *Journal of Applied Developmental Psychology*, 15, 13-32.
30
31 Trawley, S.L., Law, A.S. & Logie, R.H. (in press). Event based prospective remembering in a virtual
32 world. *Quarterly Journal of Experimental Psychology*.
33
34 Trawley, S.L., Law, A.S. & Logie, M.R. & Logie, R.H. (under review). Virtual reality for a realistic
35 price: methodology for using the Source 3D game engine in psychological research.
36
37 Van der Meulen, M., Logie, R.H., Freer, Y., Sykes, C., McIntosh, N. & Hunter, J. (2010). When a
38 Graph is Poorer than 100 Words: A Comparison of Computerised Natural Language
39 Generation, Human Generated Descriptions and Graphical Displays in Neonatal Intensive
40 Care. *Applied Cognitive Psychology*, 24, 77-89.
41
42 Ward, G., & Allport, A. (1997). Planning and problem-solving using the 5-disc Tower of London task.
43 *Quarterly Journal of Experimental Psychology*, 50, 49–78.
44
45 Wickens, C.D. (2008). Multiple Resources and Mental Workload. *Human Factors*, 50, 449–455.
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1. EVET errand lists (A & B).

Errand List	A	B
1	Pickup Brown Package in T4 and take to G6	Pickup Computer in G4 and take to T7
2	Pickup Newspaper in G3 and take to Desk in S4	Pickup Milk Carton in T3 and take to Desk in F4
3	Get Keycard in F9 and unlock G6 (via G5)	Get Keycard in S9 and unlock T7 (via T6)
4	Meet person S10 before 3:00 minutes	Meet person F10 before 3:00 minutes
5	Get stair-code from notice board in G8 and unlock stairwell	stair-code from notice board in T10 and unlock stairwell
6	Turn on Cinema S7 at 5:30 minutes	Turn on Cinema F7 at 5:30 minutes
7	Turn off Lift G Floor	Turn off Lift T Floor
8	Sort red and blue binders in room S2. Sort as many binders as you can.	Sort red and blue binders in room F2. Sort as many binders as you can.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 2. Number of bonus points added, and penalty points deducted for EVET score based on percentage of participants performing within the ranges shown in the table.

Bonus points added	+4	+3	+2	+1	0
Number of folders sorted	30+ (4%)	23-29 (11%)	15-22 (16%)	8-14 (21%)	1-7 (46%)
Cinema (absolute time discrepancy in seconds from 5:30 min)	0-2 (40%)	3-5 (10%)	6-7 (7%)	8-10 (5%)	11+ (4%)
Meeting (time discrepancy in seconds over 3:00min)	<=3mins (28%)	1-12 (9%)	13-25 (11%)	26-37 (6%)	38+ (24%)
Penalty points deducted	-4	-3	-2	-1	0
Number of objects picked up that were not on the errand list	4+ (1%)	3 (1%)	2 (6%)	1 (33%)	0 (59%)
Number of rooms entered that were not on the errand list	4+ (12%)	3 (5%)	2 (7%)	1 (15%)	0 (62%)
Number of times stair rule broken	5+ (10%)	4 (2%)	3 (7%)	2 (14%)	1 (67%)

Table 3. Descriptive Statistics and Correlation matrix of EVET performance and predictive measures

	M	SD	2	3	4	5	6	7	8	09	10	11	12	13
1 EVET score (%)	49.30	28.46	.27**	.29**	.23**	-.32**	-.02	-.44**	.46**	.27**	.46**	.23**	.23**	.35**
2 Word recall (%)	48.65	7.66		.44**	.13	-.02	.02	-.11	.27**	.07	.07	.17*	.22**	.09
3 Working memory verbal span (%)	82.85	16.09			.29**	-.18*	.09	-.14	.27**	.17*	.06	.02	.01	-.06
4 Working memory spatial span (%)	73.19	24.23				-.09	.02	-.19*	.18*	.08	.06	.00	.06	-.06
5 Travelling salesman problem (% above optimum distance)	10.10	6.42					.06	.29**	-.15	-.12	-.20*	-.08	-.20*	-.20**
6 Breakfast task (seconds)	16.12	14.75						-.01	.05	-.12	-.02	-.00	-.08	.04
7 EVET travel time (seconds)	305.10	42.08							-.20*	-.17*	-.27**	-.02	-.15	-.37**
8 EVET learn (%)	65.51	15.49								.06	.48**	.04	.18*	.04
9 EVET recount (%)	81.96	9.35									.22**	.13	-.04	.25**
10 EVET remember (%)	89.46	13.03										.10	.09	.28**
11 EVET pretest plan (%)	46.36	18.59											.36**	.23**
12 EVET posttest plan (%)	40.92	17.63												.22**
13 EVET plan follow (%)	62.29	24.09												

**p < 0.01 (2-tailed).

*p < 0.05 (2-tailed).

