Practice Effect and Beyond: Reaction to Novelty as an Independent Predictor of Cognitive Decline Among Older Adults

Yana Suchy, Matthew L. Kraybill, AND Emilie Franchow

Department of Psychology, University of Utah, Salt Lake City, Utah

Abstract

Practice Effects (PE) have been gaining interest as an early marker of pathological cognitive decline among older adults, with cognitively compromised individuals exhibiting diminished or absent PE, presumably due to reduced ability to learn. However, the opposite pattern has also been observed, with MCI participants showing *larger* PEs than controls. In this prospective cohort study, we examined the possibility that individuals with incipient cognitive decline may be more "thrown" by task novelty, which may inflate PE due to diminished performance during the first exposure to the task. We assessed Novelty Effect (NE) and Learning (LRN) on a motor task in 50 community-dwelling independent older adults who expressed a concern about their cognition. Results showed that larger NE was associated with greater cognitive decline 17 months later, reliably classifying participants into decliners and nondecliners. LRN did not independently explain any variance in future cognitive *change*, but moderated the relationship between NE and decline and correlated with the *level* of cognition at baseline and follow-up. These findings highlight the differing contributions of NE and LRN to PE, and demonstrate that NE may be sensitive to depletion of cognitive reserve among individuals who are on the verge of exhibiting a reliable cognitive decline. (*JINS*, 2011, *17*, 101–111)

Keywords: Practice effect, Novelty, Learning, Aging, Cognitive reserve, Cognitive decline

INTRODUCTION

Practice Effect (PE) refers to an improvement in performance observed on repeated administrations of the same, or similar, measure (Falleti, Maruff, Collie, & Darby, 2006; Maassen, Bossema, & Brand, 2008). PE is thought to reflect both explicit (i.e., declarative) learning of the test items, and the implicit (i.e., procedural) learning of the test taking strategy independent of the specific test content (Busch, Chelune, & Suchy, 2005). Research has shown that the size of the PE is in part dependent on the characteristics of the test-taker, such that PE decreases with age (Roennlund, Loevden, & Nilsson, 2008; Thompson, 1997) and is smaller in people with lower IQs (Rapport, Brines, Axelrod, & Theisen, 1997). These individual differences suggest that PE may reflect more than just a psychometric nuisance; rather, it may reflect a cognitive ability in its own right.

Consistent with this latter notion, research has shown that PEs generated by repeated administration of cognitive tests both within a single testing session (Darby, Maruff, Collie, & McStephen, 2002) and 1 week after the initial testing (Cooper, Lacritz Weiner, Rosenberg, & Cullum, 2004) were smaller than expected or absent among individuals with mild cognitive impairment (MCI). Because MCI represents a predictor of future cognitive decline and conversion to dementia, Duff and colleagues reasoned that small or absent PE could also predict future cognitive performance. Consistent with their hypothesis, they found that PEs[1] on memory, attentional, and processing speed measures contributed significantly to the prediction of cognitive performances on two follow-up visits 3 to 12 months later among MCI and HIV patients (Duff et al., 2007). These findings have since been replicated, with PEs predicting 1-year follow-up scores in healthy, community-dwelling older adults with either amnestic MCI or intact cognition (Duff, Beglinger, Moser, Paulsen, Schultz, & Arndt, in press). However, some inconsistencies in findings exist, as another study found that MCI

Correspondence and reprint requests to: Yana Suchy, Department of Psychology, University of Utah, 380 S. 1530 E., Room 502, Salt Lake City, UT 84112-0251. E-mail: yana.suchy@psych.utah.edu

[1] PEs were operationalized as the difference between test scores at screening and baseline sessions, which occurred 2 weeks apart. PEs were controlled for baseline performance.

participants exhibited larger, *not* smaller, PEs than controls on two delayed recall measures (Duff et al., 2008). In fact, once practice was taken into account, some MCI participants' performances became comparable to those of controls[2].

Although the latter findings (Duff et al., 2008) could be interpreted as paradoxically demonstrating that MCI patients learned more than controls, an alternative explanation can be offered. In particular, given that the performances of some of the MCI participants were comparable to those of controls on the second test administration, it is possible that the effect observed in the latter study was *not* a reflection of greater improvement in performance due to extra learning, but rather a reflection of greater initial *interference* with performance caused by the test novelty. In other words, the MCI participants may have been more "thrown" than the controls by their unfamiliarity with the tests, requiring multiple test administrations to overcome this initial effect.

The evidence for deleterious effects of novelty on task performance among individuals who are on a declining neurocognitive trajectory comes from several lines of research. First, it has been repeatedly demonstrated by functional imaging research that task novelty requires extra cognitive processing, evidenced by greater and more widespread brain activation early on during task performance (Brovelli, Laksiri, Nazarian, Meunier, & Boussaoud, 2008), followed by reduced demands on working memory systems once a task becomes familiar (Jansma, Ramsey, Slagter, & Kahn, 2001). Second, it is now well understood that cognitive reserve[3] serves as a protective factor against cognitive decline despite neuropathological changes in the brain (Andel, Vigen, Mack, Clark, & Gatz, 2006; Corral, Rodriguez, Amenedo, Sanchez, & Diaz, 2006; Stern, 2002). Cognitive reserve can also be viewed as an extra cognitive buffer that protects not only against the effects of neuropathology, but also against the deleterious effects of temporary increases in task demands, such as the presence of distractors, fatigue, unfamiliarity with the task, and stress levels (Garrett, Grady, & Hasher, 2010). Third, decreases in cognitive reserve are associated with increased and more wide-spread brain activation (Dickerson et al., 2005; Lenzi et al., 2009), thought to reflect an underlying biological marker of cognitive inefficiency, and/or greater cognitive effort. In fact, even the simple act of verbal comprehension shows more wide-spread functional magnetic resonance imaging (fMRI) activation among individuals with amnestic MCI or mild Alzheimer's disease (Bosch et al., 2010), suggesting that even processing the instructions for a new task is more taxing for individuals who are on a declining trajectory.

Taken together, research suggests that (a) dealing with novelty requires extra cognitive resources, (b) experiencing preclinical neurocognitive decline also requires extra cognitive effort to achieve performance on par with expectations, and (c) the extra cognitive demands imposed simultaneously by task novelty on the one hand and depleted cognitive reserve on the other may well overwhelm the limited available resources among individuals who are on a trajectory of a preclinical neurocognitive decline. In turn, when resources are overwhelmed, a temporary decrement in performance may occur as a task is first introduced, followed by a rebound in performance as task familiarity increases. Such temporary decrements followed by a rebound likely contribute, at least in some cases, to what is typically considered the effect of practice (in the form of improved performance on subsequent trials).

