Naturalistic action impairments in dementia

Tania Giovannetti a, David J. Libon b,c,d, Laurel J. Buxbaum a,e, Myrna F. Schwartz a,f

a Moss Rehabilitation Research Institute, Philadelphia, PA 19141, USA
b Crozer Chester Medical Center, Upland, PA 19013, USA
c Hahnemann University, Philadelphia, PA 19102, USA
d Drexel University, Philadelphia, PA 19104, USA
e Thomas Jefferson University Medical School, Philadelphia, PA 19107, USA
f Temple University, Philadelphia, PA 19122, USA

Received 2 November 2000; received in revised form 30 October 2001; accepted 16 November 2001

Abstract

Naturalistic actions are everyday tasks (e.g. cooking) that require one to use multiple objects and sequence steps to achieve a goal. Naturalistic action impairment has been attributed to executive dysfunction [Higher cortical functions in man. New York: Basic Books, 1966], semantic knowledge degradation [Brain 111 (1988) 1173], and, more recently, general limitations in cognitive resources [Neuropsychology 12 (1998) 13]. Action impairments were explored in 51 dementia participants with the short form of the multi-level action test (MLAT-S). A clinical neuropsychological test protocol was also administered. Regression analyses including measures of executive functioning, semantic knowledge, and global cognitive functioning showed that global cognitive functioning was the best predictor of MLA T-S errors. Furthermore, task demands significantly influenced the type and frequency of errors, and dementia participants showed a pattern of errors similar to that reported in other clinical populations [Cognitive Neuropsychology 15 (1998) 617; Neuropsychologia 37 (1999) 51; Neuropsychology 12 (1998) 13]. Taken together, the present findings are inconsistent with semantic and executive accounts, but support the limited-capacity resource theory of naturalistic action impairment. © 2002 Elsevier Science Ltd. All rights reserved.

Keywords: Activities of daily living; Errors of action; Semantic knowledge; Executive functions; Dementia severity

1. Introduction

Naturalistic action is behavior in the service of everyday, simple tasks (e.g. tooth brushing) and extended activities (e.g. grooming), which require one to use objects and sequence multiple steps to achieve nested goals. The topic of this paper is the breakdown of naturalistic action in persons with dementia. Historically, naturalistic action has been studied from two different perspectives. On the one hand, it has provided a sensitive methodology for studying the left-hemisphere apraxia syndromes, specifically, the nature and functional consequences of the gesture- and object-use deficits central to these syndromes [13,18]. On the other hand, naturalistic action has proven to be interesting in its own right, posing as yet unresolved questions about the organization of complex, routine action, including the conceptual knowledge base and role of executive functions and general resources [5,6,26,53,55–57]. These are the questions that motivated the present study, and while others have used the label “apraxia” when describing breakdown of complex, routine action in persons with dementia [43,46], we prefer to adopt the more neutral, and less controversial descriptor, “naturalistic action deficit”. We will, however, refer to the apraxia literature whenever relevant throughout the paper.

Patients with dementia frequently experience naturalistic action impairment [1], and this impairment has been associated with institutionalization [27] and increased caregiver burden [12]. Neuropsychologists who evaluate patients for the presence of dementia are often expected to make recommendations regarding their patients’ ability to independently execute naturalistic tasks. However, the relationship between neuropsychological test scores and naturalistic action is poorly understood, making prediction difficult.

Investigators have consistently shown a relationship between dementia severity and informant-report questionnaires of everyday task performance [7,24,31,44,63]. When studies have included extensive neuropsychological protocols, scores on executive function [7] or visuoperceptual tests [24,31] have been associated with questionnaire data. Significant correlations have also been reported between performance-based measures of naturalistic action and neuropsychological tests (e.g. dementia severity [2,4,15,59,62]; dementia severity & executive functions [40]; dementia...
severity & visuo perceptual skills [24]; dementia severity, executive functions, & visuo perceptual skills [3,37]).

With the exception of Hill et al. [24], investigators have used constructional tasks (e.g. figure copy) to test visuo perceptual skills and have interpreted the correlation between these tasks and naturalistic action as evidence for the importance of visuomotor coordination in the latter. Freeman et al. [19], however, have demonstrated a strong relationship between constructional tasks and tests of executive function in dementia. Thus, an alternate explanation is that constructional tasks and naturalistic action impose similar demands on executive functions.

In fact, there is evidence for a link between executive deficits and naturalistic action impairment in dementia [3,7,37,40]. In a detailed investigation of the naturalistic action performance of two mild to moderate dementia patients (DM and HB), Buxbaum et al. [5] showed that patient HB, who exhibited poor executive functioning and preserved semantic knowledge on background tests, showed significant impairment on a test of naturalistic action. By contrast, patient DM, who demonstrated impaired semantic knowledge but good executive functions, performed well on naturalistic tasks.

Correlations between tests of executive function and naturalistic action impairment have also been reported for participants with neurological impairments other than degenerative dementia (e.g. closed head injury [56]). Reports of disorganized and ineffectual everyday functioning are common after focal frontal lobe injury, and find ready explanations in theories of executive functioning that postulate deficient sequencing [36,52], impaired working memory, perseveration, disinhibition of automatic behaviors (e.g. utilization behavior) [20], poor planning [42], or the inability to select the necessary goals to complete a task [14]. However, when Schwartz et al. [56] evaluated specific predictions of these theories against performance data from patients with closed head injury, none of these executive function accounts were supported. Thus, the validity of the executive function account remains unclear.