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

Table 4. Results of multiple regression with backwards stepwise elimination to assess the contribution to common variance between scores on the Edinburgh Virtual Errands Test and scores on five different measures of mental ability as described in the text.

Variable	<i>B</i>	<i>SE B</i>	<i>Beta</i>	<i>t</i>	<i>p</i>
Travelling Salesman	23.746	6.707	.27	3.54	<.001
Word Recall	.30	.09	.25	3.32	<.001
Spatial working memory span	.05	.02	.15	1.96	<.055

$F=10.69, df=3, 152, p<.001 R^2=.18$
 Model selection procedure: backwards stepwise elimination.
 Excluded (Not Significant): Prospective memory task.
 Verbal working memory (VWM) did not make a significant additional contribution.
 When VWM was forced into the regression equation ahead of word recall,
 values for VWM were: $B=0.05, SE B=0.04, Beta=0.11, t=1.24, p=0.22$
 See text for explanation.

Figure Captions

Figure 1. An illustration of the structural equation model of multitasking from Burgess et al. (2000), redrawn from the original with permission.

Figure 2. Screen shot of the ground floor concourse area of the Edinburgh Virtual Errands Test (EVET)

Figure 3. Kernel density estimates of participant movement for lower (A) and upper (B) quartile plan followers.

Figure 4. A proposed model of EVET multitasking. Selected model fit indices: χ^2 ($df = 41, N = 153$) = 64.91, $p < 0.01$, CFI = .90 RMSEA = .06 (.03-.9), AIC = -17.09.

