



Published in final edited form as:

*Neuropsychologia*. 2017 July 28; 102: 95–108. doi:10.1016/j.neuropsychologia.2017.05.027.

## Cognitive persistence: Development and validation of a novel measure from the Wisconsin Card Sorting Test

Susan Teubner-Rhodes<sup>\*</sup>, Kenneth I. Vaden Jr., Judy R. Dubno, and Mark A. Eckert<sup>\*\*</sup>

Dept. of Otolaryngology-Head & Neck Surgery, Medical University of South Carolina, 135 Rutledge Ave MSC550, Charleston, SC 29425

### Abstract

The Wisconsin Card Sorting Test (WCST) has long been used as a neuropsychological assessment of executive function abilities, in particular, cognitive flexibility or “set-shifting.” Recent advances in scoring the task have helped to isolate specific WCST performance metrics that index set-shifting abilities and have improved our understanding of how prefrontal and parietal cortex contribute to set-shifting. We present evidence that the ability to overcome task difficulty to achieve a goal, or “cognitive persistence,” is another important prefrontal function that is characterized by the WCST and that can be differentiated from efficient set-shifting. This novel measure of cognitive persistence was developed using the WCST-64 in an adult lifespan sample of 230 participants. The measure was validated using individual variation in cingulo-opercular cortex function in a sub-sample of older adults who had completed a challenging speech recognition in noise fMRI task. Specifically, older adults with higher cognitive persistence were more likely to demonstrate word recognition benefit from cingulo-opercular activity. The WCST-derived cognitive persistence measure can be used to disentangle neural processes involved in set-shifting from those involved in persistence.

### Keywords

Persistence; Set-shifting; Wisconsin Card Sorting Test; Speech recognition; Prefrontal cortex

## 1. Introduction

Psychologists have long recognized that achievement on goal-directed tasks emerges not only as a result of cognitive or intellectual ability, but also from the motivation, drive, or will to succeed (Wechsler, 1950). Thus, “persistence”—applying effort to overcome a mental challenge—is thought to be an essential component underlying performance on cognitive tasks. However, the contribution of persistence to inter-individual variability in performance on mentally-demanding tasks is often neglected. This may be due, in part, to the paucity of

<sup>\*</sup>Corresponding author. 843-792-3921. <sup>\*\*</sup>Corresponding author. 843-792-5914.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

neuropsychological assessments that disentangle the effects of persistence and cognitive ability on performance.

Existing measures of persistence and related motivational factors typically take the form of subjective self-report or observer-report surveys that may be biased by prior knowledge of achievement (Choi, Mogami, & Medalia, 2010; Cloninger, Przybeck, Svrakic, & Wetzel, 1994; Doherty-Bigara & Gilmore, 2016; Onatsu-Arvilommi & Nurmi, 2000; Pintrich, Smith, Garcia, & McKeachie, 1991, 1993; Steinberg et al., 2007) and are often domain-specific (e.g., academic achievement; Onatsu-Arvilommi & Nurmi, 2000; Pintrich et al., 1991, 1993; Zhang, Nurmi, Kiuru, Lerkkanen, & Aunola, 2011). Of the few behavioral measures of persistence, most examine time spent on challenging tasks before deciding to quit (for review, see Leyro, Zvolensky, & Bernstein, 2010). While persistent individuals may be likely to engage effort in task performance for longer periods, the reverse does not follow, as higher ability levels may also increase how long individuals choose to work on a task. The goal of the current study was to establish a behavioral measure of persistence that was independent of task ability and reflected the application of effort to overcome performance difficulty, using the Wisconsin Card Sorting Task (WCST).

The WCST is a commonly used neuropsychological assessment that was developed to characterize frontal lobe function (Drewe, 1974; Milner, 1963; Nelson, 1976). The standard version of the task (Grant & Berg, 1948; Heaton, 1981; Kongs, Thompson, Iverson, & Heaton, 2000; Milner, 1963) involves matching a target card to one of four sample cards that vary in color, shape, and number, without knowing a priori how to match the cards. Participants learn the sorting rule (color, shape, or number) through trial-and-error from feedback on each trial, and the rule is changed after 10 consecutive correct responses.

The traditional index of frontal lobe function on the WCST is perseverative errors, that is, the number of errors made because participants sorted a card based on the previously reinforced rule instead of the current rule. The prefrontal cortex is thought to be essential for flexible rule switching, or “set-shifting”, because patients with lateral and/or dorsomedial prefrontal lesions exhibit more perseverative errors on the WCST than healthy controls or patients with non-frontal damage (Barceló & Knight, 2002; Drewe, 1974; Milner, 1963; Nelson, 1976; Stuss et al., 2000). Additionally, rule switches during the WCST and similar tasks elicit activity in lateral and dorsomedial prefrontal cortex (Hampshire, Gruszka, Fallon, & Owen, 2008; Konishi et al., 2002; Konishi et al., 1998; Nagahama et al., 1998; Ravizza & Carter, 2008), suggesting that these regions support set-shifting. However, non-prefrontal regions including posterior parietal cortex, occipital cortex and the striatum have also been implicated in set-shifting (Dang, Donde, Madison, O’Neil, & Jagust, 2012; Graham et al., 2009; Konishi et al., 2002; Konishi et al., 1998; Nagahama et al., 1998; Ravizza & Carter, 2008; Wang, Cao, Cai, Gao, & Li, 2015). Moreover, patients with prefrontal lesions can exhibit deficits in non-perseverative errors in addition to perseverative errors (Barceló & Knight, 2002; Drewe, 1974). These findings call into question the specificity of prefrontal cortex function in set-shifting, and appear to reflect the multi-faceted nature of cognitive processes that support WCST performance. The WCST requires not only set-shifting to flexibly switch rules, but also problem solving to deduce the correct sorting rule and working memory to maintain and retrieve task goals.

Due to its varied behavioral demands, neuroimaging and patient studies find that the WCST actually engages widespread prefrontal, parietal and occipital regions (Berman et al., 1995; Konishi et al., 2002; Konishi et al., 1998; Nagahama et al., 1998; Nyhus & Barcelo, 2009). Recent research has attempted to specify the contributions of distinct cortical areas to WCST performance by breaking down the task into its constituent components or manipulating task demands (Barceló, 1999; Barceló & Knight, 2002; Dang et al., 2012; Graham et al., 2009; Lange et al., 2016; Nyhus & Barcelo, 2009; Ravizza & Carter, 2008; Stuss et al., 2000; Wang et al., 2015). In particular, Barceló and colleagues (1999, 2003; Barceló & Knight, 2002) introduced the concept of an “efficient error,” which occurs when a participant happens to switch to the wrong sorting rule after receiving feedback that the rule has changed. A participant who is performing optimally is expected to commit efficient errors on 50% of trials following the detection of a rule change, because there are two remaining rules that could possibly be correct (see Figure 1). Importantly, when efficient errors were coded separately, patients with prefrontal lesions made fewer efficient errors, more perseverative errors, and more non-perseverative errors compared to controls (Barceló & Knight, 2002). These results suggest a dissociation between efficient errors and other error types, wherein prefrontal damage increases perseverative and non-perseverative errors while selectively decreasing efficient errors. Indeed, a factor analysis of error types on the WCST confirmed this distinction: efficient errors were negatively correlated with all other error types, whereas perseverative and non-perseverative errors were strongly positively correlated and did not load onto separate factors (Godinez, Friedman, Rhee, Miyake, & Hewitt, 2012). Because efficient errors reflect optimal shifting processes whereas other errors indicate suboptimal shifting, scoring efficient and non-perseverative errors together likely obscured the effects of frontal damage in prior studies (Drewe, 1974; Milner, 1963; Nelson, 1976; Stuss et al., 2000).

Clearly defining efficient errors and recognizing that they index normal and adaptive shifting processes has enabled targeted investigation of the neural correlates of set-shifting. Research using event-related potentials has shown that efficient errors elicit larger parietal-occipital N1 and frontal P2 amplitudes than perseverative errors (Barceló, 1999), suggesting that set-shifting involves visual attention and frontal control. Relative to rule maintenance trials, switching rules evokes a robust frontal P3a component (Barceló, 2003) that is modulated by uncertainty of decision-response outcomes (Kopp & Lange, 2013). Efficient shifting in healthy young adults engages left precuneus, left inferior frontal gyrus (IFG), right dorsal anterior cingulate cortex (dACC) and right middle frontal gyrus (MFG) compared to correctly repeating a rule (Lao-Kaim et al., 2015). In addition, parametrically increasing rule search demands elicits greater activity in the bilateral IFG and MFG, bilateral inferior parietal lobe, right angular gyrus, superior parietal lobule, precuneus and putamen (Wang et al., 2015). Thus, successful set-shifting recruits a composite network of lateral and medial prefrontal cortex and posterior parietal cortex.

While our understanding of the specific neural networks involved in successful set-shifting on the WCST has advanced, other cognitive processes important to WCST performance have received less attention. In particular, WCST performance depends not only on the ability to flexibly shift between sorting rules, but also on the continued willingness to search for and apply the correct rule after encountering negative feedback that inevitably occurs

following a rule switch. In other words, individuals have to apply effort to overcome the performance difficulty that arises from switching rules, or use cognitive persistence, in order to respond correctly. The present study leverages the Barceló and Knight (2002) approach to isolate a cognitive persistence component of WCST performance and relate it to a prefrontal neural signature of persistence. Specifically, cognitive persistence was related to activity in a cingulo-opercular network that responds to performance difficulty and subsequently leads to better performance during a challenging speech recognition in noise task (Eckert, Teubner-Rhodes, & Vaden, 2016; Vaden, Kuchinsky, Ahlstrom, Dubno, & Eckert, 2015; Vaden et al., 2016; Vaden et al., 2013).

Although the role of cognitive persistence in WCST performance has not been explicitly studied, there is some evidence that it may be important. Reports from studies of patients with prefrontal lesions suggest that receiving frequent negative feedback during WCST administration caused some patients to become so frustrated that they refused to complete the task (Drewe, 1974; Nelson, 1976). The decision to quit a challenging task has previously been used as an inverse measure of persistence (Daughters, Lejuez, Bornoalova, et al., 2005; Daughters, Lejuez, Kahler, Strong, & Brown, 2005; Daughters, Lejuez, Strong, et al., 2005; Leyro et al., 2010; Quinn, Brandon, & Copeland, 1996; Steinberg, Williams, Gandhi, Foulds, & Brandon, 2010; Steinberg et al., 2012; Ventura, Shute, & Zhao, 2013). Choosing to terminate the WCST represents an extreme and relatively rare consequence of low persistence, but persistence may still affect performance among those who complete the task. Indeed, ratings of task effort on an intrinsic motivation survey (Choi et al., 2010) predicted overall WCST performance in patients with Schizophrenia (Tas, Brown, Esen-Danaci, Lysaker, & Brune, 2012). Additionally, depressive symptoms, which often occur with decreased motivation and persistence (Potter et al., 2007; Ravizza & Delgado, 2014), predicted the total number of WCST errors made by adolescents (Han et al., 2016).

The importance of cognitive persistence becomes apparent when considering two hypothetical individuals who have equal difficulty with set-shifting, but differ in persistence. By definition, task difficulty is identical for these individuals, due to their equivalent ability. However, the more persistent person works harder to systematically search for and apply a new rule after experiencing an error. In contrast, the less persistent person is not willing to work as hard to identify the new rule and therefore tends to respond haphazardly and make more errors. The less persistent person may still discover the new rule by chance, but will have lower accuracy compared to the person with good persistence who applied effort to switch rules.

The above example illustrates that performance on a given task is determined by both the ability to perform that task and the effort applied to it. Differences in performance that are not explained by differences in the ability to perform a task reflect, in part, differences in persistence to overcome mental challenges. In the case of the WCST, set-shifting ability is critical to successful performance and (as shown below) explains the majority of variance in sorting accuracy. Nevertheless, some individuals perform better than predicted for their set-shifting ability while others underperform. We propose that differences in sorting accuracy from what is expected by set-shifting ability can be used to index the application of effort to overcome difficulty, i.e., cognitive persistence. Controlling for set-shifting ability equates

task difficulty across individuals. Thus, remaining accuracy differences reflect the extent to which individuals continue to strive toward the task goal of discovering and using the correct sorting rule, despite encountering performance difficulty. We thus reasoned that we could obtain a measure of cognitive persistence from the WCST by controlling for set-shifting ability, measured by the number of efficient shifts, from sorting accuracy scores.

