Journal of Clinical and Experimental Neuropsychology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/ncen20

To err is human, to monitor divine: Environmental adaptations reduce everyday errors but do not improve monitoring

Brianne Magouirk Bettcher [a] , Tania Giovannetti [b] , Elizabeth Klobusicky [b] , Denene Wambach [b c] , Joel Eppig [b c] & David J. Libon [c]

[a] Department of Neurology, Memory and Aging Center, University of California, San Francisco, CA, USA

[b] Department of Psychology, Temple University, Philadelphia, PA, USA

[c] Department of Neurology, Drexel University College of Medicine, Philadelphia, PA, USA

Available online: 01 Aug 2011

Psychology Press
Taylor & Francis Group

To err is human, to monitor divine: Environmental adaptations reduce everyday errors but do not improve monitoring

Brianne Magouirk Bettcher[1], Tania Giovannetti[2], Elizabeth Klobusicky[2], Denene Wambach[2,3], Joel Eppig[2,3], and David J. Libon[3]

[1]Department of Neurology, Memory and Aging Center, University of California, San Francisco, CA, USA
[2]Department of Psychology, Temple University, Philadelphia, PA, USA
[3]Department of Neurology, Drexel University College of Medicine, Philadelphia, PA, USA

The current study aimed to address error monitoring impairments in dementia using an intervention for execution deficits. Thirty-eight participants completed the Naturalistic Action Test (NAT) under two conditions: Standard and User-Centered. The Standard NAT followed the manual procedures; in the User-Centered NAT, objects were arranged sequentially, and distractor items were separated from target objects. While participants committed fewer errors in the User-Centered condition, there was no difference in the proportion of errors detected. However, the neuropsychological processes associated with monitoring differed across conditions. The results have implications for a neuropsychological model of error monitoring in dementia.

Keywords: Naturalistic action; Activities of daily living; Instrumental activities of daily living; Functional abilities; Errors of action; Error monitoring; Error detection.

The ability to accurately carry out everyday activities in a timely and efficient manner is an integral component of functional autonomy. Although everyday actions (e.g., making and eating lunch, grooming) are overlearned and often include perfunctory tasks, they necessitate a range of complex cognitive processes that are frequent targets of neurodegenerative disease. As such, everyday actions are particularly sensitive to changes in neurological functioning and represent a significant concern for individuals diagnosed with a dementia (American Psychiatric Association, 2000; Mioshi et al., 2007). The limited corpus of literature on this topic indicates that dementia patients exhibit distinct action deficits characterized by high error rates (Giovannetti, Libon, Buxbaum, & Schwartz, 2002; Ito & Kitagawa, 2006), reduced task accomplishment (Giovannetti, Libon, Buxbaum, et al., 2002), and a diminished capacity to handle errors once they have been committed (Bettcher, Giovannetti, MacMullen, & Libon, 2008; Mathalon, Whitfield, & Ford, 2003). Empirical research with older adults suggests that impaired action error detection and correction may signify a more immediate, pressing concern for independent living in this population (Bettcher et al., 2008; Falkenstein, Hoormann, & Hohnsbein, 2001; Giovannetti, Libon, & Hart, 2002; Nieuwenhuis et al., 2002), but attempts to devise a neuropsychological model and translate findings into feasible rehabilitation methods have been limited (Bettcher & Giovannetti, 2009). The primary aim of this study is to evaluate an environmental intervention designed to improve action performance and error monitoring in individuals diagnosed with a dementia. The action intervention techniques employed in this study were informed by current conceptualizations of the neuropsychological factors responsible for impoverished error monitoring.

A portion of this paper was presented at the 2009 International Neuropsychological Society conference. This research was supported in part by funds from an American Psychological Association, Division 40 Early Career Award, awarded to Brianne Bettcher.
Address correspondence to Brianne Bettcher, University of California, 350 Parnassus Ave, Suite 905, San Francisco, CA 94143–1207, USA (E-mail: bbettcher@memory.ucsf.edu).

http://www.psypress.com/jcen DOI: 10.1080/13803395.2011.595395

Error monitoring deficits in dementia

Before commencing a discussion on relevant empirical work related to error monitoring deficits in dementia, it is important to provide a working definition of the context for error detection and correction in this study—namely, everyday action tasks. Everyday action is defined as behavior in the service of everyday tasks that entails sequencing multiple steps and utilizing objects to accomplish nested goals (Giovannetti, Libon, Buxbaum, et al., 2002). Although everyday actions are related to the well-known concept of activities of daily living (ADLs), their scope is more generally limited to goal-directed behavior (i.e., as opposed to nondirective ADLs, such as ambulation).

Research with younger adults indicates that the impact of action errors is substantially less taxing if blunders are recognized and rectified in a timely manner (Reason, 1990; Sellen & Norman, 1992); given the concerning number of errors that dementia patients produce, how do they respond? Recent studies using highly controlled laboratory reaction time tasks and forced-choice response tasks suggest that dementia patients detect and correct significantly fewer errors than do age-matched controls (Ito & Kitagawa, 2006; Mathalon et al., 2003), who in turn monitor their errors less than do healthy younger adults (Band & Kok, 2000; Nieuwenhuis et al., 2002). Results from two naturalistic studies corroborate these findings and further indicate that dementia patients detect between 20% (Giovannetti, Libon, & Hart, 2002) and 34% (Bettcher et al., 2008) of their everyday action errors. Notably, patients diagnosed with a dementia correct the majority of errors they detect (76%), suggesting that the source of difficulty lies in recognizing and/or evaluating the action error (Bettcher et al., 2008). Taking into consideration the observation that dementia patients correct a high proportion of detected errors, facilitating immediate or delayed detections remains essential.