Figure 1.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

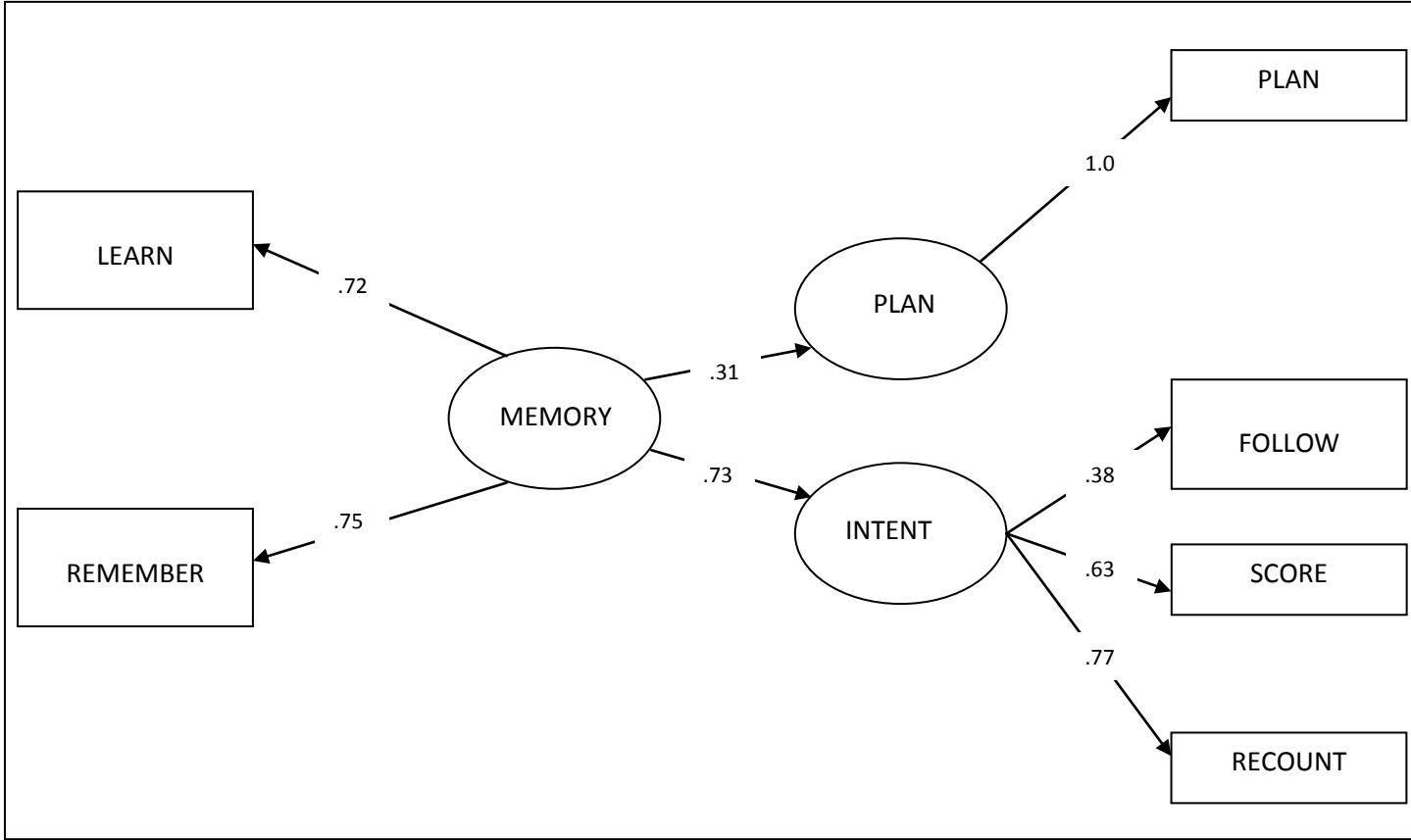


Figure 2.