The aforementioned brain regions engaged during the WCST include cingulo-opercular regions that appear to serve as the neural instantiation of cognitive persistence. The cingulo-opercular network, consisting of the right and left IFG, anterior insula/frontal operculum (AIFO), and dACC, is upregulated when the task at hand becomes more difficult and/or performance drops (Botvinick, Braver, Barch, Carter, & Cohen, 2001; Botvinick, Cohen, & Carter, 2004; Botvinick, Nystrom, Fissell, Carter, & Cohen, 1999; Carter, 1998; Crone, Wendelken, Donohue, & Bunge, 2006; Durston et al., 2003; Eichele et al., 2008; Luks, Simpson, Dale, & Hough, 2007; Ridderinkhof, Ullsperger, Crone, & Nieuwenhuis, 2004; Weissman, Roberts, Visscher, & Woldorff, 2006; Weissman, Warner, & Woldorff, 2009). Furthermore, cingulo-opercular engagement is associated with improved performance on the next trial (Botvinick et al., 2001; Botvinick et al., 2004; Kerns, 2006; Kerns et al., 2004; Orr & Weissman, 2009; Ridderinkhof et al., 2004; Weissman et al., 2006). Cingulo-opercular activation during challenging conditions has been observed across a variety of tasks, including understanding speech in multi-talker babble. A recent meta-analysis demonstrated that cingulo-opercular regions respond to decreased speech intelligibility when increased listening effort is needed (Eckert et al., 2016). Moreover, cingulo-opercular activity was elevated prior to successful word recognition in challenging listening conditions (Vaden et al., 2015; Vaden et al., 2013). One interpretation of these observations is that cingulo-opercular activity reflects not only a performance monitoring response, but also the application of effort to optimize performance (Eckert et al., 2016). In other words, cingulo-opercular function may reflect, at least in part, cognitive persistence.

In the present study, we used the WCST to develop a measure of cognitive persistence in a lifespan sample of typical adults. Persistence was defined as the extent to which participants performed better or worse than expected from their ability to flexibly shift rules. We then assessed the relationship between individual differences in persistence as estimated from the WCST and two measures of cingulo-opercular activity that are thought to reflect persistence during a challenging speech recognition in noise task. Specifically, we examined the extent to which persistence predicted 1) cingulo-opercular activity in response to word recognition errors and 2) cingulo-opercular activity that was associated with better subsequent performance in a group of middle-aged to older adults. We focused on this age group for the neuroimaging validation of persistence because it is well established that speech recognition in background noise is more effortful and challenging for older adults than younger adults (Anderson Gosselin & Gagné, 2011; Edwards, 2007; Kuchinsky et al., 2013; Zekveld, Kramer, & Festen, 2011). Thus, although recognizing speech in noise may require persistence for listeners of all ages, this is especially true later in life. We hypothesized that middle-aged to older adults with higher persistence would exhibit greater cingulo-opercular activity in response to errors and a stronger relationship between cingulo-opercular activity and subsequent word recognition, thereby providing validation for the persistence measure.

## 2. Method

### 2.1. Participants

The sample included 246 adults across the lifespan (age range: 19.41 to 88.17 years) whose results were studied to develop a measure of cognitive persistence using the abbreviated WCST-64 (Kongs et al., 2000). Thirty-one of the participants had completed a speech recognition task during fMRI scanning, as described in Vaden et al. (2015). The remaining 215 participants did not undergo neuroimaging. Participants were recruited as part of a longitudinal study on age-related hearing loss conducted at the Medical University of South Carolina. Evidence of conductive hearing loss or otologic/neurologic disease served as exclusionary criteria.

Two participants were excluded from all analyses because the WCST-64 was terminated by the experimenter prior to completion. An additional 14 participants, including three with neuroimaging data, were excluded because the task could not be scored for accuracy and efficient shifting as described below. These participants either never reached criterion for the first sorting rule or reached it so late in the task that they never encountered a detectable change in sorting rule. Thus, 230 participants contributed to the final WCST analysis (127 females; age:  $M = 59.1$  years,  $SD = 17.9$ ). These 230 cases were used to estimate median accuracy for each number of efficient shifts from which the cognitive persistence score was computed. This enables the cognitive persistence score to be replicated by other research groups using the median values derived from this relatively large sample of typical adults.

The sub-sample of participants with neuroimaging data were 31 middle-aged to older adults between 50- and 81-years-old (19 females; age:  $M = 60.2$  years,  $SD = 8.1$ ) from the Vaden et al. (2015) study. Participants were excluded from the neuroimaging study for a history of head trauma, seizures, self-reported CNS disorders, and contraindications for safe MRI scanning. Participants were selected because they were older than 50 years and exhibited normal hearing or mild hearing loss (mean pure tone thresholds  $< 32$  dB HL from 0.25 to 8 kHz in the better ear with no more than 10 dB differences between ears). Again, three participants whose WCST-64 data could not be scored were excluded from the validation analyses, resulting in a sample comprising 28 participants (17 females; age:  $M = 60.6$  years,  $SD = 8.3$ ).

### 2.2. Development of Cognitive Persistence Measure

**2.2.1. Procedure**—Participants completed a computerized version of the standard WCST-64 (Kongs et al., 2000). Scores on this abbreviated 64-card version are generally comparable to those obtained on the full-length WCST (Greve, 2001), in which cards are administered until the participant completes 6 categories or 128 trials (Grant & Berg, 1948; Heaton, 1981). A test card containing a set of figures was presented on each trial, with the number (1–4), shape (triangle, circle, cross, star) and color (red, blue, yellow, green) of the figures varying from trial to trial. Four distinct key cards remained on screen for all trials, and participants were asked to sort each test card into one of the four key card groups. A correct response required matching the test card to the key card that had the same number, shape, or color, depending on the current sorting rule. Participants were not instructed how

to sort the cards, but received feedback on their response accuracy. The sorting rule changed every time participants correctly sorted ten cards in a row. All participants completed 64 trials, encountering between 1 and 6 sorting sequences depending on performance. A measure of cognitive persistence was obtained from the WCST-64 by controlling accuracy scores for set-shifting ability, as described below.

**2.2.2. Scoring**—We used the WCST-64 to examine set-shifting ability and the new estimate of cognitive persistence. Following Barceló and Knight (2002), we assessed set-shifting ability by coding the number of “efficient shifts.” An efficient shift occurs when a sequence of trials with a given sorting rule is completely free from errors following a rule change, excepting errors that result from optimal switching. Thus, participants make an efficient shift when they switch to a new rule on the trial immediately following the first trial that signals a rule change (i.e., the first error; see Figure 1). Optimal switching will yield errors on approximately 50% of trials immediately following detection of a rule change, because the participant must guess between the two remaining rules upon discovering that the former rule is incorrect (Barceló & Knight, 2002). Feedback after trying the new rule dictates whether to keep using the new rule or to switch to the remaining rule.

A critical insight of the Barceló and Knight (2002) scoring system is that errors resulting from the strategy that maximizes performance should not be counted against the participant, because they reflect efficient shifting processes. Thus, detection errors and efficient errors, which were expected from optimal performance, were treated differently from other errors when coding accuracy. Table 1 provides definitions and scoring of different error types, with specific examples given in Figure 1. Sequences of trials that were completely correct or correct except for an efficient error were considered efficient shifts.

Efficient shifts after a rule change were counted only for blocks where participants reached the criterion of 10 correct trials in a row. Thus, if the participant reached the end of the task before reaching criterion for a given sequence, that sequence was not counted as efficient because it could not be determined if the participant would have responded correctly for the entire sequence. However, the trials from such incomplete sequences were still included in the computation of accuracy. Only trials/sequences that occurred after the first detectable rule change (i.e., post-switch trials) were included in scoring, as shifting processes cannot be evaluated prior to a rule change. The first sequence of trials was excluded from the accuracy measure and used to estimate task learning or establishment of task-set. Accuracy thus reflects the proportion correct of post-switch trials.

Participants who did not finish any sequences after the first one were assigned a score of  $-1$  for number of efficient shifts, as efficient sequences were not possible for these participants. Thus, the number of efficient shifts ranged from  $-1$  to 4, where 4 was the maximum number of sequences that could be completed following the first rule change. Participants who had no trials after a detectable rule change were excluded from analyses because they had no post-switch trials that could be scored for accuracy. This exclusionary criterion included participants who never completed the first sequence and participants who completed the first sequence just before the end of the task.

The WCST-64 contains some trials for which the test card matches a key card on more than one dimension. These required special coding rules, which are described in Table 2. Detailed examples of three of these special cases are depicted in Figure 2. All of the coding rules dealt with such multidimensional responses by favoring the response that followed the optimal strategy, thereby maximizing participants' number of efficient shifts and accuracy. Thus, multidimensional responses were handled consistently across instances. As shown in the results below (see Figure 4), participants' responses ranged the entire spectrum in number of efficient shifts and only a small proportion (<5%) of participants achieved the maximum number of efficient shifts.

**2.2.3. Scoring Quality Control**—There was 100% intra-rater consistency for scoring efficient shifts and 99% intra-rater consistency for scoring trial accuracy across the data for a selected subset of 30 participants. There was 96% inter-rater consistency for scoring efficient shifts and 99% inter-rater consistency for scoring trial accuracy for the 28 neuroimaging participants. Raters met to discuss and resolve differences in scoring.

**2.2.4. Statistical Analyses**—We used non-parametric analyses to evaluate significance of our results because the number of efficient shifts was ordinal and not normally distributed. Specifically, non-parametric bootstrapping with 10,000 samples in the 'boot' package (version 1.3–18; Canty & Ripley, 2016; Davison & Hinkley, 1997) for R (version 3.3.1; R Core Team, 2016) was used to compute the bias-corrected and accelerated (BCa; Efron, 1987) 95% Confidence Interval (CI) for each statistic of interest. Correlation and mean difference statistics were considered significant if their 95% CIs did not include 0. Group medians (see below) were considered significantly different if their 95% CIs did not overlap.

**2.2.5. Calculation of Cognitive Persistence**—The measure of cognitive persistence was obtained by residualizing accuracy for set-shifting ability by controlling for the number of efficient shifts. Again, the rationale behind this approach was that accuracy on the WCST emerges primarily as a function of set-shifting ability and that individual differences in accuracy that are not explained by set-shifting ability are a result of differences in the ability or willingness to apply effort to overcome difficulty, a.k.a., cognitive persistence. After controlling for set-shifting ability, individuals with greater persistence will have higher accuracy because they continue to apply effort toward successful sorting even after receiving negative feedback, systematically searching for and using the correct rule.

An assumption underlying this residual accuracy approach is that the number of efficient shifts predicts accuracy. To ensure that this was the case, we conducted a non-parametric Spearman's rank correlation to assess the strength of the relationship between the number of efficient shifts and accuracy. Then, the median accuracy at each efficient shift value was used to compute accuracy residuals, i.e., the difference between a participant's observed accuracy and the median accuracy given their number of efficient shifts.

While the number of efficient shifts primarily reflects set-shifting ability, it may also reflect other cognitive abilities. Specifically, individuals who learn the task more quickly may have more efficient shifts because they have more opportunities to make an efficient shift.