The purpose of the present study was to examine whether the inability to benefit from practice on the one hand and the susceptibility to the deleterious effects of novelty on the other could serve as two *independent* predictors for future cognitive decline. To that end, we conducted two screenings of cognitive status approximately 1 year apart in a sample of community-dwelling older adults who expressed some concerns about their cognition. To assess the participants' susceptibility to the effects of novelty and their ability to learn from practice, we administered a computerized motor learning[4] task that generates indices of (a) the deleterious effect of task novelty on motor planning latencies and (b) motor learning capacity across several practice trials (Suchy, Derbidge, & Cope, 2005; Suchy & Kraybill, 2007). We examined these two indices as predictors of future decline in cognition, while controlling for other potential confounds, including demographic characteristics and baseline cognitive performance[5], consistent with methodology used in prior research on performance *change* (Attix et al., 2009; Duff et al., 2007). We hypothesized that (a) larger Novelty Effects (NE) and poorer Learning (LRN) would be associated with future cognitive declines, and (b) NE and LRN would be mutually dissociable, representing two independent predictors.

METHOD

Participants

Seventy-five Caucasian older adults were originally recruited from the Salt Lake City community for this prospective cohort study, which advertised (*via* flyers and newspaper

---

[2] PEs were operationalized as the difference between baseline and follow-up standard scores after controlling for baseline performance. The relatively smaller PEs among controls in this study could not be fully explained by ceiling effects, as most scores did not appear to reach the ceiling.

[3] Cognitive reserve is typically operationalized as educational achievement or crystallized intelligence (Stern, 2002).

[4] A motor learning task, as opposed to more traditional cognitive measures, was used because motor output appears to be particularly sensitive to preclinical alterations in brain processing. For example, transfer of learned information into action has been shown to be impaired among individual with mild preclinical hippocampal atrophy (Gluck et al., 2006), and sensorimotor connectivity is altered among MCI individuals (Agosta et al., 2010). This particular motor learning task was selected because (a) it has been shown to detect the deleterious effects of novelty independent of observable levels of cognition (Suchy & Kraybill, 2007), and (b) performance on the task has been shown to correlate with cognitive (in particular executive) functioning (Kraybill & Suchy, 2008; Suchy, Kraybill, & Larson, 2010; Suchy & Kraybill, 2007), suggesting that performance on the task reflects central, rather than peripheral, processes.

[5] Controlling for initial level of cognition also controls for regression to the mean.

Table 1. Demographic, depressive, cognitive, and health characteristics of the sample divided by return status

	Returners (*n* = 50)	Dropouts (*n* = 25)
Age (years)		
M (SD)	69.46 (6.42)	70.75 (7.34)
Range	58–87	61–85
Education (years)		
M (SD)	14.70 (2.24)	14.64 (3.24)
Range	10–18	10–22
GDS (raw score)		
M (SD)	4.52 (4.16)	6.16 (6.69)
Range	0–23	0–26
DRS-2 baseline (raw score)		
M (SD)	139.44** (3.32)	135.08** (6.46)
Range	130–144	117–144
% Female	64%	60%
Hypertension	38%	52%
Heart disease	2%	16%
Chronic obstructive pulmonary disease	2%	8%
Sleep apnea	10%	20%
Small stroke	6%	12%
Seizure disorder	4%	0%

Note. M = mean; *SD* = standard deviation; DRS-2 = Mattis Dementia Rating Scale-2nd edition; GDS = Geriatric Depression Scale.
**Values differ at *p* < .001.

advertisements) seeking individuals who had concerns about their cognition but lived independently. Exclusion criteria accessed *via* self-report included color blindness, uncorrected vision or hearing problems, difficulty using the right hand, left-handedness, moderate to severe health problems, and a diagnosis of dementia (dementia was an exclusion criterion at baseline only). No participants had a history of brain injury or brain tumor. Some participants reported "mild" chronic health problems, including hypertension (*n* = 30), heart disease (*n* = 5), chronic obstructive pulmonary disease (*n* = 3), sleep apnea (*n* = 10), small stroke (*n* = 6), and seizure disorder (*n* = 2).

All 75 participants underwent baseline testing, but only 50 participants (66%) returned for follow-up testing that took place approximately 17 months after the initial evaluation. Attrition rate was in line with similar prospective cohort studies (Niti, Yap, Kua, & Ng, 2009; Stewart, Rand, Muldoon, & Kamarck, 2009). Attrition was due to participants not returning for follow-up testing for unspecified reasons (*n* = 11), being unreachable (e.g., disconnected phone number; *n* = 7), having other personal obligations such as caregiving responsibilities (*n* = 3), feeling that it was too far to travel (*n* = 2), and functional changes that made it too difficult to participate (i.e., onset of dementia; *n* = 2). Participants who completed follow-up testing were comparable to dropouts in terms of age, education, gender, and depression scores. The dropouts did, however, have significantly lower scores on the baseline Mattis Dementia Rating Scale as compared to those who returned, *t*(73) = 3.87, *p* < .001. See Table 1 for a summary of participant characteristics. Additional analyses on the characteristics of the dropouts, as they relate to the present study, are presented in the Supplementary Analyses section.

Procedures

Participants were screened over the telephone at both the initial and follow-up time points regarding exclusion criteria. On the day of testing, participants underwent informed consent procedures, followed by administration of the assessment measures (listed below). Participants received $30 for the initial session and $20 for the follow-up session (the follow-up testing lasted 2 hours, whereas the initial testing lasted 3 hours, as additional instruments were administered at baseline as part of a larger study). Participants were also provided brief feedback regarding their cognitive and depression screening results. This study and its procedures were approved by the University of Utah Institutional Review Board.

Instruments

Cognitive status and cognitive decline

Cognitive status was screened at both time points using the *Mattis Dementia Rating Scale 2nd edition* (DRS-2; Mattis, Jurica, & Leitten, 1988), which is a global measure of general cognitive status and specifically evaluates attention, abstraction, visual-constructional abilities, initiation and perseveration, and verbal and nonverbal short-term memory. We operationalized cognitive change as the difference between the baseline and the follow-up DRS-2 total raw scores[6], adjusted for the test–retest

---

[6] There were four participants whose DRS-2 scores at baseline were 144, that is, who reached the ceiling on this measure. Consequently, these participants could either stay the same, or show a decline in their scores at follow-up. Although reaching a ceiling constrained the range of possible change scores, this was not a major issue as the interest of the present study was not improvements, but rather declines. Although it is possible that participants with high scores at baseline would exhibit a regression to the mean at

interval. Specifically, we first subtracted the DRS-2 total raw score at baseline from the DRS-2 total raw score at follow-up, and then divided the difference by the given participant's test–retest interval (expressed in years). These scores then reflected annualized rate of change, in which all participants' change scores were equated regardless of their actual test–retest interval.

Depression screening

Depression was screened using the Geriatric Depression Scale (GDS; Yesavage, 1988), a measure developed for use with older adults. It was included as a means of characterizing the depressive symptomatology of the sample.