Cognitive and neuropsychological foundations of error monitoring

Although patients diagnosed with a dementia detect very few of their action errors, little is known about the derivation of their error monitoring difficulties. The literature on cognitive and human factors in error monitoring sheds some light on this issue in younger adults, as it offers a plethora of taxonomies that carefully pinpoint the individual's role in the *production* of errors as well as their likelihood for *detecting* these action errors (Blavier, Rouy, Nyssen, & De Keyser, 2005; Reason, 1990). Within the context of Reason's 1990 model, if an individual formulates an incorrect intention (i.e., planning error; he or she does not establish appropriate task parameters), then the subsequent error is referred to as a "mistake." Alternatively, if an individual incorrectly executes a correctly formulated plan, then the resulting error is referred to as a "slip" (i.e., execution error). In order to identify an action as incorrect (i.e., detection), external behavior is monitored and is compared to an internal representation of the desired state. As such, planning versus execution errors are inherently associated with different probabilities for detection. Specifically, mistakes represent the most difficult type of error to detect, as there is no dissonance between the outcome and the initial intention. In order to detect and correct a mistake, an individual must rely on third-party intervention or limiting external factors to facilitate monitoring (e.g., leaking thermos suggests improper application of lid). In contrast, "slips" are more readily detected due to the dissonance between the intention and executed action.

Although several studies have investigated the production of slips versus mistakes in younger adults (Norman, 1981; Reason, 1990), there is a paucity of literature on this issue in older adults diagnosed with a dementia. This is unfortunate, as the "slips versus mistakes" framework may help conceptualize everyday error monitoring deficits in this population. For instance, impaired error monitoring in dementia may be due to high rates of mistakes (i.e., errors due to deficient task knowledge). If so, then this might imply that error monitoring abilities are relatively preserved in the face of degraded everyday-task knowledge. On the other hand, impaired error monitoring may be explained by undetected slips (i.e., errors due to inaccurate execution). This would imply serious deficits in response monitoring, which is often subsumed under the superordinate construct of executive functioning (Miller & Cohen, 2001; Ullsperger & von Cramon, 2006), with relative preservation of task knowledge (Cosentino, Chute, Libon, Moore, & Grossman, 2006; Sirigu et al., 1995). A third possibility is that both knowledge and monitoring deficits are at play. If so, then the relative contribution of each deficit should be characterized.

Current study and hypotheses

The current study focused on reducing the number of slips (execution errors) and improving the monitoring of slips made by dementia patients in everyday tasks. Based on the younger adult literature, we postulated that everyday action interventions designed to minimize executive control demands might facilitate the detection and correction of everyday action slips but would do little to boost degraded knowledge or reduce mistakes. However, this hypothesis has never been evaluated.

The primary aim of the current study was to evaluate the impact of an environmental everyday action intervention on error monitoring processes in dementia. As an exploratory analysis, the current study also aimed to better characterize error monitoring deficits by exploring the relations between error monitoring and neuropsychological processes in participants and examining whether the intervention altered these relations. In order to accomplish these aims, participants were administered the Naturalistic Action Test (as described in the test manual; henceforth referred to as *Standard NAT*; Schwartz, Segal, Veramonti, Ferraro, & Buxbaum, 2002) as well as an environmental intervention designed to reduce the executive demands of the Standard NAT, the User-Centered Naturalistic Action Test (see Giovannetti, Bettcher, et al., 2007; referred to as *UC-NAT*). Action error monitoring rates were evaluated and compared across both conditions. A brief neuropsychological protocol that included an assessment of executive control (as well as additional measures of language, general cognitive functioning, and episodic memory to evaluate specificity) also was administered; correlations between neuropsychological test scores and error monitoring variables were analyzed.

In a previous study from our lab, participants diagnosed with a dementia demonstrated a reduction in total error rates on the UC-NAT relative to the Standard NAT (Giovannetti, Bettcher, et al., 2007); however, the impact of this intervention on error monitoring has yet to be evaluated. In contrast to the Standard NAT, the UC-NAT included the strategic placement of objects and an external cue to promote task monitoring. The UC-NAT environmental adaptations were designed to improve sequencing, reduce environmental distractibility, and increase monitoring, thereby addressing errors and monitoring failures due to impaired executive control. As such, our first hypothesis was that the UC-NAT intervention would be associated with higher error monitoring than the Standard NAT. Thus, we predicted that participants would detect and correct a higher proportion of their action errors in the UC-NAT condition than in the Standard NAT. We also considered, however, the possibility that error monitoring may become more burdensome due to the shift in error pattern towards mistakes. That is, although the overall number of errors is reduced in the UC-NAT condition, the few errors that remained might be due to deficient task knowledge, making them less likely to be detected/corrected. This possibility might work against our first prediction. Therefore, to evaluate this possibility, we conducted an exploratory correlation analysis. Our hypothesis was that the UC-NAT would reduce the likelihood of slips associated with executive control difficulties and free executive resources to facilitate behavioral monitoring. Consequently, we predicted that while error monitoring in the Standard NAT would be strongly associated with executive control abilities (Bettcher et al., 2008), this relation would be significantly weakened in the UC-NAT. That is, we anticipated a shift in neuropsychological processes associated with error monitoring as a result of the intervention.

METHOD

Participants

Retrospective data from 38 participants were collected from a larger sample of 46 participants from a previously published study (Giovannetti, Bettcher, et al., 2007). Participants in the original study were recruited from an outpatient program that included evaluations by a geriatrician and neuropsychologist, magnetic resonance imaging (MRI) of the brain, and laboratory studies. All participants met the following criteria: (a) probable Alzheimer's disease (AD; McKhann et al., 1984); (b) Mini Mental Status Exam (MMSE; Folstein, Folstein, & McHugh, 1975) ≤ 26 and ≥ 15; (c) native English speaker; (d) no evidence of a focal lesion(s) or cortical strokes on MRI; (e) no comorbid mood disorder or record of prior brain damage or disease, alcohol and/or drug abuse, or major psychiatric disorder. Participants were excluded from the present study if video recordings of test performance were lost ($n = 1$) or MMSE < 15 ($n = 8$). The more stringent MMSE cutoff was chosen to limit our evaluation to individuals with mild to moderate dementia. For a repeated measures design, the final sample ($n = 38$) provided sufficient power (.85) to detect a medium effect (.5) when alpha (p-value) is set at .05 (Erdfelder, Faul, & Buchner, 1996).