Originating from the apraxia literature, a second viable account is that naturalistic action deficits in dementia reflect impaired semantic knowledge. Semantic knowledge deficits are widely reported in dementia, particularly in Alzheimer’s disease [50]. Ochida et al. [43] have shown that multiple aspects of tool and object knowledge are impaired in patients with Alzheimer’s disease.

Investigators have linked deficient semantic knowledge to action deficits. For example, DeRenzi and Lucchelli [13] reported a significant correlation between single tool use and action performance in participants with focal left-hemisphere damage. They also observed a high rate of omission and object misuse errors (i.e. instances when objects were used in a conceptually inappropriate manner) relative to sequence and other errors in the naturalistic action performance of these participants. With this evidence DeRenzi and Lucchelli argued that knowledge of single tool/object-use is important for successful naturalistic action performance. Further support was reported by Schwartz et al. [55] who showed that non-frontal tests of object recognition and object knowledge were most strongly correlated with naturalistic task performance in stroke and closed head injury patients.

By contrast, Buxbaum et al. [5] reported preserved action performance in semantic dementia patient DM and concluded that semantic knowledge deficits are neither necessary nor sufficient for naturalistic action impairment. Furthermore, Hodges et al. [25] propose that naturalistic action performance in patients with semantic deficits may be mediated by non-semantic object affordances, mechanical problem solving skills, task context familiarity, and/or residual category knowledge. Thus, the semantic knowledge account, like the executive function account, clearly warrants further exploration in dementia.

1.1. Resource theory of naturalistic action impairment

Investigators have consistently reported a relationship between dementia severity and naturalistic action. In fact, Feyereisen et al. [15] showed that correlations between action performance and tests of specific neuropsychological functions were no longer significant when dementia severity was statistically controlled. This raises the possibility that the relationship between tests of specific cognitive deficits and action may be more parsimoniously explained by a link between action impairment and general level of dementia.

A recent account of naturalistic action suggests that action deficits are best understood as an exaggeration of the normal tendency to commit errors of action in states of fatigue or distraction [41,48]. Schwartz et al. [56] demonstrated a continuum between normal and patient behavior in their laboratory tests of naturalistic action. Mild closed head injury patients performed similarly to controls on tasks like making toast and wrapping a present unless the tasks were complicated by additional planning and prospective memory requirements. In that context, their rate of errors on toast making, etc. diverged from controls, and omission errors appeared for the first time. Close head injury patients who were clinically more severe made omission errors even without the complex task requirements; they also made more overt (“commission”) errors, including some that were more egregious than any produced by controls or mildly impaired patients.2

---

1 Hill et al. [24] showed a relationship between informant ADL questionnaire responses and performance on Poppelreuter’s figures and interpreted this finding as evidence for the importance of peripheral sensory processes for successful performance of everyday tasks in dementia.

2 For example, one severely impaired patient put raw hotdogs (instead of luncheon) between two slices of bread when making a sandwich and another placed postage stamps on the wrong side of an envelope when preparing a letter for mailing; controls never made such errors [56].
Other studies have confirmed that omission rates are particularly sensitive to patients’ clinical severity and task complexity [6,53]. Furthermore, when task demands are kept constant, error patterns have proven to be quite similar across right stroke, left stroke, and closed head injury patient groups [6,53,56]. On the other hand, task demands and situational context may influence the quality and quantity of errors, even within a single patient [55,57].

These various lines of evidence have led to the hypothesis that naturalistic action is resource demanding and highly sensitive to fluctuation—transient or chronic—in general cognitive or attentional resources [17,53,56]. On this account, brain damage affects the capacity of cognitive resources, in proportion to the overall severity of the damage (as indexed by clinical measures). The more severe the resource limitation, the higher the rate of error and the higher the ratio of omissions to commission errors. Moreover, whereas the other accounts we have discussed consider that error rates for object substitutions, perseveration, and others are reflective of the specific nature of the deficit [17], the resource account views them as subject to fluctuation in response to the demands of the task and context.

To summarize, the semantic knowledge and executive function accounts predict that deficits in these specific neuropsychological functions will predict the precise pattern of naturalistic action impairment. Specifically, greater executive deficits should be associated with sequence and perseverative errors, a tendency toward utilization behavior, and more difficulty performing two tasks concurrently. Semantic knowledge deficits, on the other hand, should be associated with substitution errors and difficulty when distractor objects are present. By contrast, the resource theory predicts that action performance should be associated strongly with measures of global cognitive functioning and weakly with measures of specific cognitive processes. In addition, an increase in errors, particularly omissions, should be observed as a function of task complexity, while the pattern of commission errors should vary across tasks. Finally, the pattern of action errors observed in dementia should be similar to those seen in other clinical populations [6,53,56].

2. Methods

2.1. Participants

Fifty-four participants were recruited from an outpatient memory clinic (Crozer Chester Medical Center, Chester, PA) where they were assessed by a social worker, psychiatrist, geriatrician, neurologist and neuropsychologist. Participants were included if they exhibited mild to moderate dementia (mini-mental state examination (MMSE) scores between 12 and 26 [16]), regardless of etiology. Patients were excluded if they demonstrated insufficient arousal/attention to tolerate testing, motor/sensory deficits precluding object grasping and use, and/or history of head injury, epilepsy, premorbid neurological illness, or long-standing psychiatric illness. Three patients of the 54 tested did not complete the semantic and executive measures used to test the study hypotheses and were not included in the final analyses (N = 51).

2.2. Procedures

Participants were tested over two sessions scheduled within a 2-week period. A 2–3 h neuropsychological assessment was administered during the first session. Naturalistic action was assessed during a second, 1–2 h testing session. In addition, during the 2-week assessment period, each participant’s caregiver was asked to complete a questionnaire regarding the patient’s home performance of everyday activities.