Novelty effect (NE) and learning (LRN) assessment

NE and LRN were assessed using the *Push-Turn-Taptap* (PTT) Task from the BDS-EV battery (Suchy et al., 2005). In this task, participants are asked to learn four different sequences of increasing length that consist of different permutations of three different hand movements, using a specialized response console (Figure 1). All participants were able to learn all sequences within the allotted number of trials (i.e., 10 trials). In the present study, we used two scores generated by this task: (a) an index of the effect of novelty (NE), and (b) an index of learning (LRN). These are described below.

(a)   NE: The NE was operationalized as the difference in the median motor planning latencies between the first and the second Block of the PTT task, with larger values representing increasingly deleterious effects of novelty. Motor planning latencies are assessed by measuring the amount of time (in ms) that elapses between completion of one trial and initiation of the next correctly executed trial (see Figure 2). Planning latencies have been shown to correlate with executive functioning, even after motor and processing speeds have been accounted for (Suchy, Kraybill, & Larson, 2010; Suchy & Kraybill, 2007). Please note that typically planning latencies are a function of the complexity, or the length, of a given sequence (Wright, Black, Immink, Brueckner, & Magnuson, 2004). Thus, progressively longer latencies are normally evident on Blocks 3 and 4 of the task, as the Blocks get progressively longer (Suchy et al., 2005; Suchy & Kraybill, 2007). However, past research has shown that the planning latencies on Block 1 of the PTT task are longer than those on Block 2 (Suchy et al., 2005; Suchy & Kraybill, 2007), despite the fact that Block 1 is the easiest (i.e., consisting of a sequence of only two hand movements, as compared to three hand movements on the Block 2). This paradoxical effect has been attributed to task novelty (Suchy & Kraybill, 2007). Greater NE values reflect a more deleterious impact of novelty on performance.
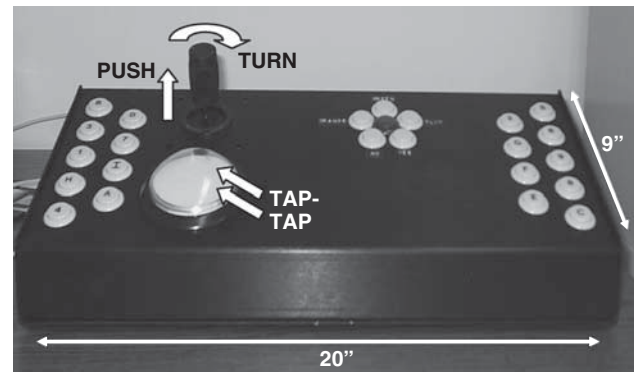
Fig. 1. The figure depicts the Behavioral Dyscontrol Scale-Electronic Version (BDS-EV) response console. It was used in the present study to assess learning and the effect of novelty. The figure originally appeared in Suchy et al. (2010). Understanding design fluency: Motor and executive contributions. *Journal of the International Neuropsychological Society, 16*(1), 26–37. Copyright 2010 by Cambridge University Press.

(b)   LRN: LRN was operationalized as the number of errors across the four blocks of the PTT task, with greater values reflecting poorer performance. Because participants are given sample trials at the beginning of each block, a lack of errors demonstrates the ability to benefit from practice, whereas a high number of errors reflects a failure to learn from practice. This score reflects both declarative learning (as participants can rehearse the words that go with the sequence), as well as procedural learning (as participants repeat the same motor sequence several times), making it conceptually parallel to the types of learning presumably underlying PE (Busch et al., 2005).

RESULTS

Preliminary Analyses

Validity check

To ascertain that the PTT task performances behaved as expected in terms of generating motor planning latencies that were susceptible to the effect of novelty (Block 1), as well as the normal expected increase in latencies as a function of increased length of the motor sequence (Blocks 3 and 4), we conducted three paired *t* tests, comparing each consecutive pair of blocks. The results were consistent with expectation (all *p* values <.035). See Figure 3.

Zero order correlations

Zero order correlations among predictors and covariates are presented in Table 2. As can be seen, while LRN was associated with both baseline and retest DRS-2 scores, NE was not. Also, NE and LRN were correlated, such that greater ability to learn was associated with larger NE, demonstrating that the two processes are not entirely orthogonal (see Figure 4). This likely

---

(*footnote continued*)
follow-up, essentially exhibiting a decline, this, too, was not a major issue, as we controlled for regression to the mean in our statistical analyses.
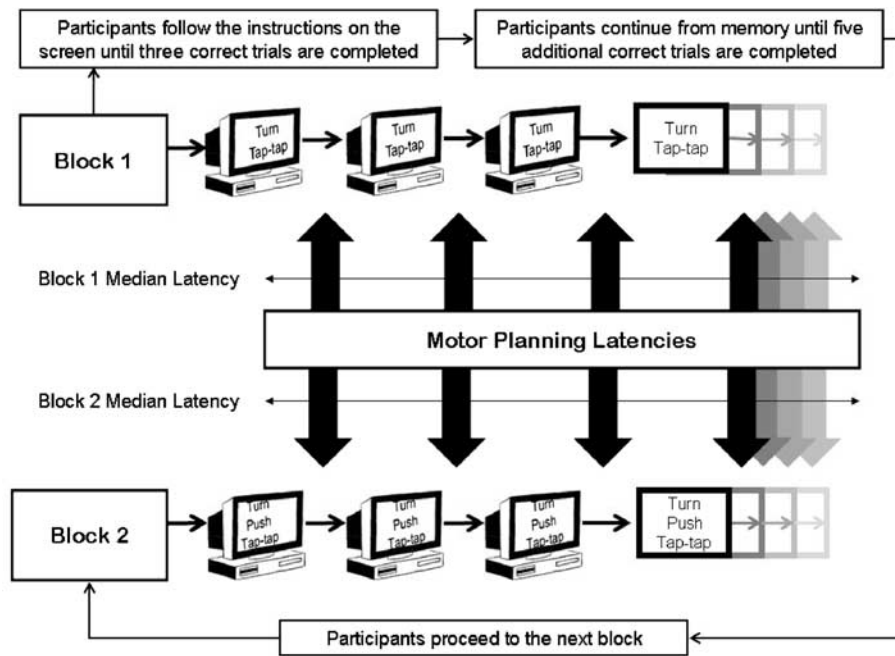
Fig. 2. The BDS-EV Push-Turn-Taptap task (Suchy et al., 2005) requires that participants learn different sequences (or permutations) of three specified hand movements, using a specialized response console. The three hand movements are "Push" – pushing the joystick forward; "Turn" – turning the joystick clockwise; and "Taptap" – double-tapping on the white dome of the response console (Figure 1). The task begins with Block 1, in which a two-movement sequence is presented on the computer screen, until three *correct* trials are completed. Following these three learning trials, participants continue to perform the sequence from memory, until accomplishing five additional *correct* trials. This completes the Block 1 of the task. After completing Block 1, participants move on to Block 2. In Block 2, a new, *longer* sequence is presented on the computer screen, and the above-described process is repeated. There is a total of four Blocks, each characterized by different and progressively longer sequences (only the first two Blocks are presented in the figure). Mistakes are followed by an audible tone, along with the presentation of the correct sequence on the computer screen and the highlighting of the next movement to be performed. Motor planning latencies (which were used for computation of the effect of novelty) are indicated in the figure by the thick black vertical arrows, and reflect the preparation time *before* initiation of each *correct* trial (only latencies preceding correctly executed trials are considered). For each block, median latencies across all correct trials are computed (indicated in the figure by double-headed horizontal arrows). The effect of novelty is operationalized in this study as the difference between the median latencies for Block 1 minus median latencies for Block 2. Adapted from "Understanding design fluency: Motor and executive contributions," by Y. Suchy, M. Kraybill, and J. G. L. Larson, 2010. *Journal of the International Neuropsychological Society, 16*(1), 26–37. Copyright 2010 by Cambridge University Press.

means that the inability to learn may have precluded participants from overcoming the initial effect of novelty evident on Block 1, resulting in a smaller improvement in motor planning on Block 2 and therefore, paradoxically, smaller NE.