Neuropsychological testing

All participants completed neuropsychological testing as part of their clinical evaluation. From this evaluation, tests were chosen to assess a range of cognitive functions, including overall dementia severity, executive control, language/naming, and episodic memory. Neuropsychological tests are described in Table 1.

Everyday action measures

The everyday action measures employed in this study have been described in a previous publication (Giovannetti, Bettcher, et al., 2007). In brief, participants performed the NAT under two conditions: (a) Standard and (b) User-Centered (repeated measures design). In both conditions, the NAT required participants to perform three everyday tasks with minimal guidance from the examiner: Item 1—prepare toast with butter and jelly and prepare coffee with cream and sugar; Item 2—wrap a gift while distractor objects that are visually and/or semantically similar to target objects are available on the table; 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 are stored out of view in a drawer with potentially distracting objects. NAT instructions, cuing procedures, and scoring are standardized and are described in the test manual (Schwartz, Buxbaum, Ferraro, Veramonti, & Segal, 2003).

In the Standard condition, the NAT was administered exactly according to the NAT manual, which specifies the placement of each of the objects on the tabletop (Schwartz et al., 2003). Objects for each item were evenly distributed to the right and left of the participant on the U-shaped table, but the objects were not grouped/ordered in any meaningful way. In the User-Centered condition, participants performed the same tasks and were presented with the same objects on the U-shaped table as those in the Standard NAT. However, three environmental adaptations were implemented. First, objects were arranged on the tabletop from left to right in the order they should be used. Second, objects used for the same step were in closer proximity than those used for different task steps. Third, to encourage task monitoring, a call bell mounted on a red stand was placed to the right of the last object for each item. The words "check your work" were printed on the red

TABLE 1
Neuropsychological protocol

Cognitive function	*Test*	*Description*	*References*
Dementia severity	MMSE	30-item measure of dementia severity. Total scores range from 0 to 30; higher scores indicate better global cognitive functioning.	Folstein et al., 1975
Executive functions	Phonemic Fluency (FAS)	The dependent variable is the number of words produced in 60 s beginning with F, A, or S, excluding proper nouns.	Spreen & Strauss, 2006; Gourovitch et al., 2000
	Clock Drawing Test	Participants were asked to (a) draw a clock with the hands set to ten after eleven and (b) copy a drawing of a clock. Ten possible errors were scored on each trial.	Consentino, Jefferson, Chute, Kaplan, & Libon, 2004; Goodglass & Kaplan, 1983; Libon, Malamut, Swenson, Sands, & Cloud, 1996; Libon, Swenson, Barnoski, & Sands, 1993; Samton et al., 2005
Language	Boston Naming Test	The dependent variable is the number of pictures correctly named.	Kaplan, Goodglass, & Weintraub, 1983; Baum, Edwards, Yonan, & Storandt, 1996
	Category Fluency (Animals)	The dependent variable is the number of animals generated in 60 seconds.	Spreen & Strauss, 2006; Mummery et al., 1996
Episodic memory	PVLT	Participants were asked to remember a 9-word list, as on the CVLT. The primary dependent variable was the accuracy on the delayed recognition memory task (recognition discriminability). Total learning across Trials 1–5 was also calculated as an ancillary measure of learning.	Libon et al., 2005

Note. MMSE = Mini-Mental State Examination. PVLT = Philadelphia Repeatable Verbal Learning Test. CVLT = California Verbal Learning Test.

stand. The Standard and UC-NATs were administered across two sessions, separated by at least 1 week but no more than 4 weeks, to minimize the effects of practice. The order of conditions was counterbalanced across participants.

Evaluation of everyday action measures

Video recordings of the Standard NAT and UC-NAT were performed at the time of testing and were available for reanalysis. All recordings were coded for errors of action (i.e., acts that deviate from the prescribed code of behavior or fail to achieve what should be done given the procedural context and instructions). The NAT manual describes an extensive coding system for error analysis (Comprehensive Error Score, CES; Schwartz et al., 2003). The CES for participants in the Standard and User-Centered conditions was available from records of prior studies. The CES does not capture errors that are not fully executed (i.e., microslip). Microslips are operationally defined as the initiation and termination of an incorrect action before the error was completed (see also Blavier et al., 2005; Botvinick & Bylsma, 2005; Giovannetti, Schwartz, & Buxbaum, 2007). This includes reaching for or picking up an incorrect item or initiating a behavior that is dissonant with the task goal. For example, picking up, but not using, the garden shears instead of the scissors in the gift-wrapping task would be coded as a microslip. Microslips reflect instances of rapid and less effortful detection/correction and are highly relevant to the study of error monitoring (Norman, 1981; Smid, Mulder, & Mulder, 1990); thus, they were coded for the present study. For all analyses, microslips were added to CES errors (i.e., total errors).

Error detection and correction

Error monitoring was analyzed using a coding system for error detection and correction modeled on that of previous studies (Giovannetti, Libon, & Hart, 2002; Hart, Giovannetti, Montgomery, & Schwartz, 1998) and recently validated in dementia (Bettcher et al., 2008). Two raters blind to participant characteristics independently coded NAT videotapes for total errors, error detection, and error correction (see Bettcher et al., 2008, for description of coding method and examples; cursory description provided in Table 2). Disagreements between the coders were resolved through discussion and re-review of videotapes. Interrater reliability was assessed for 15 participants selected randomly from the sample. The raters demonstrated 97.8% agreement in coding all errors as detected or corrected (Cohen's kappa .95).