2.3. Informant ADL questionnaire

Everyday task performance in the home was assessed with a modified version of the instrumental activities of daily living and physical self-maintenance scales developed by Lawton and Brody [30], which requires caregivers to rate participants’ performance of 15 tasks (e.g. grooming, transportation, etc.) Each task is assigned one score from the following three-point scale: 1: participant performs task independently; 2: participant requires assistance to perform task; or 3: participant is entirely dependent on others to perform task. Behavioral descriptions were anchored to each number on the scale to make the rating less subjective. The total score could range from 15 (participant performs all tasks independently) to 45 (participant is completely dependent on others to perform all tasks). Only caregivers who had at least weekly contact with the participant completed this measure.

2.4. Neuropsychological assessment

2.4.1. Dementia severity

All participants were administered the MMSE [16] as a measure of dementia severity or global cognitive functioning.

2.4.2. Executive functioning

Participants completed the Boston Revision of the Wechsler Memory Scale—mental control subtest (MC) [10,29], which includes three non-automatized tasks (reciting the months of the year backward, alphabet rhyming & alphabet visualization). Accuracy indices were calculated for each task using the following algorithm: \[1 - ((\text{false positive} - \text{false negative}) / N)\].

Lawton and Brody [30] required a social worker to complete the questionnaire based on information gathered from multiple sources, including the family, the participant, clinicians, and so on. The original questionnaire rated participants’ behavior on a scale of 1–5 for some items. In this version, a family member or caregiver with close contact with the participant completed the questionnaire, and we adopted the 1–3 scale for all items.
Tiger

Does it have hooves or claws

Rhinoceros

Does it have antlers or a horn

Semantic probe test sample items

<table>
<thead>
<tr>
<th>Item</th>
<th>Sample animal probes</th>
<th>Type</th>
<th>Item</th>
<th>Sample tool probes</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rhino</td>
<td>Does it have antlers or a horn</td>
<td>F</td>
<td>Scissors</td>
<td>Does it have 2 moving parts or is it made of 1 piece</td>
<td>P</td>
</tr>
<tr>
<td></td>
<td>Does its hair have spots or is it 1 solid color</td>
<td>P</td>
<td>Is the edge sharp or blunt</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Does it eat grain &amp; leaves or animals</td>
<td>F</td>
<td>Is this for measuring or cutting</td>
<td>F</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Does it live in open fields or the desert</td>
<td>F</td>
<td>Is it used on cloth or stone</td>
<td>F</td>
<td></td>
</tr>
<tr>
<td>Tiger</td>
<td>Does it have hooves or claws</td>
<td>P</td>
<td>Wrench</td>
<td>Does it have an opening at the end or in the middle</td>
<td>P</td>
</tr>
<tr>
<td></td>
<td>Does its fur have spots or stripes</td>
<td>P</td>
<td>Is it blunt or sharp</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Does it eat animals or grain &amp; leaves</td>
<td>F</td>
<td>Is it used for tightening or cutting</td>
<td>F</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Does it live in the desert or jungle</td>
<td>F</td>
<td>Is it used on nails or bolts</td>
<td>F</td>
<td></td>
</tr>
</tbody>
</table>

Note: The correct response is italicised. Key for probe type: F: functional; P: perceptual. See [9] for more details regarding test construction and administration.

Semantic knowledge is often assessed with probe tests. The names of these items were printed on a large card and shown to the participant. (The examiner also read the word aloud several times.) Two functional and two perceptual questions were asked about each item. Questions were posed in a forced-choice format with two alternatives (Table 1). Behavioral studies of neurologic patients (see [50] for a review) and functional neuroimaging studies of healthy controls [21,38,45] have shown that semantic knowledge is organized categorically, such that animals are represented distinctly from inanimate objects, or tools. Furthermore, tool knowledge may be more relevant than animal knowledge for naturalistic action [13]. Therefore, errors were summed separately for the tool (probe tools) and animal items (probe animals).

Our selection of semantic measures is based, in part, on studies that have demonstrated that participants diagnosed with Alzheimer’s disease, who are known to have semantic knowledge deficits, perform significantly worse on these tests relative to participants of equal dementia severity diagnosed with Parkinson’s disease or ischemic vascular dementia [9,21,33]. Thus, these studies provide external validity for MC and GST as measures of executive function.

2.4.3. Semantic knowledge

Participants were asked to report as many different animal names as they could think of in 1 min (animal naming). All responses were coded on each of the following attribute categories: size (big, small); geographic location (foreign, North America); diet (herbivore, carnivore, omnivore); zoological class (insect, mammal, bird, etc.); habitat (farm, Africa/jungle, widespread, etc.); and biological order/related groupings (feline, canine, bovine, etc.). The association index (AI) was calculated by averaging the number of shared attributes for each participant’s consecutive animal naming responses. The AI provides a measure of the semantic relatedness between animal naming responses independent of the number of words produced, which can be low due to retrieval deficits [21]. An example of how this measure is calculated, is provided in Appendix A.

Semantic knowledge is often assessed with probe tests [50]. We used a new probe test for dementia participants that assesses semantic knowledge with 64 questions [9]. The items probed include eight animals and eight tools. The names of these items were printed on a large card and shown to the participant. (The examiner also read the word aloud several times.) Two functional and two perceptual questions were asked about each item. Questions were posed in a forced-choice format with two alternatives (Table 1). Behavioral studies of neurologic patients (see [50] for a review) and functional neuroimaging studies of healthy controls [21,38,45] have shown that semantic knowledge is organized categorically, such that animals are represented distinctly from inanimate objects, or tools. Furthermore, tool knowledge may be more relevant than animal knowledge for naturalistic action [13]. Therefore, errors were summed separately for the tool (probe tools) and animal items (probe animals).