Learning the task requires abstract reasoning to discern and test different sorting rules (Heaton, 1981; Kongs et al., 2000) and working memory to update the space of possible rules following feedback (Gold, Carpenter, Randolph, Goldberg, & Weinberger, 1997; Hartman, Bolton, & Fehnel, 2001; Lange et al., 2016). Additionally, achieving an efficient shift following a rule change requires working memory to successfully maintain the new rule for the remainder of the set (Buchsbaum, Greer, Chang, & Berman, 2005; Steinmetz & Houssemand, 2011). Moreover, research on the latent structure of executive functions has demonstrated that individual differences in set-shifting measures, like number of efficient shifts, also reflect variance in an inhibition function that is applied to all executive function tasks (Miyake & Friedman, 2012). Thus, adjusting accuracy for number of efficient shifts not only controls for set-shifting ability, but should also control for abstract reasoning, working memory, and inhibition. Such additional controls would help ensure that residual accuracy reflects differences in persistence rather than differences in other cognitive functions that affect WCST performance.

### 2.3. Behavioral Validation of Cognitive Persistence

We examined behavioral measures of abstract reasoning, working memory, and performance consistency to assess the convergent and divergent validity of cognitive persistence. If adjusting accuracy for number of efficient shifts successfully controls for other cognitive abilities employed during WCST performance, then number of efficient shifts should correlate with measures of abstract reasoning and working memory, but cognitive persistence should not. In contrast, cognitive persistence, but not number of efficient shifts, should correlate with consistent performance over the course of a task.

**2.3.1. Abstract Reasoning**—The number of trials to complete the first category on the WCST-64 is a standard metric of learning, where fewer trials indicate better abstract reasoning and problem solving (Kongs et al., 2000). Participants who complete the first category in fewer trials may have more opportunities to make efficient shifts.

**2.3.2. Working Memory**—The logical memory and family pictures immediate recall subtests from the Wechsler Memory Scale-III Abbreviated (WMS; Wechsler, 1997, 2002) were used to assess auditory and visual working memory, respectively. Although these subtests do not directly index working memory, both auditory and visual memory are significantly correlated with working memory in factor analyses of the WMS (Pauls, Petermann, & Lepach, 2013; Price, Tulsky, Millis, & Weiss, 2002). The logical memory subtest required participants to repeat stories told to them by the experimenter as accurately as possible for a maximum score of 75. The family pictures memory subtest required participants to recall details from a set of visual scenes—specifically, the characters, their locations, and their actions—for a maximum score of 64.

**2.3.3. Performance Consistency**—Participants completed the Continuous Performance Test-2 (CPT; Conners, 2000). The CPT entails responding as quickly as possible to a series of visually presented letters, but withholding responses for the letter 'X'. The standard error of a participant's reaction time for hits (correct responses to target letters) was used to evaluate performance consistency over the course of a task. Higher standard errors indicate

that participants exhibit more varied response times, reflecting inconsistent performance. Performance consistency has been linked to the ability or motivation to sustain effort for the duration of a task (Douglas, 1999; Leth-Steensen, King Elbaz, & Douglas, 2000). Specifically, evidence from clinical populations including attention-deficit/hyperactivity disorder and traumatic brain injury demonstrates that individuals exhibiting lower cognitive effort (i.e., sub-par performance unrelated to their condition) have less consistent response times on the CPT (Busse & Whiteside, 2012; Ord, Boettcher, Greve, & Bianchini, 2010; Suhr, Sullivan, & Rodriguez, 2011).

## 2.4. Cingulo-opercular Validation of Cognitive Persistence

Cingulo-opercular regions are engaged during the WCST and similar tasks (Hampshire, Gruszka, Fallon, et al., 2008; Konishi et al., 2002; Konishi et al., 1998; Nagahama et al., 1998), and there is evidence from lesion, animal, and imaging studies that cingulo-opercular regions function to identify and overcome impediments to a desired goal (di Pellegrino, Ciaramelli, & Ladavas, 2007; Holec, Pirot, & Euston, 2014; Kerns, 2006; Kerns et al., 2004; Ridderinkhof et al., 2004; Tow & Whitty, 1953). In other words, activity in cingulo-opercular regions indicates when people persist despite task difficulty.

For these reasons, measures of cingulo-opercular function were obtained from a challenging task that required participants to listen to words presented in babble at two relatively poor signal-to-noise ratios (SNRs) and repeat the words they heard. This task elicits significant cingulo-opercular activity when participants make an error in repeating a word (Vaden et al., 2015), reflecting a neural response to uncertainty in difficult task conditions. No explicit response feedback is given, but prior work has shown that the cingulo-opercular network responds to error commission even in the absence of feedback (Ham et al., 2013; Mars et al., 2005; Neta, Schlaggar, & Petersen, 2014; van Veen, Holroyd, Cohen, Stenger, & Carter, 2004), perhaps as a result of ongoing response conflict or uncertainty (Botvinick et al., 2001; Yeung, Botvinick, & Cohen, 2004). We define increases in cingulo-opercular activity associated with performance errors on the immediately preceding trial as “error responses.” Moreover, this task has been used to show that elevated cingulo-opercular activity predicts word recognition on a subsequent trial (Vaden et al., 2015), indicating that this activity helps to overcome task difficulty and improve performance. We define variation in cingulo-opercular activity that predicts correct performance on the next trial as “adaptive control.”

To validate that the WCST measure of persistence corresponded to neural activity that responds to and overcomes task difficulty (Eckert et al., 2016), we examined the extent to which individual differences in cognitive persistence were related to 1) error responses and 2) adaptive control. Except where otherwise indicated, fMRI analyses were conducted using SPM8 (Wellcome Trust Centre for Neuroimaging, University College London). The fMRI speech recognition task and analyses are summarized here for clarity. Complete details of data acquisition and preprocessing are given in Vaden et al. (2015).

**2.4.1. Speech Recognition in Noise fMRI Paradigm**—Participants heard 120 monosyllabic words (Dirks, Takayanagi, Moshfegh, Noffsingler, & Fausti, 2001) presented binaurally at +3 or +10 dB SNR (Vaden et al., 2015; Vaden et al., 2013) with a multi-talker

babble (Kalikow, Stevens, & Elliott, 1977) presented at 82 dB sound pressure level. SNR was blocked so that between four and six words were presented consecutively at each SNR. Whole-brain functional images were collected every 8.6 s using a sparse single-shot echo-planar imaging sequence. Each word presentation involved image acquisition (from 0–1.6 s), a delay (from 1.6–3.1 s) to limit forward masking of the speech by the scanner noise, word presentation (from 3.1–4.1 s), a response window to repeat the word (from 4.1–6.1 s), and an inter-trial interval (from 6.1–8.6 s). This interval was selected to allow head motion to diminish prior to acquisition of the next functional image and to capture the BOLD response when it was maximal, while attempting to limit the length of each TR and the duration of the experiment. Babble played continuously throughout the task. Additional images were obtained at rest (30 TRs) and during babble-only periods (30 TRs). Responses were scored as correct only when participants repeated the stimulus word exactly as presented, and missing and uninterpretable responses were excluded from analyses.

**2.4.2. Analysis of Adaptive Control**—Cingulo-opercular regions of interest (ROIs) that predicted subsequent word recognition were defined from an independent sample of 18 younger adult participants who had completed the same speech recognition in noise task (Vaden et al., 2013). BOLD contrast from the preceding trial was used to predict word recognition on the next trial. As described in that study, *lme4* (version 0.999375.42; Bates, Maechler, & Bolker, 2011) and *AnalyzefMRI* (version 1.1–14; Bordier, Dojat, & de Micheaux, 2011) packages for R statistics software (version 2.15.0; R Core Team, 2012) were used to compute logistic mixed-effects models of word recognition accuracy in each voxel, with preceding BOLD activity, SNR, preceding BOLD  $\times$  SNR interaction and random subject effects as predictors. The first word of each SNR block was excluded so that recognition accuracy did not reflect flexibility in shifting between SNRs. There were four clusters within the cingulo-opercular network that demonstrated significant main effects (voxel  $p_{\text{unc}} < 0.001$ , cluster  $p_{\text{FWE}} < 0.05$ ) of preceding BOLD activity on word recognition: the left IFG, the right AIFO, the anterior dACC, and the posterior dACC. A mask of the significant cingulo-opercular clusters was warped into a study-specific template generated from the native anatomical images of all 31 middle-aged to older adult participants (Vaden et al., 2015; Vaden et al., 2016) using Advanced Normalization Tools (ANTs version 1.9, <http://picsl.upenn.edu/software/ants/>; Avants et al., 2011).

The independently-defined cingulo-opercular mask (see Figure 3) was used to calculate standard estimates (betas) of the relation between cingulo-opercular activity and word recognition on the next trial for each older adult participant. Logistic regression models of word recognition accuracy (excluding the first word of each SNR block) were computed separately for each participant, with preceding BOLD activity averaged across all cingulo-opercular voxels, SNR, and preceding BOLD  $\times$  SNR interaction as factors. Each beta estimate of preceding BOLD activity thus reflects the strength of the relationship between cingulo-opercular activity and subsequent word recognition performance for that participant. A multiple regression analysis examined the extent to which cognitive persistence and number of efficient shifts predicted variability in these adaptive-control betas. Follow-up analyses of significant effects examined betas that were estimated separately for each of the adaptive-control ROIs (dACC, left IFG and right AIFO). Initial analyses indicated that beta

estimates from the anterior dACC and the posterior dACC were highly correlated ( $r = 0.73$ ,  $p < 0.001$ ), so these regions were collapsed into a single ROI.

**2.4.3. Analysis of Error Responses**—Cingulo-opercular ROIs where cortex was responsive to word recognition errors were defined from the independent sample of 18 younger adults described above. A general linear model was used to model the BOLD time-series as a function of +3 dB SNR, +10 dB SNR, babble-only trials, transition trials (marked by changes between rest, babble-only, and task blocks), and parametric modulators for word recognition accuracy on each SNR predictor (see Vaden et al., 2013). Significantly greater BOLD activity for incorrect relative to correct word recognition (voxel  $p_{\text{unc}} < 0.001$ , cluster  $p_{\text{fwe}} < 0.05$ ) was observed in 3 clusters within the cingulo-opercular network, which were warped into a study-specific template as described above. The error-response ROIs (dACC, left IFG, right AIFO) had a greater spatial extent than the adaptive-control ROIs (859 vs. 410 voxels). The two sets of ROIs had 264 voxels in common (26.3% overlap; see Figure 3).

The BOLD contrast for word recognition errors (incorrect > correct) was computed for the 28 middle-aged to older adult participants using the same general linear model predictors used to identify the error-response ROIs (see Vaden et al., 2015). Then, the mean beta value across voxels in the SPM contrast map that fell within the error-response mask was computed for each participant. Follow-up analyses used mean beta values calculated separately for each of the three error-response ROIs.

### 3. Results

The relationship between number of efficient shifts and accuracy is shown in Figure 4. As expected, the number of efficient shifts was strongly related to accuracy (Spearman's rank correlation:  $\rho = 0.90$ ,  $p < 0.001$ ). We used the median accuracy at each efficient shift value to determine the expected accuracy for each participant (see Table 3); for instance, a participant with 0 efficient shifts had an expected accuracy of 0.7143. To define cognitive persistence, we subtracted expected accuracy,  $\hat{y}$ , from observed accuracy,  $y$ . A positive cognitive persistence value indicates that an individual is performing better than predicted by their number of efficient shifts, while a negative value indicates that they are performing worse (see Figure 4). The magnitude of the cognitive persistence value indicates how much an individual's actual proportion correct deviates from their expected proportion correct.

#### 3.1. Predictors of Cognitive Persistence

Correlations between demographic variables and scores on WCST measures are reported in Table 4. Participants exhibited fewer efficient shifts when they were older and when they had fewer years of education. In contrast, cognitive persistence was not significantly correlated with age or education. Thus, controlling for number of efficient shifts effectively controlled for education and age-related declines in task ability.

Correlations of the number of efficient shifts and persistence with measures of abstract reasoning, working memory and performance consistency are reported in Table 5. As hypothesized, the number of efficient shifts was related to abstract reasoning and working memory, but not to performance consistency, whereas cognitive persistence showed the

reverse pattern. Specifically, participants who had more efficient shifts required fewer trials to complete the first category (see Table 5), suggesting that number of efficient shifts in part reflects abstract reasoning abilities required to learn the task. They also had higher scores on both the auditory and visual subtests of the WMS, indicating the role of working memory in updating and maintaining the sorting rule to achieve efficient shifts. In contrast, cognitive persistence was only related to performance consistency—participants with higher cognitive persistence had smaller standard errors of reaction time on the CPT, indicating that they responded more consistently across trials. This pattern of results demonstrates that the cognitive persistence measure reflects sustaining performance for the duration of a task rather than other abilities employed during the WCST.