Error detection is operationally defined as an acknowledgement of mismatch between an individual's executed activity and the prescribed code of behavior for the task. Error detection rates cannot be meaningfully interpreted without the context of the number of errors generated. For example, a raw score of 5 detected errors reflects very

TABLE 2
Error detection and correction coding guide

Monitoring process		*Description*
Error detection	Verbalization	Individual offers a verbalization indicative of error recognition or a verbalization suggestive of possible task mismatch
	Failed attempt to correct	Individual demonstrates a failed attempt to rectify the error. This must include a visual effort to alter the previously committed error.
	Error correction	See entries below
Error correction	Microslip correction	Individual initiates an incorrect action, but stops himself or herself before the errant action is completed
	Overt correction	Individual rectifies or negates their action at any point in the task.

different error monitoring abilities in an participant with 5 total errors versus a participant with 20 total errors. As such, the proportion of total errors that was detected was calculated for each participant (proportion detected = total errors detected/total errors; see Bettcher et al., 2008; Giovannetti et al., 2002; Hart et al., 1998). In addition to the number of errors detected, subsequent rectification of these detected errors was calculated (i.e., error correction). There is overlap between the detection and correction categories, as we assumed that all corrected errors also were "detected." Therefore, with respect to correction, we were interested in knowing the proportion of "detected" errors that were subsequently corrected. This was calculated as follows: proportion detected corrected = total errors corrected / total errors detected (see Table 3).

Procedure

Participants signed institution review board informed consent forms (in compliance with institution regulations for human data) and were compensated $15.00 per session. Neuropsychological data were obtained as part of a clinical evaluation on the same day that participants were enrolled.

Statistical analyses

In order to answer the question of whether the UC-NAT intervention was associated with higher rates of error detection and correction than the Standard NAT, separate within-sample comparisons were performed for NAT proportion detected and proportion detected–corrected. Paired-sample *t* tests were used for variables that were normally distributed (or could be transformed). Wilcoxon tests were performed with non-normal variables that could not be transformed. Specifically, the outcome variable proportion detected was normally distributed and thus was not transformed. The outcome variables total errors (positive skew) and proportion detected–corrected (negative skew) did not meet assumptions for normality, and thus square-root and logarithmic transformations were performed; however, this did not correct the skewed distribution for proportion detected–corrected. As a result, nonparametric, Wilcoxon tests and Spearman rank order correlations were used for statistical analyses involving this outcome variable (as denoted in Table 4 and Table 5). Effect sizes for all UC-NAT versus Standard NAT analyses were estimated using Cohen's *d* calculations (Cohen, 1988).

To address the exploratory hypothesis that there would be a shift in neuropsychological processes responsible for error monitoring across the NAT conditions, Pearson correlation coefficients were calculated to evaluate the relations between UC-NAT/Standard NAT proportion detected and proportion detected–corrected and neuropsychological test scores (i.e., MMSE and independent neuropsychological measures; Spearman correlations were conducted for all analyses involving proportion detected–corrected). When more than one neuropsychological variable was significantly correlated to the detection/correction variables, stepwise regression analyses were performed to determine the multivariate relations (i.e., which cognitive measures or combinations of measures were predictive of error monitoring).[1]

TABLE 3
Error detection and correction variables

NAT errors	*Total errors (CES errors + microslips)*
Error detection	Proportion detected (total detected/total errors)
Error correction	Proportion detected–corrected (total corrected/total detected)

Note. NAT = Naturalistic Action Test. CES = Comprehensive Error Score.

RESULTS

Sample characteristics and NAT errors

Participants had an average age of 79.9 years (*SD* = 5.58) and 11.7 years of education (*SD* = 1.87). Mean MMSE

[1]Stepwise regressions are not considered ideal when a large corpus of research literature is available to support one's hypothesis; however, given the limited empirical work conducted on error monitoring and dementia, this form of regression was employed to isolate predictors.

TABLE 4
Error monitoring performance across conditions

	Standard NAT M (SD)	*UC-NAT M (SD)*	*z/t*	*p*	*Cohen's d*
Total errors	18.2 (11.6)	13.2 (9.3)	$z = -3.67$	.00	.50
Total detected	5.37 (4.8)	4.37 (3.3)	$t = 1.23$	.23	.24
Total corrected	4.74 (4.2)	3.29 (2.6)	$t = 1.84$	.08	.42
Proportion detected	32.9 (21.5)	37.1 (21.4)	$t = -1.04$	.30	.19
Proportion detected–corrected	82.5 (22.2)	78.7 (24.4)	$z = -0.58$	.50	.16

Note. NAT = Naturalistic Action Test. UC-NAT = User-Centered Naturalistic Action Test.

TABLE 5
Correlations for neuropsychological and error monitoring variables by condition

	Standard NAT		*UC-NAT*	
	Proportion detected r	*Proportion detected–corrected r_s*	*Proportion detected r*	*Proportion detected–corrected r_s*
MMSE	.34*	.19	.43**	.39**
FAS	.33*	.45**	–.01	–.06
Clock Drawing	–.23	–.07	–.19	–.20
Boston Naming Test	.23	.29	.13	.33*
Animal Fluency	.06	.23	.20	.16
PVLT Trials 1–5	.38**	.30	.19	.40**
PVLT Discriminability	.02	.18	.03	.10

Note. MMSE = Mini-Mental State Examination. PVLT = Philadelphia Repeatable Verbal Learning Test. NAT = Naturalistic Action Test. UC-NAT = User-Centered Naturalistic Action Test.
$^*p < .05$. $^{**}p < .025$.