Our selection of semantic measures is based, in part, on studies that have demonstrated that participants diagnosed with Alzheimer’s disease, who are known to have semantic knowledge deficits, perform significantly worse on these tests relative to participants of equal dementia severity diagnosed with Parkinson’s disease or ischemic vascular dementia [9,21,33].

2.4.4. Additional neuropsychological tests

Participants were also administered measures of episodic memory, motor functioning, visuoconstructual skills, and verbal fluency. These additional neuropsychological tests are described in Table 2.

2.5. Naturalistic action assessment

Naturalistic action was assessed using a shortened form of Schwartz and co-workers’ [6,53,56] multi-level action test (MLAT). The MLAT and the shortened form (MLAT-S) require participants to complete a series of everyday tasks independently, with little guidance from the examiner. Instructions, cues, termination procedures, and placement of objects is standardized [6,53,56]. Three of the four MLAT items are administered in the MLAT-S: Item 1—prepare toast with butter and jelly and prepare coffee with cream and sugar; Item 2—wrap a gift with related distractor objects.
Table 2
List of additional neuropsychological tests

<table>
<thead>
<tr>
<th>Test</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Episodic memory</td>
<td></td>
</tr>
<tr>
<td>California verbal learning test-dementia version (CVLT)</td>
<td>The dependent variable was the accuracy on the delayed recognition memory task (discriminability index)</td>
</tr>
<tr>
<td>Motor functioning</td>
<td></td>
</tr>
<tr>
<td>Finger tapping test</td>
<td>The number of taps in 10 s was averaged over five trials per hand</td>
</tr>
<tr>
<td>Visuospatial skills</td>
<td></td>
</tr>
<tr>
<td>The clock drawing test-copy condition</td>
<td>Participants copied a clock with hands set to 10 after 11. Ten possible errors were scored</td>
</tr>
<tr>
<td>Verbal fluency</td>
<td></td>
</tr>
<tr>
<td>Phonemic word list generation (FAS)</td>
<td>The dependent variable is the number of words produced in 60 s beginning with F, A, or S, excluding proper nouns</td>
</tr>
</tbody>
</table>

CVLT [35]; finger tapping test [49,61]; clock drawing test [34], FAS [47,60].

(gardening clippers, stapler, etc.) in the array; and Item 3—pack a lunch box with a sandwich, snack, and a drink and pack a school bag with supplies for school, while several of the necessary objects (knife, thermos lids) are stored out of view in a drawer with additional, and potentially distracting objects (ice tongs, measuring tape, etc). These specific items were selected and ordered in this sequence because previous studies of healthy controls and neurologically-impaired participants have shown that each item elicits more errors than the preceding item [6,53,56].

MLA T-S performance was videotaped, and an accomplishment score, reflecting the percentage of steps completed without error, was calculated. A detailed error analysis, identical to that used with the original MLA T, was performed to examine the number and types of errors made. An inter-rater agreement rate of 78% has been reported for this scoring procedure [56]. The error taxonomy is briefly described in Table 3. Three “global error scores” are derived from this coding system: (1) total errors (sum of all MLA T-S errors); (2) commissions (sum of all errors excluding omissions); and (3) omissions (total omission errors). Individual error types (i.e. “specific error scores”) are also analyzed separately. Errors were also summed for each item (“Item 1 errors,” etc.).

MLA T-S data are also reported as standardized error rates. Standardized rates for omission, sequence (reversal and anticipation-omission) and substitution errors were calculated by dividing the number of errors by the maximum number that could occur for that item and multiplying by 100. The standardized error rate for action addition errors was calculated by dividing these errors by the number of objects available for that item and multiplying by 100. Past studies show MLA T performance (i.e. accomplishment score and total errors) does not correlate with education, nor do men and women differ in performance [6,53,56].

A significant correlation between the MLA T and age was noted for R CVA participants [53], but was not observed for healthy controls or other clinical populations [6,56]. The MLA T-S (with a simplified scoring system), has been

Table 3
Multi-level action test error taxonomy

<table>
<thead>
<tr>
<th>Error category</th>
<th>Definition</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Substitution</td>
<td>Semantically related or perceptually similar alternate object selected in place of target object</td>
<td>Spread butter with spoon (instead of knife)</td>
</tr>
<tr>
<td>Gesture substitution</td>
<td>Correct object is used, but with incorrect gesture</td>
<td>Spoon (rather than pour) cream into coffee</td>
</tr>
<tr>
<td>Spatial misorientation</td>
<td>Misorientation of the object relative to the hand or another object</td>
<td>Grasp wrong end of scissors</td>
</tr>
<tr>
<td>Spatial misestimation</td>
<td>Spatial relationship between two or more objects is incorrect</td>
<td>Wrapping paper cut too small for gift</td>
</tr>
<tr>
<td>Tool omission</td>
<td>Action is performed without appropriate implement</td>
<td>Spread mustard with finger</td>
</tr>
<tr>
<td>Omission</td>
<td>Coded only when a step or subtask is never attempted</td>
<td>Failure to add coffee to mug</td>
</tr>
<tr>
<td>Perservation</td>
<td>Repetitive movements or duplications of steps or subtasks</td>
<td>Stir coffee for extended period of time; make two sandwiches</td>
</tr>
<tr>
<td>Sequence/anticipation omission</td>
<td>Anticipation of step which entails a subsequent omission</td>
<td>Seal thermos before filling</td>
</tr>
<tr>
<td>Sequence/reversal</td>
<td>Steps or subtasks performed in reverse order</td>
<td>Stir mug of water, then add instant coffee grinds</td>
</tr>
<tr>
<td>Action addition</td>
<td>Action not readily interpretable as a step in the task at hand, includes “utilization behavior” and anomalous actions</td>
<td>Pack non-school items, such as mustard or aluminum foil in school bag</td>
</tr>
<tr>
<td>Quality</td>
<td>Inappropriate or inexact quantity</td>
<td>Fills thermos with juice to the point of overflow</td>
</tr>
</tbody>
</table>
psychometrically validated on stroke and closed head injury patients [54], and has been shown to contribute to the prediction of instrumental activities of daily living (IADL) outcomes [51].