### 3.2. Adaptive Control Responses and Persistence

Multiple regression revealed that cognitive persistence ( $\beta = 0.39$ , 95% CI = [0.03, 0.64])<sup>1</sup>, but not number of efficient shifts ( $\beta = 0.07$ , 95% CI = [-0.26, 0.39]), was significantly related to adaptive-control beta estimates, i.e. the magnitude of the effect of cingulo-opercular activity on subsequent word recognition<sup>2</sup>. In other words, individuals with greater cognitive persistence showed a larger performance benefit following increased cingulo-opercular activity, independently of their number of efficient shifts. Follow-up analyses demonstrated that cognitive persistence was most strongly related to the adaptive-control betas from the dACC ROI ( $r = 0.40$ , 95% CI = [0.05, 0.65]; see Figure 5), as the correlations in the other cingulo-opercular ROIs did not reach significance (left IFG:  $r = 0.28$ , 95% CI = [-0.07, 0.59]; right AIFO:  $r = 0.23$ , 95% CI = [-0.16, 0.56]). These results indicate that varied performance on the WCST reflects, in part, individual variation in dACC function linked to enhancing performance on an effortful task.

### 3.3. Error Responses and Persistence

Multiple regression demonstrated that cognitive persistence ( $\beta = 0.55$ , 95% CI = [0.29, 0.77]), but not number of efficient shifts ( $\beta = 0.36$ , 95% CI = [-0.07, 0.68]), significantly predicted cingulo-opercular activity in response to errors. Follow-up analyses indicated that cognitive persistence was significantly correlated with error responses in both the left IFG ( $r = 0.51$ , 95% CI = [0.24, 0.70]) and the dACC ( $r = 0.39$ , 95% CI = [0.11, 0.61]). However, persistence was not related to right AIFO error-responses ( $r = 0.15$ , 95% CI = [-0.21, 0.50]).

Because cognitive persistence was significantly associated with both error-responses and adaptive-control effects, particularly within the dACC, the degree of independence of the persistence results was unclear. To investigate this question, we examined the extent to which adaptive-control betas in the dACC were associated with persistence after controlling for error-response betas in the same region. Multiple regression demonstrated that individual

<sup>1</sup>All  $\beta$ s indicate standardized estimates.

<sup>2</sup>To determine whether these results were affected by preceding trial accuracy, we examined adaptive control estimates separately for trials that followed correct responses versus errors. Two participants were excluded from the post-error analysis (GLM overfitting,  $n=1$ ; beta estimate  $>3$  SD from the mean,  $n=1$ ). Neither persistence ( $\beta = 0.15$ , 95% CI = [-0.16, 0.37]) nor number of efficient shifts ( $\beta = -0.02$ , 95% CI = [-0.34, 0.33]) predicted adaptive-control beta estimates following correct responses only. However, persistence significantly predicted adaptive-control beta estimates for post-error trials ( $\beta = 0.48$ , 95% CI = [0.11, 0.71]), whereas number of efficient shifts did not ( $\beta = 0.04$ , 95% CI = [-0.36, 0.44]). This pattern of results suggests that relationship between persistence and adaptive control primarily reflected performance adjustments following errors.

variation in the performance benefit from dACC activity was no longer significantly related to persistence ( $\beta = 0.28$ , 95% CI = [-0.15, 0.68]) after accounting for error-responses ( $\beta = 0.27$ , 95% CI = [-0.10, 0.62]). Thus, adaptive-control effects and error-responses in dACC account for the same variance in cognitive persistence.

In summary, the current findings provide evidence for regional specialization within the cingulo-opercular network: while individual variation in cognitive persistence predicted error responses in left IFG and dACC, the association between persistence and adaptive-control activity emerged only in dACC. The relationship between cognitive persistence and activity in dACC, which is theorized to detect and overcome performance difficulty (Botvinick et al., 2001; Botvinick et al., 2004; Eckert et al., 2016), provides a functional validation for the novel WCST-derived persistence measure.

#### 4. Discussion

Cognitive persistence is critical for understanding variation in task performance, but there are few measures of it. The results of this study demonstrate that cognitive persistence can be measured using variation in WCST performance that is unaccounted for by set-shifting abilities. This measure was conceptually validated by showing that individual variation in WCST persistence was related to activity in cingulo-opercular regions during a word recognition task both when older adults made an error and when they effectively adjusted their subsequent behavior. Indeed, individuals with greater WCST persistence exhibited a heightened error-response and greater adaptive control in dACC. They also demonstrated more consistent performance during the CPT, an independent measure of task effort over a sustained period. These results provide support for a novel use of the WCST to characterize persistence in typical adult populations.

The measure of cognitive persistence examined in the current study characterizes variation in WCST accuracy that differed from what is expected by set-shifting ability. This residual variance in part reflects differences in the willingness and ability to continue to apply effort to the task of finding and using the correct sorting rule, even after receiving repeated negative feedback. That is, sustained effort to succeed on the task can push accuracy above set-shifting ability whereas lack of effort can lower accuracy. For this reason, the residual accuracy can be thought of as a measure of the drive to optimize performance.

Controlling for set-shifting ability may remove some variance that is actually due to persistence, as individuals who exert more effort may also achieve more efficient shifts. However, our residual accuracy approach increases the purity of the persistence measure by removing variance in performance related to the range of cognitive abilities employed during the WCST. Crucially, the persistence measure was unrelated to education, trials to complete the first category on the WCST, and immediate recall subtests of the WMS, indicating that residual accuracy differences were not due to differences in academic achievement, abstract reasoning, or working memory. Individual variability in residual accuracy is not necessarily exclusively due to persistence, but could also reflect variance due to temporary psychological states such as mood or fatigue. Future work should examine the extent to which the persistence measure is reliable across repeated administrations.

To better understand the meaning of the residual accuracy measure, it is worth considering the patterns of performance reflected in the cognitive persistence construct. For example, consider performance after the rule change for the following participants who received efficient shift scores of  $-1$  and had between 25–33 post-switch trials. These are all individuals who struggled with the WCST, taking nearly half the task to complete the first category and never completing a second one, yet their variance in performance indicated different levels of persistence. One individual with higher persistence (0.283) had six separate runs of correct trials, averaging 4 trials per run (range: 1–7), but sometimes used the wrong rule after a correct multidimensional sort. Runs of incorrect trials lasted 1.8 trials on average (range: 1–3). Another individual with moderate persistence ( $-0.016$ ) had six shorter runs of correct trials, averaging 2 trials per run (range: 1–4). That participant sometimes applied the wrong rule following a correct unambiguous response, with an average length of 2.67 incorrect trials (range: 1–6). An individual with lower persistence ( $-0.084$ ) had five even shorter runs of correct trials, averaging 1.8 trials per run (range: 1–3). In that case, runs of incorrect trials were also longer, lasting an average of 3.2 trials (range: 1–6). All three individuals “discovered” the new sorting rule, in that they received correct feedback for an unambiguous response. However, with higher persistence, individuals were more likely to repeat a correct response and to return to the correct rule more quickly after an incorrect response, consistent with a greater drive to optimize performance.

Our persistence measure follows from recent efforts to decompose complex behavioral tasks into their component processes to better understand why patient populations exhibit atypical performance. For example, defining efficient or strategically-appropriate errors was used to determine that patients with prefrontal lesions make few of these errors (Barceló & Knight, 2002). We leveraged this approach to characterize WCST performance after accounting for efficient errors and their related efficient shifts.

The idea that performance on mentally-challenging tasks depends on individual differences in persistence in addition to cognitive ability is not new. Historically, psychologists have invoked the “concept of conation to support the intuition that the intention and motivation to act works in conjunction with cognition to influence performance and learning (Deci, Ryan, & Williams, 1996; Kanfer, Ackerman, & Heggestad, 1996; Militello, Gentner, Swindler, & Beisner, 2006; Snow, 1996; Wechsler, 1950; Wolters, Yu, & Pintrich, 1996). Conation, which is akin to will or drive, is abstractly defined as, “The aspect of mental processes or behavior directed toward action or change and including impulse, desire, volition, and striving” (“The American Heritage® Dictionary of the English Language,” 2016). This broad definition has perhaps made it difficult to assess conation empirically (Militello et al., 2006), and measures of conation have typically relied on participants’ self-evaluation (Kanfer et al., 1996; Wolters et al., 1996). By reconceptualizing conation as persistence, or the sustained application of effort to overcome task difficulty, we were able to concretely operationalize the concept as performance differences that are independent of task ability on the WCST.

To our knowledge, this study is the first to measure cognitive persistence from the WCST. Using a large lifespan sample, we have generated a look-up table (Table 3) to estimate a participant’s expected performance from their number of efficient shifts during the

computerized WCST-64 (Kongs et al., 2000). This table can be used to compute accuracy residuals as an index of cognitive persistence for healthy adults completing the WCST-64. However, the neural validation of this novel persistence measure was only performed in middle-aged to older adults, because recognizing speech in noise is known to require greater effort in this age group than in younger adults (Anderson Gosselin & Gagné, 2011; Edwards, 2007; Kuchinsky et al., 2013; Zekveld et al., 2011). The extent to which the persistence measure generalizes to other populations, including younger adults and neuropsychological patients, merits additional investigation.

An advantage of our novel persistence measure is that it simply requires a single administration of the standard WCST-64 task, so that all of the traditional metrics of WCST performance can be obtained in addition to the novel persistence measure. It is also straightforward to interpret: positive values indicate higher-than-expected and negative values lower-than-expected accuracy given set-shifting ability, and the magnitude indicates the size of the deviation from expected proportion correct. That is, a participant with a +0.1 persistence score has accuracy that is 10% higher than expected whereas a participant with a -0.1 persistence score has accuracy that is 10% lower than expected.

The look-up table can be applied to the original 128-card version of the WCST (Grant & Berg, 1948; Heaton, 1981) by scoring only the first 64 trials, which are identical to the abbreviated WCST-64 used here. Thus, scoring only the first 64 trials of the original WCST should yield comparable distributions of number of efficient shifts and proportion correct as obtained for the WCST-64. However, the table cannot be applied to scores derived from all completed trials of the original 128-card version, which is administered until the participant either achieves six categories or completes 128 trials. Because participants are able to accumulate more efficient shifts by completing more than 64 trials, scoring the entire 128-card version would alter the relationship between efficient shifts and accuracy. In particular, the number of efficient shifts could become inflated for participants who require more trials to complete six categories.

Some WCST versions, such as the Modified Card Sorting Test (Nelson, 1976) and the Madrid Card Sorting Test (Barceló, 2003), have been modified to omit ambiguous cards that match key cards on more than one stimulus dimension. These versions also typically provide more explicit instructions about the task, such as notifying participants when the rule changes (Nelson, 1976) or identifying the possible sorting rules in advance (Barceló, 1999, 2003; Barceló & Knight, 2002; Barceló, Sanz, Molina, & Rubia, 1997). These tasks have the advantages of being less frustrating, thereby reducing participant drop-out (Nelson, 1976), and permitting unambiguous scoring of responses, enabling precise trial-level analyses of different response types for neuroimaging studies (Barceló, 1999, 2003; Barceló & Knight, 2002; Barceló et al., 1997). However, it has been argued that the removal of ambiguous cards fundamentally changes the nature of the task (de Zubicaray & Ashton, 1996; Greve, 2001; Robinson, Kester, Saykin, Kaplan, & Gur, 1991), so it is not clear that these tasks could be used to measure persistence in the same way as the standard WCST-64. Indeed, because eliminating ambiguity is intended to reduce task difficulty and frustration, the modified versions may remove the very element of the task that makes it a useful measure of persistence. Persistence is important when participants must overcome an obstacle to



achieve a level of performance that is aligned with ability. Persistence should not materially affect performance for relatively easy tasks, except perhaps when the obstacle is sustaining a high level of performance for a long period of time. Whether or not persistence can be measured by controlling accuracy for number of efficient shifts in these modified versions remains an empirical question, but we caution that the presence of ambiguity may be essential to the utility of the WCST as a persistence measure.