Prediction of Cognitive Change

To determine whether NE and LRN can predict a *change* in the DRS-2 scores across a one year interval, we conducted a hierarchical regression, using (a) the DRS-2 annualized change score as the criterion variable, and (b) DRS-2 baseline score (to account for the initial level of performance and susceptibility to regression to the mean) and age and education at baseline (to control for demographic influences on future cognitive change) as predictors on Step 1. Next, in one analysis, we added NE and LRN as predictors at Steps 3 and 4, respectively, and, in a separate analysis, we reversed

the last two Steps (i.e., LRN and NE at Steps 3 and 4, respectively), to allow examination of unique contributions of NE and LRN as predictors. As can be seen in Table 3, while demographics and the initial level of performance did not account for a significant amount of variance in cognitive change, NE uniquely contributed approximately 9% of variance in future change above and beyond all the covariates, as well as above LRN. In contrast, LRN did *not* contribute to the model, regardless of whether it was entered on Step 3 or 4.

Of interest, the effect of NE increased somewhat when LRN was entered into the model first (Table 3), suggesting a possible moderating effect of LRN on NE. In other words, some variance in the presumed effect of novelty may have been accounted for by the capacity to learn from practice. However, when the statistical significance of this moderation was tested by entering an interaction term (i.e., NE*LRN) at Step 3 of the equation, the results did not
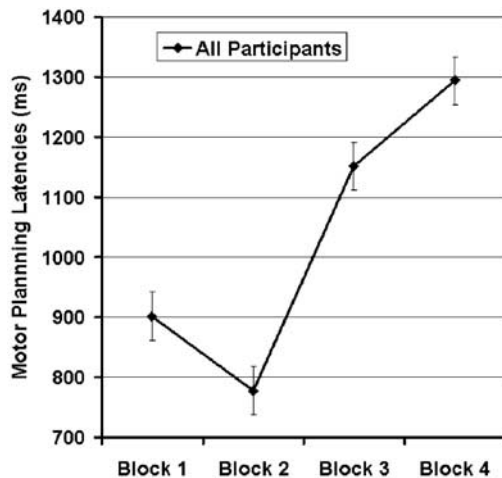
Fig. 3. The figure demonstrates the size of the motor planning latencies on the Push-Turn-Taptap task (Suchy et al., 2005) across four learning blocks. As can be seen, the latencies become progressively longer from Block 2 to Block 4, reflecting increases in sequence length. However, latencies on Block 1 are longer than those on Block 2, despite the fact that the Block 1 sequence is the shortest. The difference between the latencies on the first two blocks has been interpreted as the effect of novelty (NE), and represents the operationalization of NE in the present study. The differences between each pair of adjacent blocks are significant at $p < .05$.

reach significance, F_{change} (1,45) = 1.52, p = .224, suggesting that this moderating effect was relatively negligible. Importantly, as mentioned above, NE accounted for variance in cognitive change, whether LRN was included in the model or not.

Classifying Participants

To determine whether larger NE would flag individuals who are at risk for cognitive decline, we first identified participants whose DRS-2 scores exhibited a reliable annualized decrement from baseline. To accomplish this, we used an empirically derived cut score based on the reliable change index (RCI) methodology (Chelune, Naugle, Lueders, Sedlak, & Awad, 1993). Specifically, it was recently demonstrated that in a sample similar to ours and a comparable test–retest interval[7] that a decline of 6 or more points on the DRS-2 raw scores represented a reliable change at the 5th percentile (Pedraza et al., 2007). This cut score identified 7 of the 50 participants (i.e., 14%) in our sample showing a reliable decrement in performance. The characteristics of the two groups (i.e., decliners and nondecliners) are presented in Table 4. As can be seen, the groups did not differ on any characteristics (all p values >.15), except for the follow-up DRS-2 score, t = 5.19(48), p < .001.

[7] The test–retest interval in our study was approximately 1.4 years (17 months), and the test–retest interval in the Pedraza et al. (2007) study was approximately 1.3 years. We converted the annualized change scores used in the linear regressions to the expected change in 1.3 years, and used those scores to determine reliable decline.

Although the presently observed base rate of decline (i.e., 14%) is considerably higher than would be expected in the general population and is statistically higher than the RCI base rate of 5%, $\chi^2(1)$ = 4.71, p < .030, this relatively high base rate of decline is explained by the study's active recruitment of individuals who had some concerns about their cognition. In fact, subjective complaints about cognition have been shown to be associated with higher base rates of decline at follow-up, despite normal scores at baseline (Geerlings, Jonker, Bouter, Ader, & Schmand, 1999; Schofield, Marder, Dooneief, & Jacobs, 1997).

Next, we conducted a Receiver Operating Characteristic (ROC) curve analysis, using the NE as the predictor and the cognitive retest status (i.e., decliners vs. nondecliners) as criterion. The results showed that NE reliably classified participants (area under the curve [AUC] = .817, p = .008), such that NE cut scores of greater than 103 ms identified correctly 6 of the 7 decliners (i.e., 86% sensitivity), and 27 of 43 nondecliners (i.e., 63% specificity).

Lastly, even though the moderating effect of LRN was negligible in the regression analyses, this effect could nevertheless play a role in participant classification. Thus, we examined the ability of the interaction term (i.e., the product of NE and LRN) to classify participants. The results showed that the interaction term was somewhat more successful than NE alone (AUC = .850, p = .003), such that the NE*LRN product scores greater than 1722 were associated with a sensitivity of 86% (correctly classifying 6 of 7 decliners) and specificity of 81% (correctly classifying 35 of 43 nondecliners). This means that in this prospective cohort study, individuals whose NE*LRN product was greater than 1722 had a relative risk of showing a reliable cognitive decline that was 15 times higher than those whose scores were below that cutoff (i.e., risk ratio = 15.43).