2.6. Semantic knowledge and executive function composite scores

A principal-component analysis (PCA) that included each of the putative measures of executive functioning (GST-D & MC) and semantic knowledge (AI, probe animals & probe tools) was performed to determine whether these measures aligned as expected. The eigenvalue criterion (eigenvalue >1) was used to determine the number of factors extracted, and varimax rotation was performed to simplify the interpretation of the factors. A two-factor solution accounting for 66.5% of the variance resulted, with GST-D and MC as one factor (executive function) and probe animals and AI as a separate factor (semantic knowledge). Probe tools loaded with tests of executive function, which is consistent with the fact that knowledge of tool attributes has been associated with frontal lobe functioning in neuroimaging studies of healthy controls (e.g. premotor cortex [23,38]; left dorso-lateral frontal cortex [45]) and case reports of patients with tool-specific knowledge deficits (see [50] for a review).

Standardized scores (Z scores) were derived for all semantic and executive measures and then averaged to produce a semantic knowledge (S) and executive function (E) composite score. Because probe tools—a putative semantic measure—loaded with executive measures, it was not included in either composite score; rather, the relationship between probe tools and naturalistic action was assessed with separate analyses.

2.7. Data analysis

Several variables required transformation to meet assumptions of normality for bivariate (Pearson) and multivariate correlational analyses: MMSE (squared), commissions (square root), and omissions (square root). A subset of variables were not normally distributed and could not be transformed because of a high number of 0 values (GDS, clock drawing test; probe tools; MLAT-S substitutions, sequence, perseverations, action additions, Item 1 errors & Item 2 errors). For these variables, we used Spearman rank-order correlation coefficients to measure strength of associations and Wilcoxon tests to measure differences between conditions. Partial correlations were performed to control for the effects of general severity, as indexed by MMSE. Bonferroni correction was used to interpret significance in cases where a series of correlations or between-group comparisons were tested.

3. Results

3.1. Demographic and neuropsychological characterization of the participants

On average, participants were 77.5-year-old (S.D. = 6.1, range = 60–88) and had completed 11.6 (S.D. = 2.4, range = 6–19) years of education. Fifteen were men and 36 were women. All participants were diagnosed with a mild to moderate level of dementia (MMSE = 22.2, S.D. = 3.4, range = 14–26) due to a range of etiologies, including Alzheimer’s disease (n = 9), vascular dementia (n = 18), Parkinson’s disease (n = 8), dementia due to multiple etiologies (n = 13), and substance-induced persisting dementia (n = 3).

Table 4 shows the group’s average score on each neuropsychological test. Several participants did not complete the entire neuropsychological evaluation (n = 44 for clock drawing test; n = 49 for geriatric depression scale; n = 46 for finger tapping test), and only 40 caregivers were available to complete the informant ADL questionnaire. Limited time and scheduling conflicts were the most common causes of missing data. There was no significant difference

<table>
<thead>
<tr>
<th>Dementia participants (mean (S.D.))</th>
<th>Normative data (mean (S.D.))</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Executive functions</strong></td>
<td></td>
</tr>
<tr>
<td>Mental control (MC)</td>
<td>62.8 (25.2)</td>
</tr>
<tr>
<td>Graphical sequences test (GST)</td>
<td>15.6 (16.0)</td>
</tr>
<tr>
<td><strong>Semantic knowledge</strong></td>
<td></td>
</tr>
<tr>
<td>Association index (AI)</td>
<td>3.18 (.70)</td>
</tr>
<tr>
<td>Semantic probe test</td>
<td>8.5 (6.1)</td>
</tr>
<tr>
<td><strong>Additional neuropsychological tests</strong></td>
<td></td>
</tr>
<tr>
<td>CVLT discriminability index (CVLT)</td>
<td>80 (10)</td>
</tr>
<tr>
<td>Finger tapping test</td>
<td>35.8 (9.3)</td>
</tr>
<tr>
<td>Clock drawing-copy</td>
<td>2.1 (2.0)</td>
</tr>
<tr>
<td>FAS</td>
<td>17.4 (9.9)</td>
</tr>
</tbody>
</table>

* n = 51.

MC [10,29]; GST [28]; AI [21]; semantic probe test [9]; CVLT [35]; finger tapping test [49,61]; clock drawing [34]; FAS [47,60].
in MMSE score between participants who had missing test scores (n = 19; M(MMSE) = 22.6) and those who had no missing data (n = 32; M(MMSE) = 22.3), indicating that missing data were not disproportionately from more impaired participants.