Because the cognitive persistence metric was developed from the WCST-64, it is subject to the same limitations as the WCST-64 task. There may be some participant drop-out; 1% of our study participants failed to complete the task. Additionally, the task does not allow for unambiguous coding of responses. However, we provide a detailed method for handling multidimensional responses in order to code for efficient errors and consequently the number of efficient shifts on the standard WCST-64 task. This novel coding scheme is important because it enables researchers to index normal and adaptive shifting behavior on the standard WCST-64 in addition to the unambiguous Madrid Card Sorting Test (Barceló & Knight, 2002). Future studies should assess utility of number of efficient shifts on the WCST-64 for indexing prefrontal damage as well as its relationship with the same measure on the Madrid Card Sorting Test.

Some participants performed at ceiling on the WCST-64 task, making it difficult to measure their persistence. Participants who achieved 4 efficient shifts (11 of 230), the maximum number possible, were expected to have 100% accuracy on the task. As a result, participants who were performing at ceiling had cognitive persistence scores equal to 0. It is not possible to assess the extent to which participants persist beyond their ability-level when task ability alone enables perfect performance (i.e., when the task is too easy and does not require persistence). Thus, the best estimate of cognitive persistence for these participants is the “average” score of 0—they are doing as well as expected for their task ability. However, their actual cognitive persistence may be higher. This is comparable to ceiling effects on other cognitive tasks, where individuals who are performing at ceiling may actually have a better ability than can be measured on the task. The inclusion or exclusion of these few cases had limited influence on the results; however, imperfect performance and the need to overcome errors is critical for assessing persistence.

Finally, we note that 6% of our sample could not be scored for efficient shifts or persistence because they never completed the first category or had no post-switch trials. Put another way, it is not possible to measure persistence for individuals who do not appear to learn the task. One option to address this limitation may be to adopt a longer version of the WCST, thereby increasing the opportunity to complete the first category and encounter a rule change. In fact, 4 of the 14 participants who could not be scored ended the task with 8 or more correct responses in a row, suggesting that a longer task would enable scoring of these individuals. Future work should examine the possibility of measuring persistence from a longer version of the WCST.

#### 4.1. Neural Validation

To validate the cognitive persistence measure, we demonstrated that it predicted individual variability in cingulo-opercular function in middle-aged to older adults. Specifically, we

related persistence to cingulo-opercular activity that was linked to trial-level performance during a challenging word recognition in noise task. This activity is thought to reflect the persistence to overcome performance obstacles, at least in part, as it is routinely observed following increases in task difficulty due to changing task demands or decrements in performance (Botvinick et al., 2004; Botvinick et al., 1999; Carter, 1998; Crone et al., 2006; Durston et al., 2003; Eichele et al., 2008; Luks et al., 2007; Ridderinkhof et al., 2004; Weissman et al., 2006; Weissman et al., 2009) and appears to boost subsequent performance (Botvinick et al., 2004; Kerns, 2006; Kerns et al., 2004; Orr & Weissman, 2009; Ridderinkhof et al., 2004; Weissman et al., 2006). The relation with persistence was largely driven by the cingulate cortex (dACC) rather than the more lateral frontal cortex regions.

The cingulo-opercular results are consistent with the theoretical framework that the dACC detects task difficulty and/or performance decrements to adjust the recruitment of top-down control resources accordingly (Botvinick, 2007; Botvinick et al., 2004; Ridderinkhof et al., 2004). Specifically, dACC activity increases in response to conflict and predicts subsequent improvements in performance that are accompanied by increases in dorsolateral prefrontal activity (Kerns, 2006; Kerns et al., 2004). The dACC thus appears to signal to engage lateral frontal cortex to enact control when necessary, i.e., when increased effort is required to achieve a positive outcome. Moreover, damage to dACC is associated with avoidance of challenging but rewarding tasks (Holec et al., 2014; Tow & Whitty, 1953) and attenuates adjustments in behavior following stimulus conflict and response errors (di Pellegrino et al., 2007; cf. Fellows & Farah, 2005). Patients with dACC lesions also show reduced reactive changes in heart rate variability and blood pressure in response to conditions requiring mental effort (Critchley et al., 2003), and thus fail to prepare the autonomic system to handle challenges. These findings suggest that the dACC helps to improve performance in conditions when performance has declined or may decline because of changes in task difficulty.

Interestingly, set-shifting and persistence demonstrated a dissociation in their relationships with cingulo-opercular and behavioral effects. Set-shifting was related to measures of abstract learning and working memory, but not to performance consistency or the cingulo-opercular effects. In contrast, persistence was related to performance consistency and cingulo-opercular effects, but not to abstract learning or working memory. Specifically, persistence was associated with error responses in the left IFG and dACC, as well as dACC adaptive-control. The absence of a significant association between set-shifting and cingulo-opercular effects may seem to contradict prior studies that have reported increased activity in cingulo-opercular regions during rule switches (Hampshire, Gruszka, & Owen, 2008; Konishi et al., 2002; Konishi et al., 1998; Lao-Kaim et al., 2015). However, we intentionally measured cingulo-opercular activity during error detection and subsequent improvements in performance—that is, activity related to overcoming obstacles. We believe this is the reason that persistence was selectively related to cingulo-opercular function in the present study. If we had measured cingulo-opercular activity during a set-shifting task, then we would indeed expect it to relate to number of efficient shifts on the WCST-64.

The extent of dACC error responses accounted for the relationship between persistence and dACC-related performance improvements. This is consistent with the premise that the

magnitude of the dACC response to difficulty or error drives subsequent improvements in performance (Botvinick et al., 2001; Botvinick et al., 2004). Heightened sensitivity to errors may indicate that there is value in applying effort, while heightened adaptive control may indicate the successful application of effort to overcome difficulty. In this way, error responses and adaptive control in dACC may reflect the drive to perform well and follow-through in exerting effort, respectively.

We view the link between dACC activity and subsequent performance as an epiphenomenon of sustained task-related effort. Recent work by Holroyd and colleagues (Holroyd & Umemoto, 2016; Holroyd & Yeung, 2012) characterizes dACC activity as providing motivation for protracted, goal-directed behaviors. A general boost in dACC function, and thus the drive to optimize overall task performance, would mean that the dACC is more sensitive to performance difficulty and more likely to signal increases in effort to set performance back on track. For this reason, dACC function may underlie individual differences in personality traits including reward sensitivity and persistence (Holroyd & Umemoto, 2016). However, this difference in dACC function should only matter for effortful tasks, like the WCST and the speech recognition in noise task; there is no cost to overcome when good performance is achieved with little effort or control, so the dACC is not necessary in these contexts (Holroyd & Umemoto, 2016; Holroyd & Yeung, 2012). This interpretation helps explain why dACC lesions produce deficits in repeating rewarded responses across trials (Kennerley, Walton, Behrens, Buckley, & Rushworth, 2006) and overcoming obstacles to achieve a reward (Holec et al., 2014), but do not always affect trial-level performance adjustments (Fellows & Farah, 2005; Kennerley et al., 2006). Because the WCST is an effortful task without explicit rewards, we predict that individuals with dACC lesions should have lower accuracy than expected from their number of efficient shifts, and thus, negative persistence values. However, individuals who start at ceiling on the WCST might be spared from these effects due to the ease with which they perform the task.

Recent computational models of dACC function suggest that the dACC encodes the value anticipated from the effort required to improve performance or receive a reward (Shenhav, Botvinick, & Cohen, 2013; Shenhav, Cohen, & Botvinick, 2016; Shenhav et al., 2017; Verguts, Vassena, & Silvetti, 2015). From this perspective, variability in persistence across individuals may emerge from two sources, differences in the valuation of the reward and differences in the perceived efficacy of applying effort. Individuals who place less value on positive performance outcomes and individuals who believe that effort is unlikely to improve performance should be less likely to engage dACC to apply effort in the service of performance. One implication of this framework is that dACC function and consequently, persistence, can be shifted by task parameters: individuals should exhibit more persistence on tasks that are perceived as having a relatively high value from the investment of effort (Shenhav et al., 2013; Shenhav et al., 2017; Verguts et al., 2015), regardless of task modality. Indeed, performance on cognitive tasks can be improved simply by increasing incentive values (for review, see Shenhav et al., 2017). This potential dependency between value and effort is an important consideration for future development of persistence tasks. Nevertheless, the WCST persistence measure appears to characterize a trait that varies across participants, but that is consistent within participants across tasks.

Our conceptualization of cognitive persistence as the application of effort to improve performance is closely related to motivation. Individuals who are motivated to perform well are more likely to apply effort to achieve task goals, and motivation has been shown to explain about 5% of the variance in task performance after accounting for cognitive ability (Kanfer et al., 1996). However, our persistence measure focuses on the ultimate impact that motivation has on performance (via the application of effort) rather than the extent of motivation per se. The advantage to this approach is that it directly probes the amount of effort individuals employ relative to their cognitive ability.

## 4.2. Conclusions

We developed a novel measure of cognitive persistence from the WCST-64 and validated its relationship with cingulo-opercular activity that predicts subsequent performance on a speech recognition in noise task. Cognitive persistence can be thought of as the application of effort in the face of task difficulty to achieve a goal. Because successful performance on mentally-challenging tasks depends on both persistence and cognitive ability, these functions are typically confounded. By definition, our measure of cognitive persistence varies independently from set-shifting ability, a core executive function that is essential to performance on the WCST. Thus, it offers the potential to dissociate contributions of prefrontal cortex to persistence versus set-shifting and to understand how persistence versus set-shifting are affected by different neurobiological factors.

## Acknowledgments

The researchers would like to thank Lois Matthews for her assistance with data management, Jayne Ahlstrom for her assistance with stimulus development, Loretta Tsu for her assistance with data collection, and Anita Smalls for her assistance with data coding.

### Funding

This work was supported (in part) by the National Institutes of Health / National Institute on Deafness and Other Communication Disorders (P50 DC 000422 and T32 DC0014435), MUSC Center for Biomedical Imaging, South Carolina Clinical and Translational Research (SCTR) Institute, NIH/NCRR Grant number UL1 RR029882. This investigation was conducted in a facility constructed with support from Research Facilities Improvement Program (C06 RR14516) from the National Center for Research Resources, National Institutes of Health.