TABLE 6
Mean neuropsychological test scores

	M	*SD*	*Minimum*	*Maximum*
MMSE	22.55	2.90	16	28
Phonemic Fluency (FAS)	23.31	10.34	7	52
Clock Drawing Test	4.82	2.99	1	10
Boston Naming Test	37.59	13.18	11	53
Category Fluency (Animals)	9.58	4.22	3	22
PVLT Trials 1–5 (Total Learning)	20.68	6.81	7	35
PVLT Recognition Discriminability	74.12	14.58	37	100

Note. MMSE = Mini-Mental State Examination. PVLT = Philadelphia Repeatable Verbal Learning Test.

scores were 22.55 ($SD = 2.9$), suggesting mild to moderate overall impairment (See Table 6 for a summary of neuropsychological test performance). In addition, the sample was composed of individuals who were primarily female (81%).

Consistent with previously published data, participants produced significantly fewer total errors in the User-Centered (UC-NAT) condition than in the Standard condition. Note that total errors in this study included microslips, which were not coded in our prior work. The proportion of errors accounted for by microslips was relatively small, equating to approximately 12% of the total error rate (Standard condition: $M = 12.66$, $SD = 12.55$; UC-NAT condition: $M = 12.10$, $SD = 11.44$) and 35% of the detected error rate for each group (Standard condition: $M = 35.55\%$, $SD = 27.83$; UC-NAT condition: $M = 35.723\%$, $SD = 32.42$).

Hypothesis testing: Error monitoring in UC-NAT versus Standard NAT conditions

Counter to our predictions, the UC-NAT condition did not improve error detection or correction rates relative to the Standard condition (see Table 4).[2] There was no effect of condition order on error rates or error monitoring ($p > .05$ for all dependent variables).

Consistent with previous studies (Bettcher et al., 2008), the participants' total error rate for each condition was not significantly related to the proportion of errors

[2]This analysis also was performed with raw error detection scores, controlling for total errors across conditions (i.e., a repeated measures general linear model, with total errors as a covariate). Comparable, nonsignificant monitoring results were found ($p = .6$) using this approach.

detected (UC-NAT, $r = -.28$; Standard NAT, $r = -.25$). Thus, the null results cannot be attributed to differences in error production rates.[3] Although it was not statistically significant, participants evidenced a trend towards detecting more errors in the UC-NAT condition than in the Standard NAT condition. Interestingly, the UC-NAT condition resulted in slightly lower error correction rates than the Standard NAT condition, although participants corrected the majority of detected errors in both conditions.

Exploratory hypothesis testing: A shift in neuropsychological processes related to error detection/correction

Consistent with our prediction, a measure of executive control functioning, FAS (Verbal Fluency), was significantly related to both proportion detected and proportion detected–corrected in the Standard NAT condition, but was not significantly related to either error-monitoring variable in the UC-NAT condition. In fact, correlations between this neuropsychological test variable and error monitoring in the UC-NAT condition approximated 0, which stands in contrast to the more robust relations observed in the Standard NAT condition (See Table 5); t tests on the between-condition correlations were conducted, verifying the significant *difference* between FAS and error monitoring variables by condition (proportion detected: $t = 1.94$, $p = .03$; proportion detected–corrected: $t = 2.40$, $p = .01$).

Study results yielded a significant association between The Boston Naming Test (BNT) and proportion detected–corrected in the UC-NAT condition. Notably, this measure was not significantly related to proportion detected in the UC-NAT condition, nor was it significantly related to either error monitoring variable in the Standard condition. This stated, the difference in the proportion detected–corrected correlations for each group (UC-NAT vs. Standard NAT) was not statistically significant ($t = 0.18$, $p = .57$).

In addition to the executive control and language measures detailed above, the MMSE was also related to proportion detected in the Standard condition, as well as both error monitoring variables in the UC-NAT condition. Similarly, a general measure of learning (total learning, Trials 1–5; Philadelphia Repeatable Verbal Learning Test, PVLT) was related to proportion detected in the Standard condition ($r = .38$, $p = .02$), and proportion detected–corrected ($r_s = .4$, $p = .01$) in the User-Centered condition. Considering that these measures are cognitively nonspecific and multifactorial, they were both considered proxies of general cognitive functioning in the current study. There were no significant relations noted between Animal Fluency, Clock Drawing Test, or delayed PVLT (recognition memory measure) and error monitoring variables in both conditions.

Finally, regression analyses were conducted to isolate the best neuropsychological predictors for each condition. In terms of the Standard NAT condition, a stepwise regression analysis (MMSE, Total Learning, and FAS entered in to the model) indicated that the best model for proportion detected accounted for only 10.2% of the variance and included a trend towards FAS being the sole predictor ($\beta = .32$; $t = 2.00$; $p = .06$). Considering that proportion detected–corrected was solely related to FAS in the Standard condition, additional regression analyses were not conducted for this outcome variable. In terms of the UC-NAT condition, a stepwise regression analysis (MMSE, Total Learning, and BNT entered in to the model) indicated that the best model for proportion detected–corrected accounted for 42% of the variance and included only MMSE and BNT as predictors (MMSE: $\beta = .31$; $t = 2.07$; $p = .04$; BNT: $\beta = .46$; $t = 3.07$; $p = .004$).

DISCUSSION

The overarching aim of this study was to evaluate the impact of an environmental intervention on error-monitoring deficits in dementia; to our knowledge, this investigation represents the first naturalistic study to empirically test a neuropsychological model of error-monitoring breakdown in the context of an intervention. The results showed that despite a significant improvement in overall action performance (i.e., errors), the intervention did not appreciably alter error monitoring. Note that the error monitoring variables were calculated as proportions rather than raw numbers (Hart et al., 1998); thus, the reduced error production rate in the UC-NAT condition should not have affected the error detection and correction variables. Furthermore, total errors were not associated with error monitoring variables in either condition.