3.2. MLAT-S relationship to clinical and demographic variables

On the MLAT-S, participants made a mean of 16.4 (S.D. = 9.3; range = 4–60) errors and attained a mean accomplishment score of 67.1% (S.D. = 25.7; range = 6–100). MMSE correlated with both accomplishment score (r = 0.31, P = 0.022) and total errors (r = −0.35, P = 0.010). Even for participants with the mildest level of dementia (MMSE > 23; n = 24), the error rate was high (M of total errors = 13.3; S.D. = 9.2; range = 1–32) and the accomplishment score low (M = 75.2; S.D. = 23.5; range = 11–100). Accomplishment score and total errors did not correlate with age or education and there was no difference between men and women (t-test) or between participants with different dementia diagnoses (Kruskal–Wallis tests).

3.3. MLAT-S and caregiver ADL questionnaire

On average, caregivers reported that the dementia participants required assistance and supervision to perform many everyday tasks (M ADL questionnaire score = 24.15; S.D. = 6.31, range = 15–38). The relationship between the caregiver ADL questionnaire and the MLAT-S was assessed to determine whether MLAT-S performance generalized to performance of everyday activities in the home. There was a significant positive correlation between the ADL questionnaire and accomplishment score (r = 0.34, P = 0.032) and total errors (r = 0.45, P = 0.004), and partial correlations controlling for MMSE remained statistically significant (accomplishment score r = −0.38, P = 0.013; total errors r = 0.41, P = 0.009).

3.4. General resources, semantic knowledge, and executive functioning

3.4.1. MLAT-S global scores

Regression analyses were performed to evaluate the competing theories of action impairment and also to explore the possibility that a multifactorial model may explain MLAT-S errors. Three stepwise multiple regression analyses were performed with total, commission, and omission errors as the dependent variables and MMSE, E composite score, and S composite score as the independent variables. The best model for total errors accounted for 28% of the variance (F(1, 49) = 4.1, P = 0.05) and had MMSE as the only significant predictor (R = −0.28, P = 0.05). The best model for omissions accounted for 22% of the variance (F(1, 49) = 5.4, P = 0.02) and also had only MMSE as a significant predictor (R = −0.32, P = 0.02). The regression analysis for commission errors was not significant.

When probe tools was added to the models as a predictor, the results did not change. However, because probe tools was not normally distributed and was recoded as a dichotomous variable for the regression (1: 0–1 errors; 2: >1 errors), we also explored the relationship between tool knowledge and naturalistic action with correlations. Significant positive correlations were observed between probe tools and both total errors (r = 0.35, P = 0.019) and omissions (r = 0.37, P = 0.013). However, partial correlations, controlling for MMSE, were no longer significant (total errors r = 0.18, P = 0.23; omissions r = 0.12, P = 0.42).7

3.4.2. Specific error scores

Spearman correlations between specific error scores (action addition, sequence, perseveration, & substitutions) and MMSE, S composite, and E composite were not significant (Bonferroni corrected P > 0.014; r < 0.25, P > 0.07 for all). There was a positive association between probe tools and perseverations, which was nearly significant with and without controlling for MMSE (r = 0.32, partial r = 0.31, P < 0.05 for each).

3.4.3. Item error scores

Recall that one of the predictions of the executive account is that the planning and prospective memory requirements of Item 3 should make it particularly vulnerable to executive dysfunction. A stepwise multiple regression analysis was performed with Item 3 errors as the dependent variable and MMSE, S composite and E composite as the predictor variables. The best model accounted for 43% of the variance (F(1, 49) = 11.1, P = 0.002) and had MMSE as the only significant predictor (R = −0.43, P = 0.002). Item 3 errors were significantly positively correlated with probe tools (r = 0.36, P = 0.014), but the partial correlation, controlling for MMSE, was not significant (r = 0.16, P = 0.29).

Spearman rank order correlations were performed for Items 1 and 2 errors, as these variables were not normally distributed. There was no significant correlation between these error measures and MMSE, S and E composite, or probe tools (r < 0.30, P > 0.05 for all; Bonferroni corrected P = 0.013).

---

5 Hereafter, analyses examined only MLAT-S error variables, as the accomplishment score highly correlated with error measures (i.e. r = −0.80, P < 0.001 for correlation with total errors).

6 MMSE was highly correlated with probe tools (r = −0.45, P = 0.002).

7 As a further check, we recomputed the composite analyses with probe tools added to the S composite score, with no change in the results.
3.5. Additional neuropsychological tests

3.5.1. MLAT-S global scores

Partial correlation analyses, controlling for MMSE, were performed to explore the relationship between total errors, omissions and commissions and neuropsychological tests of memory (CVLT), visuoconstructional skills (clock drawing), motor speed (finger tapping), and verbal fluency (FAS). Significant positive correlations were observed between the clock drawing test and both total errors \( r = 0.38, P = 0.009 \) and commissions \( r = 0.37, P = 0.013 \). No other partial correlation reached statistical significance \( r < 0.29, P > 0.04 \) for all; Bonferroni corrected \( P = 0.013 \).

3.6. Specific error scores

Partial correlations showed a significant positive correlation between clock drawing and substitutions \( r = 0.41, P = 0.006 \). There was also a nearly significant correlation between FAS and perseverations \( r = 0.34, P = 0.018 \). None of the other analyses reached statistical significance \( r < 0.23, P > 0.10 \) for all; Bonferroni corrected \( P = 0.013 \).

3.6.1. Item errors

Partial correlations between Items 1 and 3 errors and the additional neuropsychological tests did not reach statistical significance \( r < 0.28, P > 0.06 \) for all; Bonferroni corrected \( P = 0.013 \).