## References

- The American Heritage® Dictionary of the English Language. 5th2016. Retrieved from <https://ahdictionary.com/>
- Anderson Gosselin P, Gagné JP. Older adults expend more listening effort than young adults recognizing speech in noise. *Journal of Speech Language and Hearing Research*. 2011; 54(3):944. DOI: 10.1044/1092-4388(2010/10-0069).
- Avants BB, Tustison NJ, Song G, Cook PA, Klein A, Gee JC. A reproducible evaluation of ANTs similarity metric performance in brain image registration. *NeuroImage*. 2011; 54(3):2033–2044. DOI: 10.1016/j.neuroimage.2010.09.025 [PubMed: 20851191]
- Barceló F. Electrophysiological evidence of two different types of error in the Wisconsin Card Sorting Test. *NeuroReport*. 1999; 10:1299–1303. DOI: 10.1097/00001756-199904260-00027 [PubMed: 10363943]
- Barceló F. The Madrid Card Sorting Test (MCST): A task switching paradigm to study executive attention with event-related potentials. *Brain Research Protocols*. 2003; 11(1):27–37. DOI: 10.1016/s1385-299x(03)00013-8 [PubMed: 12697260]

- Barceló F, Knight RT. Both random and perseverative errors underlie WCST deficits in prefrontal patients. *Neuropsychologia*. 2002; 40:349–356. DOI: 10.1016/S0028-3932(01)00110-5 [PubMed: 11684168]
- Barceló F, Sanz M, Molina V, Rubia FJ. The Wisconsin Card Sorting Test and the assessment of frontal function: A validation study with event-related potentials. *Neuropsychologia*. 1997; 35(4): 399–408. DOI: 10.1016/S0028-3932(96)00096-6 [PubMed: 9106269]
- Bates, D., Maechler, M., Bolker, B. *lme4: Linear mixed-effects models using eigen and jags*. R package version 0.999375-42. 2011. Retrieved from <http://cran.R-project.org/package=lme4>
- Berman KF, Ostrem JL, Randolph C, Gold J, Goldberg TE, Coppola R, Weinberger DR. Physiological activation of a cortical network during performance of the Wisconsin Card Sorting Test: A positron emission tomography study. *Neuropsychologia*. 1995; 33(8):1027–1046. DOI: 10.1016/0028-3932(95)00035-2 [PubMed: 8524452]
- Bordier C, Dojat M, de Micheaux PL. Temporal and spatial independent component analysis for fMRI data sets embedded in the AnalyzeFMRI R package. *Journal of Statistical Software*. 2011; 44(9): 1–24. DOI: 10.18637/jss.v044.i09
- Botvinick MM. Conflict monitoring and decision making: Reconciling two perspectives on anterior cingulate function. *Cognitive, Affective & Behavioral Neuroscience*. 2007; 7(4):356–366. DOI: 10.3758/CABN.7.4.356
- Botvinick MM, Braver TS, Barch DM, Carter CS, Cohen JD. Conflict monitoring and cognitive control. *Psychological Review*. 2001; 108(3):624–652. DOI: 10.1037/0033-295X.108.3.624 [PubMed: 11488380]
- Botvinick MM, Cohen JD, Carter CS. Conflict monitoring and anterior cingulate cortex: An update. *Trends in Cognitive Sciences*. 2004; 8(12):539–546. DOI: 10.1016/j.tics.2004.10.003 [PubMed: 15556023]
- Botvinick MM, Nystrom LE, Fissell K, Carter CS, Cohen JD. Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature*. 1999; 402:179–181. DOI: 10.1038/46035 [PubMed: 10647008]
- Buchsbaum BR, Greer S, Chang WL, Berman KF. Meta-analysis of neuroimaging studies of the Wisconsin Card-Sorting Task and component processes. *Human Brain Mapping*. 2005; 25(1):35–45. DOI: 10.1002/hbm.20128 [PubMed: 15846821]
- Busse M, Whiteside D. Detecting suboptimal cognitive effort: Classification accuracy of the Conner's Continuous Performance Test-II, Brief Test of Attention, and Trail Making Test. *The Clinical Neuropsychologist*. 2012; 26(4):675–687. DOI: 10.1080/13854046.2012.679623 [PubMed: 22533714]
- Canty, A., Ripley, B. *boot: Bootstrap R (s-plus) functions*. R package version 1.3-18. 2016. Retrieved from <http://cran.R-project.org/package=boot>
- Carter CS. Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science*. 1998; 280:747–749. DOI: 10.1126/science.280.5364.747 [PubMed: 9563953]
- Choi J, Mogami T, Medalia A. Intrinsic motivation inventory: An adapted measure for schizophrenia research. *Schizophrenia Bulletin*. 2010; 36(5):966–976. DOI: 10.1093/schbul/sbp030 [PubMed: 19386577]
- Cloninger, CR., Przybeck, TR., Svrakic, DM., Wetzler, RD. *The Temperament and Character Inventory (TCI): A guide to its development and use*. St. Louis, Missouri: Center for Psychobiology of Personality; 1994.
- Conners, CK. *Conners' Continuous Performance Test (CPT-2) computer program for windows, technical guide, and software manual*. Toronto, ON: Multi Health Systems Inc; 2000.
- Critchley HDM, Josephs CJ, O'Doherty J, Zanini S, Dewar B-K, Cipolotti L, Dolan R J. Human cingulate cortex and autonomic control: Converging neuroimaging and clinical evidence. *Brain*. 2003; 126:2139–2152. DOI: 10.1093/brain/awg216 [PubMed: 12821513]
- Crone EA, Wendelken C, Donohue SE, Bunge SA. Neural evidence for dissociable components of task-switching. *Cerebral Cortex*. 2006; 16(4):475–486. DOI: 10.1093/cercor/bhi127 [PubMed: 16000652]

- Dang LC, Donde A, Madison C, O'Neil JP, Jagust WJ. Striatal dopamine influences the default mode network to affect shifting between object features. *Journal of Cognitive Neuroscience*. 2012; 24(9): 1960–1970. DOI: 10.1162/jocn\_a\_00252 [PubMed: 22640392]
- Daughters SB, Lejuez CW, Bornovalova MA, Kahler CW, Strong DR, Brown RA. Distress tolerance as a predictor of early treatment dropout in a residential substance abuse treatment facility. *Journal of Abnormal Psychology*. 2005; 114(4):729–734. DOI: 10.1037/0021-843X.114.4.729 [PubMed: 16351393]
- Daughters SB, Lejuez CW, Kahler CW, Strong DR, Brown RA. Psychological distress tolerance and duration of most recent abstinence attempt among residential treatment-seeking substance abusers. *Psychology of Addictive Behaviors*. 2005; 19(2):208–211. DOI: 10.1037/0893-164X.19.2.208 [PubMed: 16011392]
- Daughters SB, Lejuez CW, Strong DR, Brown RA, Breen RB, Lesieur HR. The relationship among negative affect, distress tolerance, and length of gambling abstinence attempt. *Journal of Gambling Studies*. 2005; 21(4):363–378. DOI: 10.1007/s10899-005-5554-6 [PubMed: 16311873]
- Davison, AC., Hinkley, DV. *Bootstrap Methods and their Applications*. Cambridge: Cambridge University Press; 1997.
- de Zubicaray G, Ashton R. Nelson's (1976) Modified Card Sorting Test: A review. *The Clinical Neuropsychologist*. 1996; 10(3):245–254. DOI: 10.1080/13854049608406687
- Deci EL, Ryan RM, Williams GC. Need satisfaction and the self-regulation of learning. *Learning and Individual Differences*. 1996; 8(3):165–183. DOI: 10.1016/S1041-6080(96)90013-8
- di Pellegrino G, Ciaramelli E, Làdavas E. The regulation of cognitive control following rostral anterior cingulate cortex lesion in humans. *Journal of Cognitive Neuroscience*. 2007; 19(2):275–286. DOI: 10.1162/jocn.2007.19.2.275 [PubMed: 17280516]
- Dirks DD, Takayanagi S, Moshfegh A, Noffsingler PD, Fausti SA. Examination of the neighborhood activation theory in normal and hearing-impaired listeners. *Ear & Hearing*. 2001; 22(1):1–13. DOI: 10.1097/00003446-200102000-00001 [PubMed: 11271971]
- Doherty-Bigara J, Gilmore L. Development of the Dimensions of Adult Mastery Motivation Questionnaire. *The Australian Educational and Developmental Psychologist*. 2016; 32(2):142–157. DOI: 10.1017/edp.2015.18
- Douglas, VI. Cognitive control processes in attention deficit hyperactivity disorder. In: Quay, HC., Hogan, AE., editors. *Handbook of disruptive behavior disorders*. New York: Plenum Press; 1999. p. 105-138.
- Drewe EA. The effect of type and area of brain lesion on Wisconsin Card Sorting Test performance. *Cortex*. 1974; 10(2):159–170. DOI: 10.1016/S0010-9452(74)80006-7 [PubMed: 4844468]
- Durston S, Davidson MC, Thomas KM, Worden MS, Tottenham N, Martinez A, Casey BJ. Parametric manipulation of conflict and response competition using rapid mixed-trial event-related fMRI. *NeuroImage*. 2003; 20:2135–2141. DOI: 10.1016/j.neuroimage.2003.08.004 [PubMed: 14683717]
- Eckert MA, Teubner-Rhodes S, Vaden KI Jr. Is listening in noise worth it? The neurobiology of speech recognition in challenging listening conditions. *Ear & Hearing*. 2016; 37(Suppl 1):101S–110S. DOI: 10.1097/AUD.0000000000000300 [PubMed: 27355759]
- Edwards B. The future of hearing aid technology. *Trends in Amplification*. 2007; 11(1):31–45. DOI: 10.1177/1084713806298004 [PubMed: 17301336]
- Efron B. Better bootstrap confidence intervals. *Journal of the American Statistical Association*. 1987; 82(397):171–185. DOI: 10.2307/2289144
- Eichele T, Debener S, Calhoun VD, Specht K, Engel AK, Hugdahl K, Ullsperger M. Prediction of human errors by maladaptive changes in event-related brain networks. *Proceedings of the National Academy of Sciences*. 2008; 105(16):6173–6178. DOI: 10.1073/pnas.0708965105
- Fellows LK, Farah MJ. Is anterior cingulate cortex necessary for cognitive control? *Brain*. 2005; 128(Pt 4):788–796. DOI: 10.1093/brain/awh405 [PubMed: 15705613]
- Godinez DA, Friedman NP, Rhee SH, Miyake A, Hewitt JK. Phenotypic and genetic analyses of the Wisconsin Card Sort. *Behavior Genetics*. 2012; 42(2):209–220. DOI: 10.1007/s10519-011-9502-1 [PubMed: 21952792]

- Gold JM, Carpenter C, Randolph C, Goldberg TE, Weinberger DR. Auditory working memory and Wisconsin Card Sorting Test performance in schizophrenia. *Archives of General Psychiatry*. 1997; 54:159–165. [PubMed: 9040284]
- Graham S, Phua E, Soon CS, Oh T, Au C, Shuter B, Yeh IB. Role of medial cortical, hippocampal and striatal interactions during cognitive set-shifting. *NeuroImage*. 2009; 45(4):1359–1367. DOI: 10.1016/j.neuroimage.2008.12.040 [PubMed: 19162202]
- Grant DA, Berg EA. A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology*. 1948; 38:404–411. DOI: 10.1037/h0059831 [PubMed: 18874598]
- Greve KW. The WCST-64: A standardized short-form of the Wisconsin Card Sorting Test. *The Clinical Neuropsychologist*. 2001; 15(2):228–234. DOI: 10.1076/clin.15.2.228.1901 [PubMed: 11528544]
- Ham TE, de Boissezon X, Leff A, Beckmann C, Hughes E, Kinnunen KM, Sharp DJ. Distinct frontal networks are involved in adapting to internally and externally signaled errors. *Cerebral Cortex*. 2013; 23(3):703–713. DOI: 10.1093/cercor/bhs056 [PubMed: 22426336]
- Hampshire A, Gruszka AFSJ, Owen AM. Inefficiency in self-organized attentional switching in the normal aging population is associated with decreased activity in the ventrolateral prefrontal cortex. *Journal of Cognitive Neuroscience*. 2008; 20(9):1670–1686. [PubMed: 18345987]
- Han G, Helm J, Iucha C, Zahn-Waxler C, Hastings PD, Klimes-Dougan B. Are executive functioning deficits concurrently and predictively associated with depressive and anxiety symptoms in adolescents? *Journal of Clinical Child & Adolescent Psychology*. 2016; 45(1):44–58. DOI: 10.1080/15374416.2015.1041592 [PubMed: 26042358]
- Hartman M, Bolton E, Fehnel SE. Accounting for age differences on the Wisconsin Card Sorting Test: Decreased working memory, not inflexibility. *Psychology and Aging*. 2001; 16(3):385–399. DOI: 10.1037/0882-7974.16.3.385 [PubMed: 11554518]
- Heaton, RK. A manual for the Wisconsin Card Sorting Test. Odessa, FL: Psychological Assessment Services; 1981.
- Holec V, Pirot HL, Euston DR. Not all effort is equal: The role of the anterior cingulate cortex in different forms of effort-reward decisions. *Frontiers in Behavioral Neuroscience*. 2014; 8:12.doi: 10.3389/fnbeh.2014.00012 [PubMed: 24478659]
- Holroyd CB, Umemoto A. The research domain criteria framework: The case for anterior cingulate cortex. *Neuroscience and Biobehavioral Reviews*. 2016; 71:418–443. DOI: 10.1016/j.neubiorev.2016.09.021 [PubMed: 27693229]
- Holroyd CB, Yeung N. Motivation of extended behaviors by anterior cingulate cortex. *Trends in Cognitive Sciences*. 2012; 16(2):122–128. DOI: 10.1016/j.tics.2011.12.008 [PubMed: 22226543]
- Kalikow DN, Stevens KN, Elliott LL. Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *The Journal of the Acoustical Society of America*. 1977; 61(5):1337–1351. [PubMed: 881487]
- Kanfer R, Ackerman PL, Heggstad ED. Motivational skills and self-regulation for learning: A trait perspective. *Learning and Individual Differences*. 1996; 8(3):185–209. DOI: 10.1016/S1041-6080(96)90014-X
- Kennerley SW, Walton ME, Behrens TE, Buckley MJ, Rushworth MF. Optimal decision making and the anterior cingulate cortex. *Nature Neuroscience*. 2006; 9(7):940–947. DOI: 10.1038/nn1724 [PubMed: 16783368]
- Kerns JG. Anterior cingulate and prefrontal cortex activity in an fMRI study of trial-to-trial adjustments on the Simon task. *NeuroImage*. 2006; 33(1):399–405. DOI: 10.1016/j.neuroimage.2006.06.012 [PubMed: 16876434]
- Kerns JG, Cohen JD, MacDonald AW 3rd, Cho RY, Stenger VA, Carter CS. Anterior cingulate conflict monitoring and adjustments in control. *Science*. 2004; 303:1023–1026. DOI: 10.1126/science.1089910 [PubMed: 14963333]
- Kongs, SK., Thompson, LL., Iverson, GL., Heaton, RK. Wisconsin Card Sorting Test—64 card version. Lutz, FL: Psychological Assessment Resources, Inc; 2000.