Supplementary Analyses: Dealing with Attrition

Given that the dropouts in this study had lower DRS-2 scores than returning participants already at baseline (Table 1), it is possible that the reason for their failure to return was that they were already on a declining trajectory, with an additional decline in cognition in the intervening period. If that is the case, then these individuals should also exhibit a greater NE. To examine this, we compared the NE for three groups: (1) dropouts (n = 25), (2) nondecliners (i.e., returners who did *not* decline cognitively; n = 43), and (3) decliners (i.e., returners who declined cognitively; n = 7). We conducted two planned comparisons, using NE as the dependent variable, and group membership (i.e., dropouts vs. nondecliners, and decliners vs. nondecliners) as the independent variable. As expected, the results showed that nondecliners had a smaller NE than both those whose cognition declined and those who dropped out (Mann-Whitney U = 55.00, p = .006, Z = 2.67 and Mann-Whitney U = 377.00, p = .041, Z = 2.04, respectively). Similar examination of the LRN variable failed to produce statistically significant results (p values >.394). See Figures 5 and 6.

Table 2. Zero order correlations

Parameter	Age	Education	GDS	DRS-2$_{baseline}$	DRS-2$_{follow-up}$	LRN
NE	−.191	.162	−.067	.034	−.165	−.306*
LRN	.272	−.440**	.073	−.385**	−.296*	−

Note. $N = 50$; DRS-2 = Mattis Dementia Rating Scale-2nd edition (higher values reflect better performance); GDS = Geriatric Depression Scale (higher values reflect greater number of depressive symptoms); NE = Novelty Effect reflecting the difference in motor planning latencies on the first minus the second block of a motor learning task (higher values reflect greater NE); LRN = Learning reflecting the number of errors made on a motor learning task (higher values reflect poorer learning).

Lastly, we examined the NE of the two participants who explicitly stated that they could not return for re-testing due to onset of dementia. Both of these participants had NE above the cutting score of 103 ms (i.e., 229 and 242 ms) and NE*LRN product above the cutting score of 1723 (i.e., 4008 and 4850). Together, these findings lend some support to the suspicion that individuals who dropped out of the study did so in part due to declines in cognition.

DISCUSSION

The present study examined whether susceptibility to the effects of novelty (NE) and the inability to learn from practice (LRN) can be used as markers for future cognitive decline. Consistent with expectations, the results showed that larger NE, as assessed via planning latencies on a motor learning task, was associated with greater cognitive decline (assessed approximately 17 months after the baseline assessment), accounting for variance in cognitive change scores above and beyond demographics, baseline cognitive status, and LRN. Furthermore, NE reliably classified participants into decliners (those showing a reliable decrease on the DRS-2) and non-decliners, with 86% sensitivity and 63% specificity.

In contrast, and contrary to expectation, LRN did *not* explain any variance in future cognitive change. However, LRN was negatively correlated with NE, and appeared to moderate the relationship between NE and future cognitive

Fig. 4. The observed relationship between Novelty Effect (NE) and Learning (LRN) indices. As can be seen, larger NE was associated with fewer errors.

decline. Once this moderation effect was accounted for, the ability to classify participants improved somewhat (i.e., specificity improved from 63% to 81%, with sensitivity of 86%). Consequently, the hypothesis that NE and LRN are independent was only partially supported; NE independently predicted future cognitive decline, but LRN appeared to slightly moderate this relationship.

Of interest, although NE proved to be a reliable predictor of future cognitive decline, it was *un*related to the *level* of current or future cognitive functioning. In contrast, LRN correlated with the level of cognition at both time points, but did not predict future decline. These results highlight the fundamental difference between the two processes, such that LRN assesses cognitive abilities that are traditionally captured in the scope of neuropsychological assessments, whereas NE may tap into less-measured constructs (such as current level of cognitive reserve, neural efficiency, or the capacity to compensate for fluctuations in task demands).

Taken together, these findings offer promise for the utility of NE as a marker of preclinical neurocognitive changes that eventually convert into clinically notable cognitive declines. These findings may also offer a potential explanation for the apparent inconsistencies in the literature regarding the utility of small or absent PE as a marker of future cognitive decline, and suggest directions for future research.

Theoretical Implication: Deconstructing the Practice Effect

When an improvement in performance takes place due to repeated exposure to a task, such improvement is referred to as PE. However, any time PE is present, it could technically be conceptualized as either a result of LRN due to repeated exposure, or a result of NE interfering with performance upon the initial exposure, followed by a rebound. However, LRN and NE are not just two sides of the same coin. In fact, based on the present findings, LRN and NE appear to represent two largely separate constructs, which together may contribute in a dynamic way to improvements in test scores known as PE.

In particular, one could speculate that in healthy individuals, second exposure to a task may be associated with an improvement that is to a small extent a function of NE, and to a large extent a function of LRN. That some NE is present
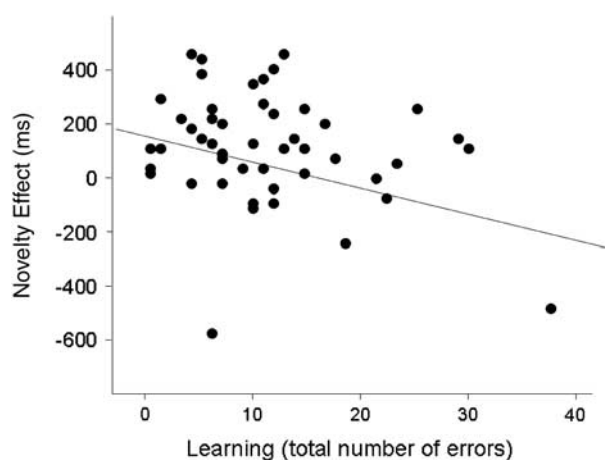
Table 3. Hierarchical regression results showing the relative contribution of learning and novelty to prediction of annualized cognitive decline

Analysis	Step	Variables	R^2_{change}	F_{change}	df1	df2	p
1 and 2	1[a]	DRS-2 raw, age, education	.074	1.23	3	46	.310
1	2	NE	.079	4.19	1	45	.046
1	3	LRN	.008	.404	1	44	.528
2	2	LRN	.000	.004	1	45	.953
2	3	NE	.087	4.57	1	44	.039

Note. DRS-2 raw = Mattis Dementia Rating Scale-2nd edition total raw score; NE = Novelty Effect reflecting the difference in motor planning latencies on the first minus the second block of a motor learning task; LRN = Learning reflecting the number of errors across four motor learning task blocks.
[a] Step 1 is identical for both the analysis 1 and the analysis 2.

among cognitively healthy individuals is known from both functional imaging research (Brovelli et al., 2008), as well as from behavioral research conducted with the present task (Suchy & Kraybill, 2007).

As a person's cognitive reserve declines, the normal effect of novelty may exert increasingly greater demands on cognitive processing, evidenced not only by the increasingly wide-spread fMRI activation observed in such individuals (Dickerson et al., 2005; Lenzi et al., 2009), but also by measurably slower planning latencies early on during task performance (observed in the present study). Furthermore, because LRN in the present study did not predict future decline, the present results suggest that as NE exerts greater influence, learning may still be relatively intact, potentially leading to greater gains in scores upon second exposure to a given test. This, in fact, was found in at least one study (Duff et al., 2008).