3.7. Standardized error rates across MLAT-S items

Analyses of errors across the MLAT-S items were performed using standardized error rates (for sequence, substitutions, omissions, and action additions) as dependent variables (Bonferroni-corrected \( P = 0.017 \)). The standardized rate for total errors was lower on Item 1 \( (M = 4.4, \text{ S.D.} = 5.7) \) as compared to Item 2 \( (M = 9.3, \text{ S.D.} = 5.7) \) and Item 3 \( (M = 10.3, \text{ S.D.} = 6.2) \). Item 1 versus Item 2: \( Z = -3.04, P = 0.002 \); Item 1 versus Item 3: \( Z = -5.34, P < 0.001 \). There was no difference between Items 2 and 3. Standardized error rates for just omissions and just sequence errors showed the predicted pattern, Item 1 < Item 2 < Item 3 (see Fig. 1). For omissions, the difference between Item 3 and each of the other two items was significant, but the difference between Item 1 and Item 2 was not (Item 1 versus Item 3: \( Z = -4.53, P < 0.001 \); Item 2 versus Item 3: \( Z = -3.43, P < 0.001 \)). For sequence errors the only significant difference was between Items 3 and 1 (\( Z = -4.11, P < 0.001 \)). Substitutions patterned differently from the others (Fig. 1), yielding the highest error rate at Item 2 (gift-wrapping with distractors). Item 2 was significantly different from Items 3 (\( Z = -2.73, P = 0.006 \)) and the difference between Items 2 and 3 just missed significance after Bonferroni correction (\( Z = -2.18, P = 0.03 \)). There was no difference in the standardized error rate for action additions across the items.

3.8. Dementia participants and other clinical populations

Although the methodology in this study is not identical to that of past MLAT studies, similar items were used and the same scoring procedure was applied. It was therefore possible to compare the dementia group to other groups tested with the MLAT with respect to the pattern of specific error proportions of each error type (e.g. omission proportion = # omissions/# total errors). Table 5 shows that the pattern of errors produced by dementia participants was strikingly similar to other clinical groups (30 CHI [56]; 30 R CVA [53]; 16 L CVA [6]). There was a significant correlation...
between the dementia group’s error pattern and the pattern for the CHI and RCV A groups (Spearman r = 0.80 and 0.82, respectively; P < 0.01 for both). The correlation between the dementia and L CVA groups (r = 0.70, P = 0.036) just missed significance after Bonferroni correction (P = 0.017).

4. Discussion

Competing accounts of the source of naturalistic action impairments in dementia were tested. Regression analyses were performed with MLAT-S error scores as the dependent variables and measures of executive functioning, semantic knowledge, and general dementia severity as predictor variables. Dementia severity was the only significant predictor for naturalistic action performance, except for the analysis of commission errors. The variance in commissions was not explained by any of the accounts tested; however, partial correlation analyses controlling for MMSE showed a significant relationship between commissions and visuoconstructional skills, indexed by clock drawing. Comparisons of standardized error rates across MLAT-S items showed that participants made more sequence and omission errors on the most complex item (Item 3). Substitution errors, by contrast, occurred more frequently when related distractors were included along with target objects (Item 2). Finally, when dementia participants were compared to other clinical populations tested with the MLA T, the distribution of error types was found to be similar.

Before further discussion of the results, it is important to highlight two points. First, significant naturalistic action difficulties were observed even in patients with mild dementia (i.e. MMSE > 23), confirming that action deficits are, indeed, a serious problem for this population. Second, the informant ADL questionnaire significantly correlated with MLAT-S scores (accomplishment score & total errors). This provides support for the validity of the MLAT-S as a measure of everyday action performance. This result is consistent with a recent study showing that MLAT-S performance (with a simplified scoring system) was a significant predictor of future IADL functioning in individuals undergoing inpatient rehabilitation for CHI or stroke [51].

The executive and semantic accounts explain action impairments in terms of specific cognitive deficits. According to executive theories, the causal contribution of executive dysfunction is manifested in sequence and perseverative errors, utilization behavior, and difficulty performing two tasks concurrently [36,42,52]. On the semantic account, faulty semantic knowledge leads to object selection errors and a greater sensitivity to distractor objects in naturalistic action [13]. To the contrary, we found that scores on executive functioning and semantic knowledge did not significantly predict or correlate with MLAT-S global scores or specific error variables.

Several authors have suggested that tool knowledge is particularly relevant to naturalistic action performance [13]. For this reason, and because factor analysis showed that probe tools loaded with tests of executive function, we explored the relationship between naturalistic action performance and probe tools in separate analyses. Except for a nearly significant positive association with perseverations, probe tools did not correlate with the various action measures nor did it contribute to the predictive models. The suggestion of an association between tool knowledge and perseverative errors is potentially interesting, but not predicted by any of the accounts we have considered.

It is important to note that object substitutions, the error most closely associated with the semantic account, did occur with relatively high frequency in a few patients; and others demonstrated a bias toward other errors (e.g. utilization behavior). However, these error patterns did not appear to be systematically related to any aspect of neuropsychological test performance. These findings contradict the notion of a simple, transparent relationship between semantic impairments and naturalistic action errors in dementia. Detailed case studies are needed to determine whether these error patterns are stable over time and across tasks, and whether they are predicted by the cognitive neuropsychological status of the patient (for relevant evidence see [5,26]).}

The results of the regression and correlation analyses accord best with the resource theory [6,53,56], and this is also
true of the analyses of errors across the MLAT-S items. In past MLAT studies the finding that patients showed more omissions on complex compared to simple items was interpreted as the consequence of the severe capacity limitations imposed by both cognitive impairment and complex task conditions [6,53,56]. That is, multiple action schema are selected simultaneously in complex task conditions, and cognitively impaired patients lack the resources necessary to resolve competition between schema and achieve activation above threshold for execution. Consistent with this, we found that even when we controlled for differences in error opportunities across items, the rate of total errors was significantly higher for more complex tasks (Items 2 and 3) compared to simpler tasks (Item 1), and the increase in error frequency was most significant for omissions.