The UC-NAT was designed to minimize environmental distractibility and improve sequence ordering, based on a body of work on stimulus–response compatibility and script execution (Fitts & Seeger, 1953; Giovannetti, Bettcher, et al., 2007). Thus, the UC-NAT does not purport to address task knowledge or language deficits, nor does it include a training session. As such, the adaptations implemented in the intervention are directed primarily at execution errors (or slips): namely, errors generated when the individual knows what to do, but experiences difficulty with executing this knowledge. In contrast, errors resulting from degraded or incomplete knowledge (planning errors; mistakes) are likely to remain undetected in this intervention, as the UC-NAT offers no explicit support for semantic knowledge deficits.

[3]Considering that individuals with mild dementia were included in the study, participants with a MMSE > 23 were examined separately to determine whether their total errors rates were large enough to adequately assess detection and correction rates. All participants with a MMSE > 23 generated errors on both tasks (Standard: $M = 15.9$, $SD = 11.6$; UC: $M = 9.0$, $SD = 5.9$), suggesting that ceiling effects did not detrimentally impact the analyses.

We proposed that the UC-NAT condition would reduce the likelihood of execution errors (slips); in addition, we hypothesized that error detection would increase due to the environmental support and structure offered in the intervention, and the results were in the direction of this hypothesis; however, it is not altogether surprising that an appreciable increase in error monitoring was not observed. We suspect that by ameliorating execution errors in the intervention (discussed further in the following section), we essentially tipped the outcome towards error monitoring problems stemming from degraded task knowledge (mistakes), which are associated with a low probability of error detection/correction. Although the exploratory neuropsychological outcome data provide preliminary support for this assertion, it is also possible that the intervention did not sufficiently reduce executive demands to facilitate error monitoring, which may necessitate greater executive control over everyday action performance. Future studies should incorporate a broad-based evaluation of executive functions as well as address and isolate the role of task knowledge in error monitoring. In addition, the raw number of error detections remains quite low for both conditions (Standard and UC-NAT), and thus the interpretation of these findings are based on a relatively small number of observations; as such, caution in generalizing the results to other interventions remains prudent. Importantly, improvements in performance as a function of increased environmental structure are a commonly reported finding in the literature; however, this study demonstrates that while overall action performance improves, detection and correction of the remaining action errors remain *relatively* unaffected by the intervention.

An additional exploratory aim of the current study was to evaluate the role of neuropsychological processes in the intervention and control conditions in the hope of developing a more comprehensive model of error monitoring in dementia. Consistent with our predictions, participants evidenced a shift in neuropsychological processes associated with action error monitoring in the UC-NAT condition. Specifically, a measure of executive functioning abilities (FAS, Verbal Fluency) was the best predictor of error detection and correction in the Standard NAT condition and notably demonstrated a robust relation with error correction ($r = .45$); however, FAS no longer served as a predictor for error monitoring in the UC-NAT condition and strikingly evidenced almost no relation with the error detection and correction measures ($r = -.01$ and $-.06$, respectively). This offers preliminary, albeit very tentative, support for our hypothesis that the UC-NAT decreased the executive demands associated with error monitoring; it is also consistent with literature citing verbal fluency as a predictor of competency status (Marson, Cody, Ingram, & Harrell, 1995).

In contrast with our predictions, error monitoring variables in the Standard NAT were not significantly related to the Clock Drawing Test. It is unclear why the Clock Drawing Test did not serve as an explanatory role in our models, as previous work from our lab has isolated this measure as relevant for error correction processes (Bettcher et al., 2008). However, even in the face of statistically significant findings, the explanatory role of traditional neuropsychological tests for action monitoring has consistently been quite small in prior studies. It may be that executive measures tapping retrieval, attention, and mental flexibility are more important in error monitoring than those mediated by visuoconstructional abilities.

It is also important to highlight that error detection in the UC-NAT was not predicted by any isolated test of language or delayed episodic memory. The lack of association with neuropsychological indices of specific cognitive functions does not appear to be due to restricted range or "floor effects," as evidenced in Table 6. One potential reason for this finding (or lack thereof) is that the cognitive protocol used in the study may not have adequately addressed the neuropsychological processes involved in error detection. Although we attempted to incorporate commonly utilized measures of executive functions, language, and memory for the current study, a comprehensive neuropsychological evaluation was not employed. As a result, measures not included in the current study may have been more optimal predictors of action error monitoring in the UC-NAT.

An additional finding of the study was that error correction processes in the UC-NAT were related to a language measure (BNT), a general cognitive functioning test (MMSE), and a general, multifactorial measure of learning (Trials 1–5, PVLT), although the BNT and MMSE were determined to be the best predictors of this monitoring variable. Importantly, however, the relation between the BNT and error correction was modest ($r = .33$) and was only slightly higher than nonsignificant correlations obtained in the Standard condition. Although language measures did not serve as stronger predictors in the UC-NAT condition relative to the Standard NAT condition, study results suggest that a more focused evaluation of language processes may be beneficial to understanding error monitoring breakdowns in dementia. It is currently unclear what role language functioning, and more specifically task knowledge, may play in everyday action monitoring for dementia patients; furthermore, recent studies suggest that semantic knowledge for objects or everyday action may be represented differently than knowledge for other domains (e.g., animals, foods, etc.; Patterson, Nestor, & Rogers, 2007). Thus, commonly used language measures (e.g., BNT) may not be ideal for assessing semantic knowledge in dementia patients, as error monitoring may be more strongly related to action-specific semantic measures.