As cognition continues to decline, improvement on test re-administration may begin to decline as well (Cooper et al.,

2004; Darby et al., 2002), due to the diminished ability to benefit from practice. It is also possible that at some point along the cognitive decline trajectory, NE may no longer play a role, as any given test may present as equally overwhelming regardless of how many times it is administered. For an illustration of these theoretical relationships, see Figure 7. Future research may focus on the separate contributions of NE and LRN to PE, as well as the potentially curvilinear relationship between PE and cognitive aging.

In addition to considering the effect of novelty, future research also needs to consider that PE likely consists of both implicit (procedural) learning and explicit (declarative) learning. The degree to which these two types of learning take place during re-administration of any given task is not understood, but is likely not uniform across different measures. These two types of learning processes are also known to be differentially affected for different clinical populations, such that, for example, the declarative aspect tends to be more affected in early Alzheimer's disease, whereas the procedural

Table 4. Demographic, depressive, and cognitive characteristics of the sample divided by cognitive decline status

	Nondecliners ($n = 43$)	Decliners ($n = 7$)
Follow-up time (months)		
M (SD)	17.40 (6.60)	14.40 (9.96)
Range	7.68–32.04	8.16–19.92
Age (years)		
M (SD)	69.63 (6.00)	68.43 (9.11)
Range	60–87	58–85
Education (years)		
M (SD)	14.88 (2.24)	13.57 (2.07)
Range	11–18	10–16
GDS score (raw score)		
M (SD)	4.53 (4.42)	4.43 (2.23)
Range	0–23	1–7
DRS-2 baseline (raw score)		
M (SD)	139.44 (3.20)	139.43 (4.28)
Range	130–144	133–144
DRS-2 follow-up (raw score)		
M (SD)	139.33** (3.80)	130.71** (5.62)
Range	127–144	123–136
% Female	65%	57%

Note. M = mean, SD = standard deviation; DRS-2 = Mattis Dementia Rating Scale-2nd edition; GDS = Geriatric Depression Scale.
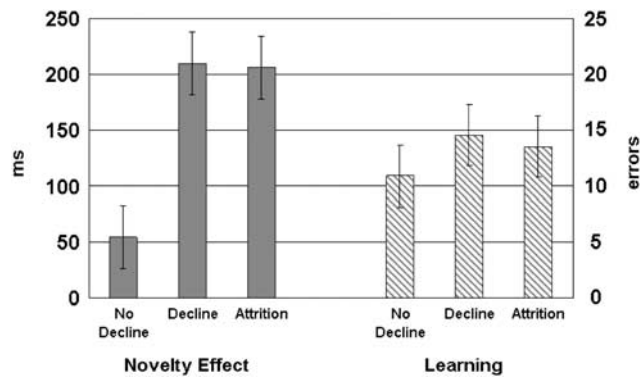*Values differ at $p < .05$. **Values differ at $p < .001$.

Fig. 5. The figure demonstrates the size of the Novelty Effect (NE; reflecting the difference in motor planning latencies between the first and the second blocks of a motor learning task) and the difficulty in Learning (LRN; reflecting the number of errors across four blocks of the motor learning task) for (a) participants who returned for follow-up testing and remained cognitively the same ("nondecliners"; $n = 43$), (b) participants who returned for follow-up testing and exhibited a reliable cognitive decline ("decliners"; $n = 7$), and (c) participants who dropped out of the study after the baseline assessment ("dropouts"; $n = 25$).

Fig. 7. The figure demonstrates a hypothetical model of dynamical changes in the size of the practice effect (PE), with variable contributions from novelty effect (NE) and learning (LRN) as a function of cognitive decline.

CLINICAL IMPLICATIONS

Clinically, the challenge has been not only to identify individuals who are on a trajectory of pathological neurocognitive decline (typically represented by the MCI diagnostic category), but also to identify individuals whose cognitive decline is still masked by sufficient cognitive reserve, that is, before the emergence of clinical symptoms. The interest in the latter group will only continue to grow as new pharmacologic interventions for neurodegenerative diseases become available. The present paradigm (i.e., the assessment of NE using the PTT task) offers promise for such clinical applications.

First, NE reflects a very discrete aspect of performance, one that is not typically quantified as part of clinical assessment, as it is generally subsumed into a more global score (i.e., the brief latency before initiation of movement represents a negligible fraction of the overall speed with which a task is completed). As such, it gets at an aspect of performance for which patients likely cannot consciously compensate, but which nevertheless emerges as a predictor of future decline, possibly reflecting a depletion of cognitive reserve. Given that such subtle NE is not detected by traditional assessment methods offers an explanation for why decreases in neural efficiency evidenced by functional imaging do not necessarily translate into a measurable cognitive compromise (Ernst, Yakupov, Nakama, Crocket, Cole, Watters et al., 2009; Rypma, Berger, Genova, Rebbechi, & D'Esposito, 2005).

Second, although NE likely can be assessed in a variety of ways, using motor planning latencies may represent a particularly sensitive method. Motor planning refers to an abstract plan that contains both general information about the intended goal and specific information about the neuromuscular control that will be required (Keele, 1968). It has been shown that longer movement sequences require more planning time (Keele, 1981; Klapp, McRae, & Long, 1978; Suchy & Kraybill, 2007; Wright et al., 2004), and that task novelty

aspect tends to be more affected in early Parkinson's disease. Thus, for the LRN component of PE to be a viable marker for future decline, it needs to be examined more cautiously, with careful analysis of PE and its components, as they relate to different measures and different clinical populations.
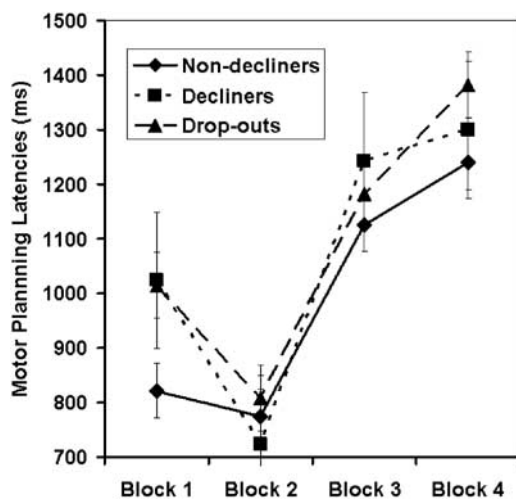