- Konishi S, Hayashi T, Uchida I, Kikyo H, Takahashi E, Miyashita Y. Hemispheric asymmetry in human lateral prefrontal cortex during cognitive set shifting. *Proceedings of the National Academy of Sciences*. 2002; 99(11):7803–7808. DOI: 10.1073/pnas.122644899
- Konishi S, Nakajima K, Uchida I, Kameyama M, Nakahara K, Sekihara K, Miyashita Y. Transient activation of inferior prefrontal cortex during cognitive set shifting. *Nature Neuroscience*. 1998; 1(1):80–84. DOI: 10.1038/283 [PubMed: 10195114]
- Kopp B, Lange F. Electrophysiological indicators of surprise and entropy in dynamic task-switching environments. *Frontiers in Human Neuroscience*. 2013; 7:300.doi: 10.3389/fnhum.2013.00300 [PubMed: 23840183]
- Kuchinsky SE, Ahlstrom JB, Vaden KI Jr, Cute SL, Humes LE, Dubno JR, Eckert MA. Pupil size varies with word listening and response selection difficulty in older adults with hearing loss. *Psychophysiology*. 2013; 50(1):23–34. DOI: 10.1111/j.1469-8986.2012.01477.x [PubMed: 23157603]
- Lange F, Kroger B, Steinke A, Seer C, Dengler R, Kopp B. Decomposing card-sorting performance: Effects of working memory load and age-related changes. *Neuropsychology*. 2016; 30(5):579–590. DOI: 10.1037/neu0000271 [PubMed: 26866348]
- Lao-Kaim NP, Fonville L, Giampietro VP, Williams SC, Simmons A, Tchanturia K. Aberrant function of learning and cognitive control networks underlie inefficient cognitive flexibility in anorexia nervosa: A cross-sectional fMRI study. *PLoS One*. 2015; 10(5):e0124027.doi: 10.1371/journal.pone.0124027 [PubMed: 25970523]
- Leth-Steensen C, King Elbaz Z, Douglas VI. Mean response times, variability, and skew in the responding of ADHD children: A response time distributional approach. *Acta Psychologica*. 2000; 104:167–190. [PubMed: 10900704]
- Leyro TM, Zvolensky MJ, Bernstein A. Distress tolerance and psychopathological symptoms and disorders: A review of the empirical literature among adults. *Psychological Bulletin*. 2010; 136(4): 576–600. DOI: 10.1037/a0019712 [PubMed: 20565169]
- Luks TL, Simpson GV, Dale CL, Hough MG. Preparatory allocation of attention and adjustments in conflict processing. *NeuroImage*. 2007; 35(2):949–958. DOI: 10.1016/j.neuroimage.2006.11.041 [PubMed: 17258912]
- Mars RB, Coles MG, Grol MJ, Holroyd CB, Nieuwenhuis S, Hulstijn W, Toni I. Neural dynamics of error processing in medial frontal cortex. *NeuroImage*. 2005; 28(4):1007–1013. DOI: 10.1016/j.neuroimage.2005.06.041 [PubMed: 16055352]
- Militello, LG., Gentner, FC., Swindler, SD., Beisner, GI. Paper presented at the International Symposium on Collaborative Technologies and Systems. Las Vegas, NV: 2006. Conation: Its historical roots and implications for future research.
- Milner B. Effects of different brain lesions on card sorting: The role of the frontal lobes. *Archives of Neurology*. 1963; 9(1):90–100. DOI: 10.1001/archneur.1963.00460070100010
- Miyake A, Friedman NP. The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*. 2012; 21(1):8–14. DOI: 10.1177/0963721411429458 [PubMed: 22773897]
- Nagahama Y, Sadato N, Yamauchi H, Katsumi Y, Hayashi T, Fukuyama H, Yonekura Y. Neural activity during attention shifts between object features. *NeuroReport*. 1998; 9:2633–2638. DOI: 10.1097/00001756-199808030-00038 [PubMed: 9721946]
- Nelson HE. A modified card sorting test sensitive to frontal lobe defects. *Cortex*. 1976; 12(4):313–324. DOI: 10.1016/S0010-9452(76)80035-4 [PubMed: 1009768]
- Neta M, Schlaggar BL, Petersen SE. Separable responses to error, ambiguity, and reaction time in cingulo-opercular task control regions. *NeuroImage*. 2014; 99:59–68. DOI: 10.1016/j.neuroimage.2014.05.053 [PubMed: 24887509]
- Nyhus E, Barcelo F. The Wisconsin Card Sorting Test and the cognitive assessment of prefrontal executive functions: A critical update. *Brain and Cognition*. 2009; 71(3):437–451. DOI: 10.1016/j.bandc.2009.03.005 [PubMed: 19375839]
- Onatsu-Arvilommi T, Nurmi JE. The role of task-avoidant and task-focused behaviors in the development of reading and mathematical skills during the first school year: A cross-lagged



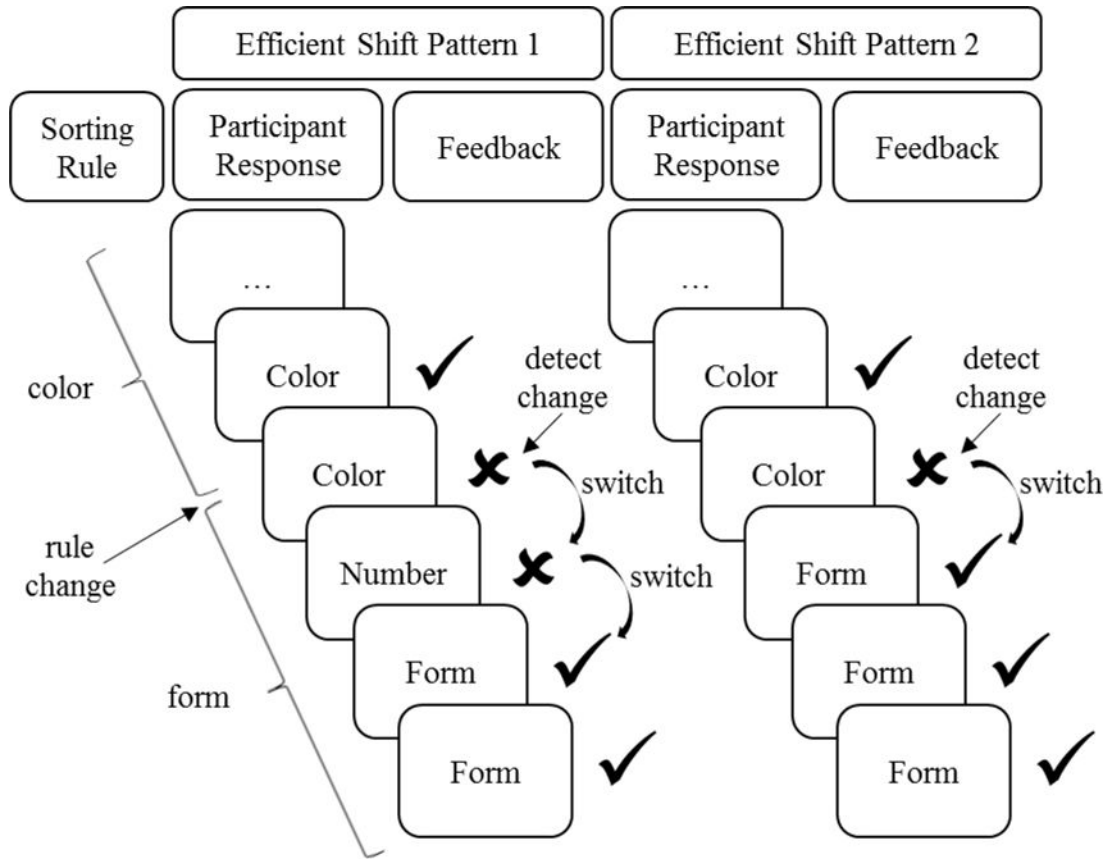
- longitudinal study. *Journal of Educational Psychology*. 2000; 92(3):478–491. DOI: 10.1037/0022-0663.92.3.478
- Ord JS, Boettcher AC, Greve KW, Bianchini KJ. Detection of malingering in mild traumatic brain injury with the Conners' Continuous Performance Test-II. *Journal of Clinical and Experimental Neuropsychology*. 2010; 32(4):380–387. DOI: 10.1080/13803390903066881 [PubMed: 19739010]
- Orr JM, Weissman DH. Anterior cingulate cortex makes 2 contributions to minimizing distraction. *Cerebral Cortex*. 2009; 19(3):703–711. DOI: 10.1093/cercor/bhn119 [PubMed: 18653665]
- Pauls F, Petermann F, Lepach AC. Memory assessment and depression: Testing for factor structure and measurement invariance of the Wechsler Memory Scale-fourth edition across a clinical and matched control sample. *Journal of Clinical and Experimental Neuropsychology*. 2013; 35(7):702–717. DOI: 10.1080/13803395.2013.820257 [PubMed: 23886016]
- Pintrich, PR., Smith, DAF., Garcia, T., McKeachie, WJ. A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ). Ann Arbor, MI: The University of Michigan; 1991.
- Pintrich PR, Smith DAF, Garcia T, McKeachie WJ. Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (MSLQ). *Educational and Psychological Measurement*. 1993; 53(3):801–813. DOI: 10.1177/0013164493053003024
- Potter GG, Blackwell AD, McQuoid DR, Payne ME, Steffens DC, Sahakian BJ, Krishnan KR. Prefrontal white matter lesions and prefrontal task impersistence in depressed and nondepressed elders. *Neuropsychopharmacology*. 2007; 32(10):2135–2142. DOI: 10.1038/sj.npp.1301339 [PubMed: 17299509]
- Price LR, Tulskey D, Millis S, Weiss L. Redefining the factor structure of the Wechsler Memory Scale-III: Confirmatory factor analysis with cross-validation. *Journal of Clinical and Experimental Neuropsychology*. 2002; 24(5):574–585. [PubMed: 12187442]
- Quinn EP, Brandon TH, Copeland AL. Is task persistence related to smoking and substance abuse? The application of learned industriousness theory to addictive behaviors. *Experimental and Clinical Psychopharmacology*. 1996; 4(2):186–190. DOI: 10.1037/1064-1297.4.2.186
- R Core Team. R: A language and environment for statistical computing (Version 2.15.0). Vienna, Austria: R Foundation for Statistical Computing; 2012. Retrieved from <https://www.R-project.org/>
- R Core Team. R: A language and environment for statistical computing (Version 3.3.1). Vienna, Austria: R Foundation for Statistical Computing; 2016. Retrieved from <https://www.R-project.org/>
- Ravizza SM, Carter CS. Shifting set about task switching: Behavioral and neural evidence for distinct forms of cognitive flexibility. *Neuropsychologia*. 2008; 46(12):2924–2935. DOI: 10.1016/j.neuropsychologia.2008.06.006 [PubMed: 18601941]
- Ravizza SM, Delgado MR. Motivational enhancement of cognitive control depends on depressive symptoms. *Emotion*. 2014; 14(4):646–650. DOI: 10.1037/a0036754 [PubMed: 24866522]
- Ridderinkhof KR, Ullsperger M, Crone EA, Nieuwenhuis S. The role of the medial frontal cortex in cognitive control. *Science*. 2004; 306:443–447. DOI: 10.1126/science.1100301 [PubMed: 15486290]
- Robinson LJ, Kester DB, Saykin AJ, Kaplan EF, Gur RC. Comparison of two short forms of the Wisconsin Card Sorting Test. *Archives of Clinical Neuropsychology*. 1991; 6(1–2):27–33. DOI: 10.1016/0887-6177(91)90019-6 [PubMed: 14589597]
- Shenhav A, Botvinick MM, Cohen JD. The expected value of control: An integrative theory of anterior cingulate cortex function. *Neuron*. 2013; 79(2):217–240. DOI: 10.1016/j.neuron.2013.07.007 [PubMed: 23889930]
- Shenhav A, Cohen JD, Botvinick MM. Dorsal anterior cingulate cortex and the value of control. *Nature Neuroscience*. 2016; 19(10):1286–1291. DOI: 10.1038/nn.4384 [PubMed: 27669989]
- Shenhav A, Musslick S, Lieder F, Kool W, Griffiths TL, Cohen JD, Botvinick MM. Toward a rational and mechanistic account of mental effort. *Annual Review of Neuroscience*. 2017; 40doi: 10.1146/annurev-neuro-072116-031526
- Snow RE. Self-regulation as meta-cognition? *Learning and Individual Differences*. 1996; 8(3):261–267. DOI: 10.1016/S1041-6080(96)90017-5