Limitations

Although the study offered a theory-based, naturalistic approach to improving action performance in dementia patients due to execution deficits, it also faced limitations. First, a brief clinical neuropsychological protocol was employed in the study; although measures were carefully chosen based on prior studies and neuroanatomical considerations (i.e., dorsolateral prefrontal cortex for assessment of executive functions; temporal lobe for assessment of task knowledge and language; Cosentino et al., 2006; Gourovitch et al., 2000; Mummery, Patterson, Hodges, & Wise, 1996; Ravnkilde, Videbech, Rosenberg, Gjedde, & Gade, 2002; Samton et al., 2005), a more comprehensive protocol that included specific measures of task knowledge and extensive indices of executive functions may have shed additional light on the error monitoring variables assessed.

Second, a potential concern regarding the assessment of error monitoring in the two conditions is the lack of explicit instruction to detect and correct errors. Thus, it is possible that participants failed to monitor their errors because they did not think it was part of the overarching task goal. Of note, previous work from our lab has documented that healthy control participants detect and subsequently correct their errors on the Standard condition at a high rate, suggesting that monitoring behaviors occur without directive requests (Bettcher et al., 2008; Giovannetti, Libon, & Hart, 2002; see also Giovannetti, Schwartz, et al., 2007).

Third, the determination of whether an individual is truly aware of his or her errors remains a challenging feature of this research. This study relied on verbalizations and failed or successful attempts to correct errors to indicate detection; yet, error awareness may have occurred without overt signs of detection. As a result, error detection variables may have underestimated participants' error awareness. However, importantly, naturalistic studies offer an opportunity to capture online error monitoring behaviors in an ecologically valid context, thereby providing insight into how dementia participants overtly respond to and handle their everyday action errors.

Future directions

Consideration of the current study's findings suggests need for additional research on the role of task knowledge and executive functions in everyday action monitoring. The current investigation focused on one piece (i.e., execution errors) of a larger error monitoring puzzle in dementia patients. Although executive control processes are important and necessary facets of error monitoring, they may not represent the linchpin of our patients' error detection and correction deficits. Thus, future studies should target task knowledge impairments in patients to isolate the relative importance of this cognitive process.

Alternatively, the breakdown of executive functions or "execution errors" may be germane to specific dementia populations and not others; for example, individuals with subcortical dementias characterized by disproportionately more executive control problems than memory/language deficits may benefit more from this intervention. If so, this would suggest that these patients experience a primary deficit in error monitoring, rather than a secondary monitoring deficit due to degraded task knowledge.

In sum, evaluating and ultimately rehabilitating error monitoring problems in dementia patients is a critical component to our understanding of this neurodegenerative disease. Errors, and more importantly poor error detection/correction, may lead to blunders that jeopardize patients' autonomy and safety. The current study purported to address error monitoring impairments in dementia using a naturalistic, theory-based intervention for execution deficits. Results suggest that while the UC-NAT intervention improves overall performance, it does not significantly increase error detection/correction relative to the Standard NAT. However, consistent with its hypothesized mechanism of action, it does appear to impact the neuropsychological processes associated with error monitoring and thus has considerable implications for a neuropsychological model of error monitoring in dementia.

Original manuscript received 3 August 2010
Revised manuscript accepted 23 April 2011
First published online day month year

REFERENCES

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.

Band, G. P. H., & Kok, A. (2000). Age effects on response monitoring in a mental-rotation task. *Biological Psychology, 51*, 201–221.

Baum, C., Edwards, D., Yonan, C., & Storandt, M. (1996). The relation of neuropsychological test performance to performance on functional tasks in dementia of the Alzheimer's type. *Archives of Clinical Neuropsychology, 11*, 69–75.

Bettcher, B. M., & Giovannetti, T. (2009). From cognitive neuroscience to geriatric neuropsychology: What do current conceptualizations of the action error handling process mean for older adults? *Neuropsychology Review, 19*, 64–84.

Bettcher, B. M., Giovannetti, T., MacMullen, L., & Libon, D. J. (2008). Error detection and correction patterns in dementia: A breakdown of error monitoring processes and their neuropsychological correlates. *Journal of the International Neuropsychological Society, 14*(2), 199–208.

Blavier, A., Rouy, E., Nyssen, A., & De Keyser, V. (2005). Prospective issues for error detection. *Ergonomics, 48*(7), 758–781.

Botvinick, M. M., & Bylsma, L. M. (2005). Distraction and action slips in an everyday task: Evidence for dynamic representation of task context. *Psychonomic Bulletin & Review, 12*(6), 1011–1017.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cosentino, S., Jefferson, A., Chute, D., Kaplan, E., & Libon, D. (2004). Clock drawing errors in dementia: Neuropsychological and neuroanatomical considerations. *Cognitive and Behavioral Neurology, 17*(2), 74–84.

Cosentino, S., Chute, D., Libon, D., Moore, P., & Grossman, M. (2006). How does the brain support script comprehension? A study of executive processes and semantic knowledge in dementia. *Neuropsychology, 20*(3), 307–318.

Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior Research Methods, Instruments, & Computers, 28*, 1–11.

Falkenstein, M., Hoormann, J., & Hohnsbein, J. (2001). Changes of error-related ERPs with age. *Experimental Brain Research, 138*(2), 258–262.