Fig. 6. The figure demonstrates the size of the median motor planning latencies on the Push-Turn-Taptap task (Suchy & Kraybill, 2007) across four learning blocks for (a) participants who returned for follow-up testing and remained cognitively the same ("nondecliners"; $n = 43$), (b) participants who returned for follow-up testing and exhibited a reliable cognitive decline ("decliners"; $n = 7$), and (c) participants who dropped out of the study after the baseline assessment ("dropouts"; $n = 25$). As can be seen, both the dropouts and the decliners exhibited longer planning latencies than nondecliners on Block 1 (evident by nonoverlapping standard errors), despite having comparable latencies on the remaining blocks (evident by overlapping standard errors).
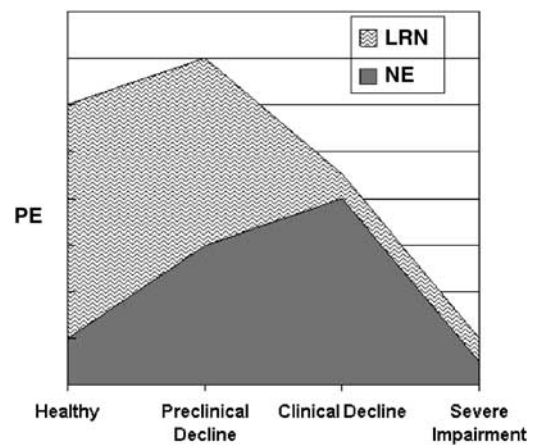
increases motor planning latencies, likely due to the added "complexity" of processing task instructions and converting those into a novel action plan (Suchy et al., 2005; Suchy & Kraybill, 2007). Interestingly, language (i.e., instructions) comprehension has been shown to require extra neural processing among cognitively compromised individuals (Bosch et al., 2010), and transfer of learned information into motor action has been shown to be impaired among individuals with mild preclinical hippocampal atrophy (Gluck, Myers, Nicolle, & Johnson, 2006). Additionally, recent evidence suggests that sensorimotor connectivity is altered among MCI individuals (Agosta et al., 2010). Together, these three mechanisms suggest that motor planning latencies may be particularly sensitive to NE among individuals who are on a trajectory of neurocognitive decline.

Although NE, as assessed in the present study, proved useful in predicting future decline, the effectiveness of this variable needs to be pitted against other potential effects of novelty already embedded in existing clinical measures. For example, TBI patients exhibit steep learning curves on verbal learning tasks, marked by very poor recall upon the first exposure to a given word list, followed by a successful rebound on subsequent exposures (Geary, Kraus, Pliskin, & Little, 2010). It is unclear whether this effect represents problems with encoding (as it is often interpreted), or an effect of novelty. Regardless, the utility of such a profile in predicting cognitive declines should be examined in future research.

LIMITATIONS

The principal limitation of the present study is the approximately 33% attrition rate, resulting in a relatively modest sample size of 50 participants at the time of follow up testing. However, this rate of attrition is common in prospective cohort studies of older adults (Niti et al., 2009; Stewart et al., 2009). The attrition rate may have been caused by cognitive decline among some of the participants who did not return. This interpretation is consistent with the dropouts' lower DRS-2 scores already at baseline. Additionally, dropouts as a group showed NEs that were (a) above the empirically derived cutting score and (b) comparable to those seen in participants who exhibited cognitive declines at follow-up.

Additionally, it is not clear to what extent the present findings would generalize to other measures of novelty, or whether LRN, as assessed in this study, is comparable to other measures of procedural and declarative learning. Replications with the presently employed instrument used in conjunction with other measures of NE and LRN would help clarify this question.

Lastly, it is not clear whether the present results would generalize to other populations. For example, given the relatively low specificity (61% for NE and 81% for NE*LRN), it is possible that false positive rates would be unacceptably high in a sample of independent adults who do *not* express concerns about their cognition. A lack of such a control group in the present study represents a significant

weakness. Alternatively, the specificity rates may improve in a clinical sample, such as individuals who seek medical attention for cognitive problems or individuals diagnosed with MCI. By the same token, different MCI subtypes may also respond differently to our task, and may exhibit differential rates of decline. These questions need to be examined in future research.

REFERENCES

Agosta, F., Rocca, M.A., Pagani, E., Absinta, M., Magnani, G., Marcone, A., … Filippi, M. (2010). Sensorimotor network rewiring in mild cognitive impairment and Alzheimer's disease. *Human Brain Mapping*, 31(4), 515–525.

Andel, R., Vigen, C., Mack, W.J., Clark, L.J., & Gatz, M. (2006). The effect of education and occupational complexity on rate of cognitive decline in Alzheimer's patients. *Journal of the International Neuropsychological Society*, 12(1), 147–152.

Attix, D.K., Story, T.J., Chelune, G.J., Ball, J.D., Stutts, M.L., Hart, R.P., & Barth, J.T. (2009). The prediction of change: Normative neuropsychological trajectories. *The Clinical Neuropsychologist*, 23(1), 21–38.

Bosch, B., Bartres-Faz, D., Rami, L., Arenaza-Urquijo, E.M., Fernandez-Espejo, D., Junque, C., … Molinuevo, J.L. (2010). Cognitive reserve modulates task-induced activations and deactivations in healthy elders, amnestic mild cognitive impairment and mild Alzheimer's disease. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*, 46(4), 451–461.

Brovelli, A., Laksiri, N., Nazarian, B., Meunier, M., & Boussaoud, D. (2008). Understanding the neural computations of arbitrary visuomotor learning through FMRI and associative learning theory. *Cerebral Cortex*, 18(7), 1485–1495.

Busch, R.M., Chelune, G.J., & Suchy, Y. (Eds.). (2005). *Using norms in neuropsychological assessment of the elderly*. New York: Guilford Press.

Chelune, G.J., Naugle, R.I., Lueders, H., Sedlak, J., & Awad, I.A. (1993). Individual change after epilepsy surgery: Practice effects and base-rate information. *Neuropsychology*, 7(1), 41–52.

Cooper, D.B., Lacritz, L.H., Weiner, M.F., Rosenberg, R.N., & Cullum, C.M. (2004). Category fluency in mild cognitive impairment: Reduced effect of practice in test-retest conditions. *Alzheimer Disease and Associated Disorders*, 18(3), 120–122.

Corral, M., Rodriguez, M., Amenedo, E., Sanchez, J.L., & Diaz, F. (2006). Cognitive reserve, age, and neuropsychological performance in healthy participants. *Developmental Neuropsychology*, 29(3), 479–491.

Darby, D., Maruff, P., Collie, A., & McStephen, M. (2002). Mild cognitive impairment can be detected by multiple assessments in a single day. *Neurology*, 59(7), 1042–1046.

Dickerson, B.C., Salat, D.H., Greve, D.N., Chua, E.F., Rand-Giovannetti, E., Rentz, D.M., … Sperling, R.A. (2005). Increased hippocampal activation in mild cognitive impairment compared to normal aging and AD. *Neurology*, 65(3), 404–411.

Duff, K., Beglinger, L.J., Moser, D.J., Paulsen, J.S., Schultz, S.K., & Arndt, S. (in press). Predicting cognitive change in older adults: the relative contribution of practice effects. *Archives of Clinical Neuropsychology*.