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

Figure 3

Kernel density map, Lower quartile plan followers, 3D visualization

Kernel density map, Upper quartile plan followers, 3D visualization

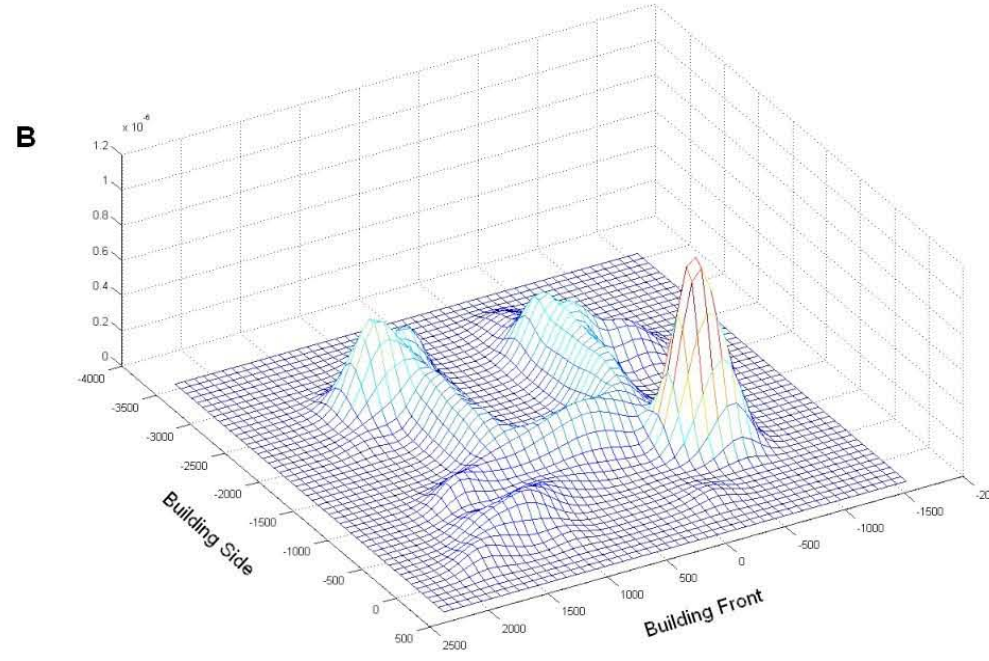
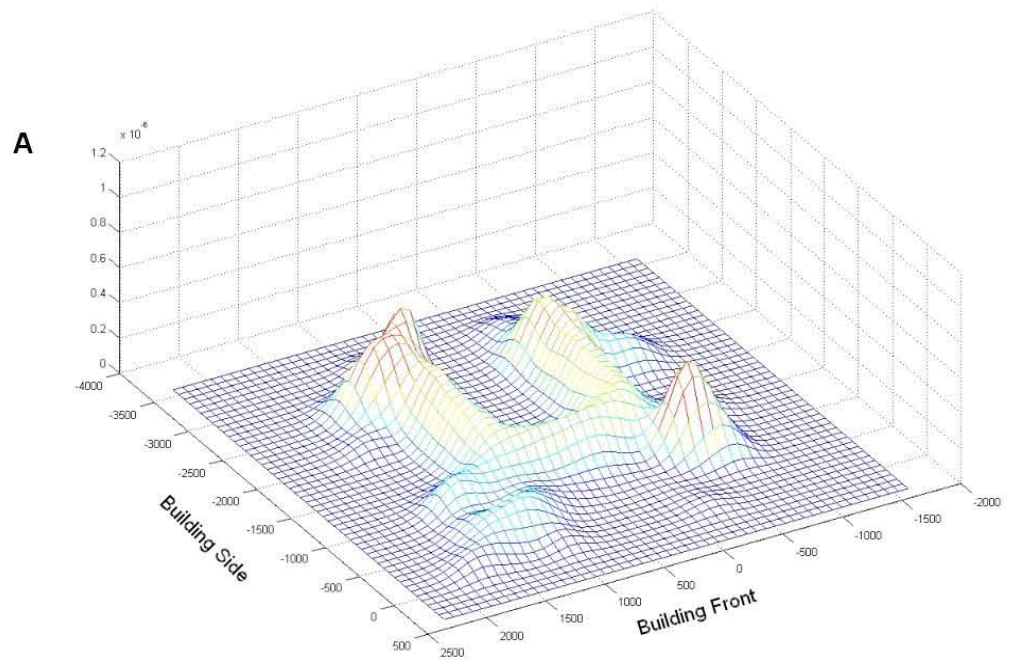
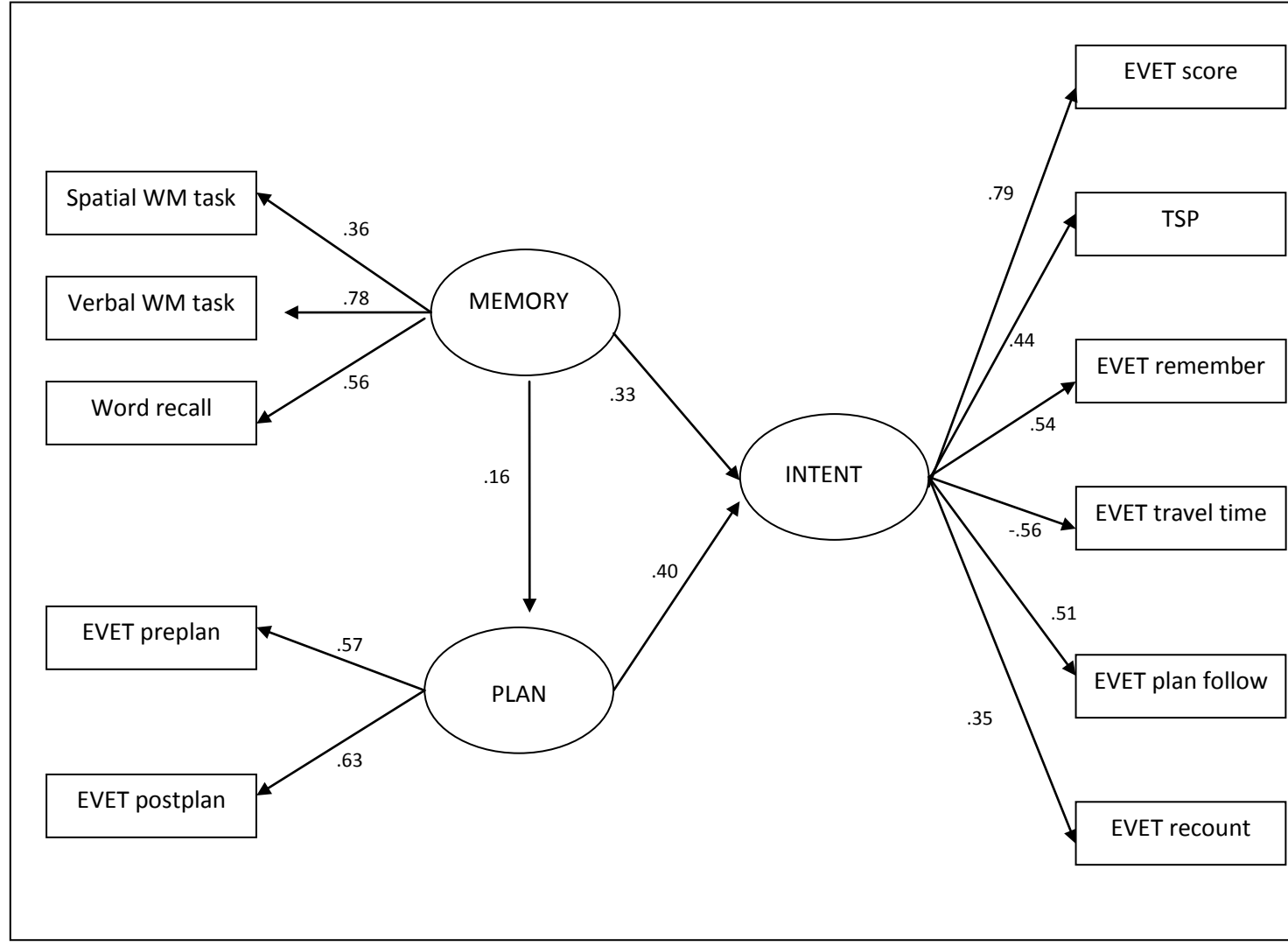


Figure 4.



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47