Fitts, P., & Seeger, C. (1953). Spatial characteristics of stimulus and response codes. *Journal of Experimental Psychology, 46*, 193–210.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state." A practical method for grading the cognitive

state of patients for the clinician. *Journal of Psychiatric Research, 12*(3), 189–198.
Giovannetti, T., Bettcher, B. M., Libon, D. J., Brennan, L., Sestito, N., & Kessler, R. K. (2007). Environmental adaptations improve everyday action performance in Alzheimer's disease: Empirical support from performance-based assessment. *Neuropsychology, 21*, 448–457.
Giovannetti, T., Libon, D. J., Buxbaum, L. J., & Schwartz, M. F. (2002). Naturalistic action impairments in dementia. *Neuropsychologia, 40*(8), 1220–1232.
Giovannetti, T., Libon, D. J., & Hart, T. (2002). Awareness of naturalistic action errors in dementia. *Journal of the International Neuropsychological Society, 8*(5), 633–644.
Giovannetti, T., Schwartz, M. F., & Buxbaum, L. J. (2007). The coffee challenge: A new method for the study of everyday action errors. *Journal of Clinical and Experimental Neuropsychology, 29*, 690–705.
Goodglass, H., & Kaplan, E. (1983). *The assessment of aphasia and related disorders* (2nd ed.). Philadelphia, PA: Lea and Febiger.
Gourovitch, M. L., Kirkby, B. S., Goldberg, T. E., Weinberger, D. R., Gold, J. M., Esposito, G., et al. (2000). A comparison of rCBF patterns during letter and semantic fluency. *Neuropsychology, 14*(3), 353–360.
Hart, T., Giovannetti, T., Montgomery, M. W., & Schwartz, M. F. (1998). Awareness of errors in naturalistic action after traumatic brain injury. *Journal of Head Trauma Rehabilitation, 13*(5), 16–28.
Ito, J., & Kitagawa, J. (2006). Performance monitoring and error processing during a lexical decision task in patients with Parkinson's disease. *Journal of Geriatric Psychiatry and Neurology, 19*(1), 46–54.
Kaplan, E. F., Goodglass, H., & Weintraub, S. (1983). *The Boston Naming Test* (2nd ed.). Philadelphia, PA: Lea & Febiger.
Libon, D. J., Malamut, B. L., Swenson, R., Sands, L. P., & Cloud, B. S. (1996). Further analyses of clock drawings among demented and nondemented older subjects. *Archives of Clinical Neuropsychology, 11*(3), 193–205.
Libon, D. J., Schmidt, K. S., Gallo, J. L., Penney, D. L., Swenson, R., Giovannetti, T., et al. (2005). The prototypicality of extra-list intrusion errors produced on a serial list-learning task by patients with mild dementia [Abstract]. *Journal of the International Neuropsychological Society, 11*(Suppl. 1), 92–93.
Libon, D. J., Swenson, R. A., Barnoski, E. J., & Sands, L. P. (1993). Clock drawing as an assessment tool for dementia. *Archives of Clinical Neuropsychology, 8*(5), 405–415.
Marson, D. C., Cody, H. A., Ingram, K. K., & Harrell, L. E. (1995). Neuropsychological predictors of competency in Alzheimer's disease using a rational reasons legal standard. *Archives of Neurology, 52*(10), 955–959.
Mathalon, D. H., Whitfield, S. L., & Ford, J. M. (2003). Anatomy of an error: ERP and fMRI. *Biological Psychology, 64*(1–2), 119–141.
McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology, 34*(7), 939–944.
Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*, 167–202.
Mioshi, E., Kipps, C. M., Dawson, K., Mitchell, J., Graham, A., & Hodges, J. R. (2007). Activities of daily living in frontotemporal dementia and Alzheimer disease. *Neurology, 68*, 2077–2084.
Mummery, C. J., Patterson, K., Hodges, J. R., & Wise, R. J. (1996). Generating "tiger" as an animal name or a word beginning with T: Differences in brain activation. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 263*(1373), 989–995.
Nieuwenhuis, S., Ridderinkhof, K. R., Talsma, D., Coles, M. G., Holroyd, C. B., Kok, A., et al. (2002). A computational account of altered error processing in older age: Dopamine and the error-related negativity. *Cognitive and Affective Behavioral Neuroscience, 2*(1), 19–36.
Norman, D. A. (1981). Categorization of action slips. *Psychological Review, 88*, 1–15.
Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews: Neuroscience, 8*, 976–987.
Ravnkilde, B., Videbech, P., Rosenberg, R., Gjedde, A., & Gade, A. (2002). Putative tests of frontal lobe function: A PET-study of brain activation during Stroop's test and verbal fluency. *Journal of Clinical & Experimental Neuropsychology, 24*(4), 534–547.
Reason, J. (1990). *Human error*. Cambridge, UK: Cambridge University Press.
Samton, J. B., Ferrando, S. J., Sanelli, P., Karimi, S., Raiteri, V., & Barnhill, J. W. (2005). The clock drawing test: Diagnostic, functional, and neuroimaging correlates in older medically ill adults. *Journal of Neuropsychiatry and Clinical Neurosciences, 17*(4), 533–540.
Schwartz, M. F., Buxbaum, L. J., Ferraro, M., Veramonti, T., & Segal, M. (2003). *The Naturalistic Action Test*: Bury St. Edmunds, UK: Thames Valley Test Company.
Schwartz, M. F., Segal, M., Veramonti, T., Ferraro, M., & Buxbaum, L. J. (2002). The Naturalistic Action Test: A standardized assessment for everyday action impairment. *Neuropsychological Rehabilitation, 12*(4), 311–339.
Sellen, A. J., & Norman, D. A. (1992). The psychology of slips. In B. J. Baars (Ed.), *Experimental slips and human error: Exploring the architecture of volition*. New York, NY: Plenum Press.
Sirigu, A., Zalla, T., Pillon, B., Grafman, J., Agid, Y., & Dubois, B. (1995). Selective impairments in managerial knowledge following pre-frontal cortex damage. *Cortex, 31*, 201–316.
Smid, H. G. O., Mulder, G., & Mulder, L. J. M. (1990). Selective response activation can begin before recognition is complete: A psychophysiological and error analysis of continuous flow. *Acta Psychologica, 74*, 169–201.
Spreen, O., & Strauss, E. (2006). *A compendium of neuropsychological tests: Administration, norms and commentary* (3rd ed.). New York, NY: Oxford University Press.
Ullsperger, M., & von Cramon, D. Y. (2006). The role of intact frontostriatal circuits in error processing. *Journal of Cognitive Neuroscience, 18*(4), 651–664.