Duff, K., Beglinger, L.J., Schultz, S.K., Moser, D.J., McCaffrey, R.J., Haase, R.F., … Paulsen, J.S. (2007). Practice effects in the prediction of long-term cognitive outcome in three patient

samples: A novel prognostic index. *Archives of Clinical Neuropsychology*, 22(1), 15–24.

Duff, K., Beglinger, L.J., Van Der Heiden, S., Moser, D.J., Arndt, S., Schultz, S.K., & Paulsen, J.S. (2008). Short-term practice effects in amnestic mild cognitive impairment: Implications for diagnosis and treatment. *International Psychogeriatrics*, 20(5), 986–999.

Ernst, T., Yakupov, R., Nakama, H., Crocket, G., Cole, M., Watters, M., Ricardo-Dukelow, M.L., & Chang, L. (2009). Declined neural efficiency in cognitively stable human immunodeficiency virus patients. *Annals of Neurology*, 65(3), 316–325.

Falleti, M.G., Maruff, P., Collie, A., & Darby, D.G. (2006). Practice effects associated with the repeated assessment of cognitive function using the CogState battery at 10-minute, one week and one month test-retest intervals. *Journal of Clinical and Experimental Neuropsychology*, 28(7), 1095–1112.

Garrett, D.D., Grady, C.L., & Hasher, L. (2010). Everyday memory compensation: The impact of cognitive reserve, subjective memory, and stress. *Psychology and Aging*, 25(1), 74–83.

Geary, E.K., Kraus, M.F., Pliskin, N.H., & Little, D.M. (2010). Verbal learning differences in chronic mild traumatic brain injury. *Journal of the International Neuropsychological Society*, 16(3), 506–516.

Geerlings, M.I., Jonker, C., Bouter, L.M., Ader, H.J., & Schmand, B. (1999). Association between memory complaints and incident Alzheimer's disease in elderly people with normal baseline cognition. *The American Journal of Psychiatry*, 156(4), 531–537.

Gluck, M.A., Myers, C.E., Nicolle, M.M., & Johnson, S. (2006). Computational models of the hippocampal region: Implications for prediction of risk for Alzheimer's disease in non-demented elderly. *Current Alzheimer Research*, 3(3), 247–257.

Jansma, J.M., Ramsey, N.F., Slagter, H.A., & Kahn, R.S. (2001). Functional anatomical correlates of controlled and automatic processing. *Journal of Cognitive Neuroscience*, 13(6), 730–743.

Keele, S. (1968). Movement control in skilled motor performance. *Psychological Bulletin* (70), 387–403.

Keele, S. (1981). Behavioral analysis of movement. In V.B. Brooks (Ed.), *Handbook of physiology, section 1, volume 2. Motor control* pp. (1391–1414). Bethesda, MD: American Psychological Society.

Klapp, S.T., McRae, J., & Long, W. (1978). Response programming vs alternative interpretations of the 'dit-dah' reaction time effect. *Bulletin of the Psychonomic Society*, 11(1), 5–6.

Kraybill, M.L., & Suchy, Y. (2008). Evaluating the role of motor regulation in figural fluency: Partialing variance in the Ruff Figural Fluency Test. *Journal of Clinical and Experimental Neuropsychology*, 30(8), 903–912.

Lenzi, D., Serra, L., Perri, R., Pantano, P., Lenzi, G.L., Paulesu, E., … Macaluso, E. (2009). Single domain amnestic MCI: A multiple cognitive domains fMRI investigation. *Neurobiology of Aging*. doi:10.1016/j.neurobiolaging.2009.09.006

Maassen, G.H., Bossema, E., & Brand, N. (2008). Reliable change and practice effects: Outcomes of various indices compared. *Journal of Clinical and Experimental Neuropsychology*, 31(3), 339–352.

Mattis, S.L., Jurica, P.J., & Leitten, C.L. (1988). *The Dementia Rating Scale*. Lutz, FL: Psychological Assessment Resources.

Niti, M., Yap, K.-B., Kua, E.-H., & Ng, T.-P. (2009). APOE-É›4, depressive symptoms, and cognitive decline in Chinese older adults: Singapore longitudinal aging studies. *The Journals of Gerontology: Series A: Biological Sciences and Medical Sciences*, 64A(2), 306–311.

Pedraza, O., Smith, G.E., Ivnik, R.J., Willis, F.B., Ferman, T.J., Petersen, R.C., … Lucas, J.A. (2007). Reliable change on the Dementia Rating Scale. *Journal of the International Neuropsychological Society*, 13(4), 716–720.

Rapport, L.J., Brines, D.B., Axelrod, B.N., & Theisen, M.E. (1997). Full scale IQ as mediator of practice effects: The rich get richer. *Clinical Neuropsychologist*, 11(4), 375–380.

Roennlund, M., Loevden, M., & Nilsson, L. (2008). Cross-sectional versus longitudinal age gradients of Tower of Hanoi performance: The role of practice effects and cohort differences in education. *Aging, Neuropsychology, and Cognition*, 15(1), 1–28.

Rypma, B., Berger, J.S., Genova, H.M., Rebbechi, D., & D'Esposito, M. (2005). Dissociating age-related changes in cognitive strategy and neural efficiency using event-related fMRI. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*, 41(4), 582–594.

Schofield, P.W., Marder, K., Dooneief, G., & Jacobs, D.M. (1997). Association of subjective memory complaints with subsequent cognitive decline in community-dwelling elderly individuals with baseline cognitive impairment. *The American Journal of Psychiatry*, 154(5), 609–615.

Stern, Y. (2002). What is cognitive reserve? Theory and research application of the reserve concept. *Journal of the International Neuropsychological Society*, 8(3), 448–460.

Stewart, J.C., Rand, K.L., Muldoon, M.F., & Kamarck, T.W. (2009). A prospective evaluation of the directionality of the depression-inflammation relationship. *Brain, Behavior, and Immunity*, 23(7), 936–944.

Suchy, Y., Derbidge, C., & Cope, C. (2005). Behavioral dyscontrol scale-electronic version: First examination of reliability, validity, and incremental utility. *Clinical Neuropsychologist*, 19(1), 4–26.

Suchy, Y., & Kraybill, M.L. (2007). The relationship between motor programming and executive abilities: Constructs measured by the Push-Turn-Taptap task from the BDS-EV. *Journal of Clinical and Experimental Neuropsychology*, 29(6), 648–659.

Suchy, Y., Kraybill, M., & Larson, J.G.L. (2010). Understanding design fluency: Motor and executive contributions. *Journal of the International Neuropsychological Society*, 16(1), 26–37.

Thompson, D.N. (1997). Practice effects of advance organization with older adult subjects. *Educational Gerontology*, 23(3), 207–212.

Wright, D.L., Black, C.B., Immink, M.A., Brueckner, S., & Magnuson, C. (2004). Long-term motor programming improvements occur via concatenation of movement sequences during random but not during blocked practice. *Journal of Motor Behavior*, 36(1), 39–50.

Yesavage, J.A. (1988). Geriatric depression scale. *Psychopharmacology Bulletin*, 24(4), 709–711.