There was also evidence that error patterns varied significantly across items. For example, dementia participants demonstrated a higher rate of substitution errors when related distractor objects were present (Item 2). A trend for more substitution errors in the presence of distractor objects was also reported with CHI patients [56]. Task effects on error patterns have also been documented in single case studies [55,57]. For example, when a stroke patient with serious action difficulties (IH) was observed during tooth brushing and coffee preparation over several weeks, he consistently made more perseverative errors in tooth brushing compared to coffee making. Object substitution errors, by contrast, occurred more frequently in coffee making. This was explained by the fact that there were more non-exhaustible commodities (e.g. water, toothpaste) and fewer related objects available at the sink where the patient brushed his teeth, and more semantically related supplies at the breakfast tray (e.g. various small packets of condiments). Taken together, these results imply that task context, more than patient characteristics or specific cognitive impairment, influence naturalistic error patterns.

Finally, we found that the error proportions in the dementia group correlated with those obtained from previously studied RCV A [53] and CHI [56] groups. The correlation with a LCVA group [6] just missed significance; inspection of Table 5 shows that the largest discrepancy was in quality statistics or specific cognitive impairment, influence naturalistic error patterns.

When the relationship between naturalistic action and additional neuropsychological tests was explored, the correlation between verbal fluency and perseverations was significant at the trend level. Because this unexpected correlation was observed in only one analysis and did not reach statistical significance, we are hesitant to draw any conclusions from this finding, but suggest that it be further explored in future investigations.

The significant correlation between MLAT-S commission (and substitutions) and the clock drawing test was also unexpected based on all of the accounts tested; however, past studies have also reported a significant relationship between measures of visuospatial/constructional skills (e.g. drawing tests) and everyday tasks [3,37]. Schwartz et al. [53] demonstrated that patients with right hemiphere stroke, who are particularly vulnerable to deficits in arousal and allocation of attention, were more severely impaired in naturalistic action as compared to patients with left-hemisphere stroke [6] or CHI [56]. Thus, one possible explanation for the association between commission errors and visuconstructional measures is that the latter are sensitive to the right hemisphere attention/arousal processes that are important for action. It may also be relevant that the clock drawing test assesses multiple cognitive processes and brain regions [8]. Replication and further exploration of the relationship between visuconstructional skills and naturalistic action are warranted.

In summary, we have shown that individuals with diverse forms of dementia were defective in naturalistic action production, and that the features of the deficit (severity; error pattern) were not predicted by executive function or semantic memory scores, once dementia severity was taken into account. We chose the executive tests because they tap functions that bear an obvious relevance to naturalistic action production (e.g. sustained attention, working memory, behavioral monitoring). Other executive functions, such as complex planning and strategy application, are more relevant to non-routine than routine action production; nevertheless, we may have missed an association by not testing for these functions in our executive battery [12,58]. We also acknowledge that a more sensitive assessment of object knowledge, involving the very items of the MLAT-S, might have revealed an association between semantic memory and action production. Finally, the present results are based on a heterogeneous sample of dementia patients; it may turn out that the predictions of the executive-function and semantic-memory accounts receive more support when tested in groups with distinct and contrasting etiologies and syndromes (e.g. frontal versus semantic dementia). These caveats notwithstanding, the present study adds to the body of evidence showing that the link between neuropsychological deficit profile and naturalistic action performance is weak, at best. This suggests the need for caution in generating recommendations and predictions about naturalistic action from tests of specific cognitive functions. Rather, recommendations relevant to everyday functioning are most appropriately based on dementia patients’ general level of cognitive impairment and task context, including the complexity of activities and the presence of distractor objects.

Acknowledgements

This research was conducted in partial fulfilment of the requirements for Tania Giovannetti’s doctoral degree and was supported in part by NIH National Research Service Grant # 5132/R07425 and internal funds from Moss Rehabilitation Hospital. We would like to thank Marissa Carey, Rhonda Freeman, Melissa Lamar, and Monica Thompson for their help with collecting the neuropsychological test data.
Table A.1
Animal naming response coding for association index

<table>
<thead>
<tr>
<th>Response from participant</th>
<th>Size</th>
<th>Location</th>
<th>Diet</th>
<th>Zoological class</th>
<th>Habitat</th>
<th>Biological order</th>
<th>Site of capture</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response 1: bear</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Response 2: swan</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Response 3: lama</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Response 4: gnu</td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Response 5: wolf</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Response 6: jackal</td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Response 7: otter</td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Total rate of abilities = 22. Association index = 3.67. All possible attributes not listed within each category in this table, see [59] for a complete list.
References


Appendix A. Association index

Table A.1 shows the seven animal naming responses generated by participant DG. Starting with the second response and moving across the page, “mule” matches with the first response, “horse” on six attributes (big, local, farm, herbivore, mammal & equine). The third response is “lion,” which matches with “mule” on only two attributes (big & mammal). After the number of shared attributes for each successive response pair is tabulated, the total sum of shared attributes is obtained for all successive responses. (For example, in Table A.1 the total sum of shared attributes is 22.) The AI is calculated by dividing the sum of shared attributes by the number of responses minus one, because the attributes of the first response are never actually figured into the sum. The AI for participant DG is 3.67 or 22/(7 – 1). As can be seen from this example, the AI provides a measure of the semantic association between all consecutive animal naming responses that is independent of the number of responses generated. See [21] for more details.