- Steinberg ML, Krejci JA, Collett K, Brandon TH, Ziedonis DM, Chen K. Relationship between self-reported task persistence and history of quitting smoking, plans for quitting smoking, and current smoking status in adolescents. *Addictive Behaviors*. 2007; 32(7):1451–1460. DOI: 10.1016/j.addbeh.2006.10.008 [PubMed: 17125931]
- Steinberg ML, Williams JM, Gandhi KK, Foulds J, Brandon TH. Lower task persistence in smokers with schizophrenia as compared to non-psychiatric control smokers. *Psychology of Addictive Behaviors*. 2010; 24(4):724–729. DOI: 10.1037/a0020972 [PubMed: 21198231]
- Steinberg ML, Williams JM, Gandhi KK, Foulds J, Epstein EE, Brandon TH. Task persistence predicts smoking cessation in smokers with and without schizophrenia. *Psychology of Addictive Behaviors*. 2012; 26(4):850–858. DOI: 10.1037/a0028375 [PubMed: 22642860]
- Steinmetz JP, Houssemand C. What about inhibition in the Wisconsin Card Sorting Test? *The Clinical Neuropsychologist*. 2011; 25(4):652–669. DOI: 10.1080/13854046.2011.568525 [PubMed: 21547856]
- Stuss DT, Levine B, Alexander MP, Hong J, Palumbo C, Hamer L, Izukawa D. Wisconsin Card Sorting Test performance in patients with focal frontal and posterior brain damage: Effects of lesion location and test structure on separable cognitive processes. *Neuropsychologia*. 2000; 38:388–402. DOI: 10.1016/S0028-3932(99)00093-7 [PubMed: 10683390]
- Suhr JA, Sullivan BK, Rodriguez JL. The relationship of noncredible performance to continuous performance test scores in adults referred for attention-deficit/hyperactivity disorder evaluation. *Archives of Clinical Neuropsychology*. 2011; 26(1):1–7. DOI: 10.1093/arclin/acq094 [PubMed: 21159792]
- Tas C, Brown EC, Esen-Danaci A, Lysaker PH, Brune M. Intrinsic motivation and metacognition as predictors of learning potential in patients with remitted schizophrenia. *Journal of Psychiatric Research*. 2012; 46(8):1086–1092. DOI: 10.1016/j.jpsychires.2012.04.027 [PubMed: 22608773]
- Tow PM, Whitty CWM. Personality changes after operations on the cingulate gyrus in man. *Journal of Neurology, Neurosurgery, and Psychiatry*. 1953; 16:186–193. DOI: 10.1136/jnnp.16.3.186
- Vaden KI Jr, Kuchinsky SE, Ahlstrom JB, Dubno JR, Eckert MA. Cortical activity predicts which older adults recognize speech in noise and when. *The Journal of Neuroscience*. 2015; 35(9):3929–3937. DOI: 10.1523/JNEUROSCI.2908-14.2015 [PubMed: 25740521]
- Vaden KI Jr, Kuchinsky SE, Ahlstrom JB, Teubner-Rhodes SE, Dubno JR, Eckert MA. Cingulo-opercular function during word recognition in noise for older adults with hearing loss. *Experimental Aging Research*. 2016; 42(1):67–82. DOI: 10.1080/0361073X.2016.1108784 [PubMed: 26683042]
- Vaden KI Jr, Kuchinsky SE, Cute SL, Ahlstrom JB, Dubno JR, Eckert MA. The cingulo-opercular network provides word-recognition benefit. *The Journal of Neuroscience*. 2013; 33(48):18979–18986. DOI: 10.1523/JNEUROSCI.1417-13.2013 [PubMed: 24285902]
- van Veen V, Holroyd CB, Cohen JD, Stenger VA, Carter CS. Errors without conflict: Implications for performance monitoring theories of anterior cingulate cortex. *Brain and Cognition*. 2004; 56(2):267–276. DOI: 10.1016/j.bandc.2004.06.007 [PubMed: 15518940]
- Ventura M, Shute V, Zhao W. The relationship between video game use and a performance-based measure of persistence. *Computers & Education*. 2013; 60(1):52–58. DOI: 10.1016/j.compedu.2012.07.003
- Verguts T, Vassena E, Silvetti M. Adaptive effort investment in cognitive and physical tasks: A neurocomputational model. *Frontiers in Behavioral Neuroscience*. 2015; 9:57. doi: 10.3389/fnbeh.2015.00057 [PubMed: 25805978]
- Wang J, Cao B, Cai X, Gao H, Li F. Brain activation of negative feedback in rule acquisition revealed in a segmented Wisconsin Card Sorting Test. *PloS One*. 2015; 10(10):e0140731. doi: 10.1371/journal.pone.0140731 [PubMed: 26469519]
- Wechsler D. Cognitive, conative, and non-intellective intelligence. *The American Psychologist*. 1950; 5:78–83. DOI: 10.1037/h0063112
- Wechsler, D. Wechsler Memory Scale®—third edition (WMS-III) administration and scoring manual. San Antonio, TX: The Psychological Corporation®.; 1997.
- Wechsler, D. Wechsler Memory Scale®—third edition abbreviated (WMS-III abbreviated) manual. San Antonio, TX: The Psychological Corporation®.; 2002.

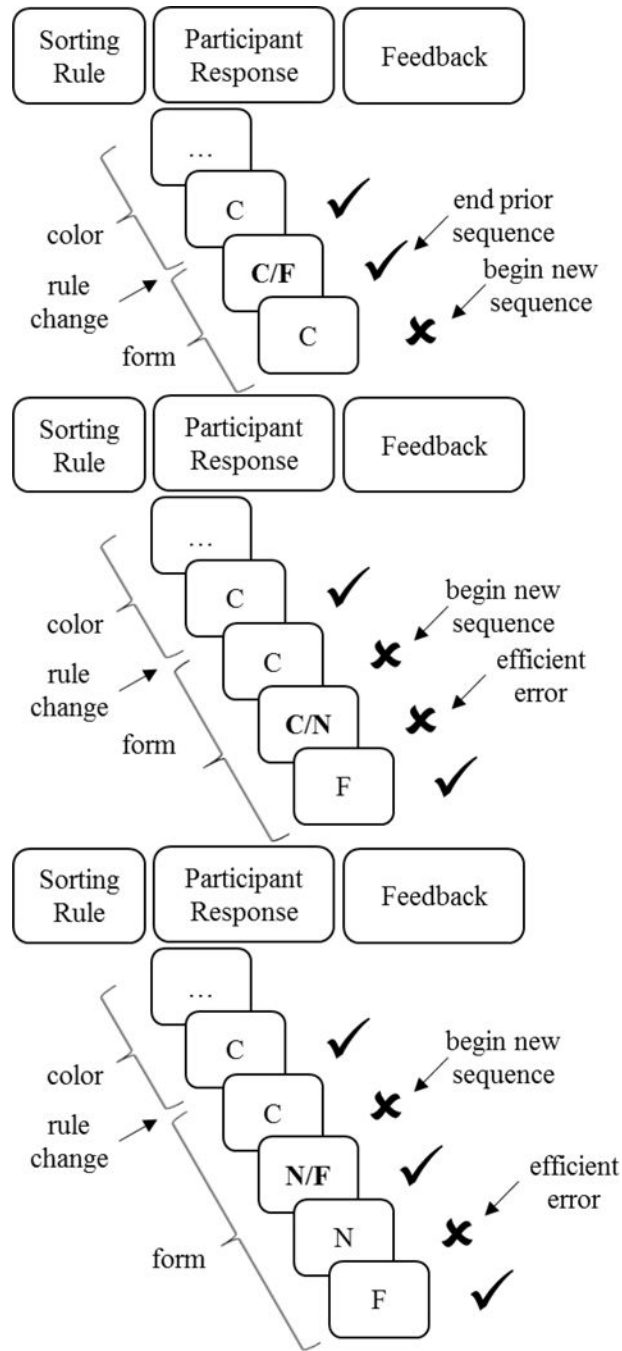
- Weissman DH, Roberts KC, Visscher KM, Woldorff MG. The neural bases of momentary lapses in attention. *Nature Neuroscience*. 2006; 9(7):971–978. DOI: 10.1038/nn1727 [PubMed: 16767087]
- Weissman DH, Warner LM, Woldorff MG. Momentary reductions of attention permit greater processing of irrelevant stimuli. *NeuroImage*. 2009; 48(3):609–615. DOI: 10.1016/j.neuroimage.2009.06.081 [PubMed: 19596451]
- Wolters CA, Yu SL, Pintrich PR. The relation between goal orientation and students' motivational beliefs and self-regulated learning. *Learning and Individual Differences*. 1996; 8(3):211–238. DOI: 10.1016/S1041-6080(96)90015-1
- Yeung N, Botvinick MM, Cohen JD. The neural basis of error detection: Conflict monitoring and the error-related negativity. *Psychological Review*. 2004; 111(4):931–959. DOI: 10.1037/0033-295x.111.4.931 [PubMed: 15482068]
- Zekveld AA, Kramer SE, Festen JM. Cognitive load during speech perception in noise: The influence of age, hearing loss, and cognition on the pupil response. *Ear & Hearing*. 2011; 32(4):498–510. [PubMed: 21233711]
- Zhang X, Nurmi JE, Kiuru N, Lerkkanen MK, Aunola K. A teacher-report measure of children's task-avoidant behavior: A validation study of the behavioral strategy rating scale. *Learning and Individual Differences*. 2011; 21(6):690–698. DOI: 10.1016/j.lindif.2011.09.007

### Highlights

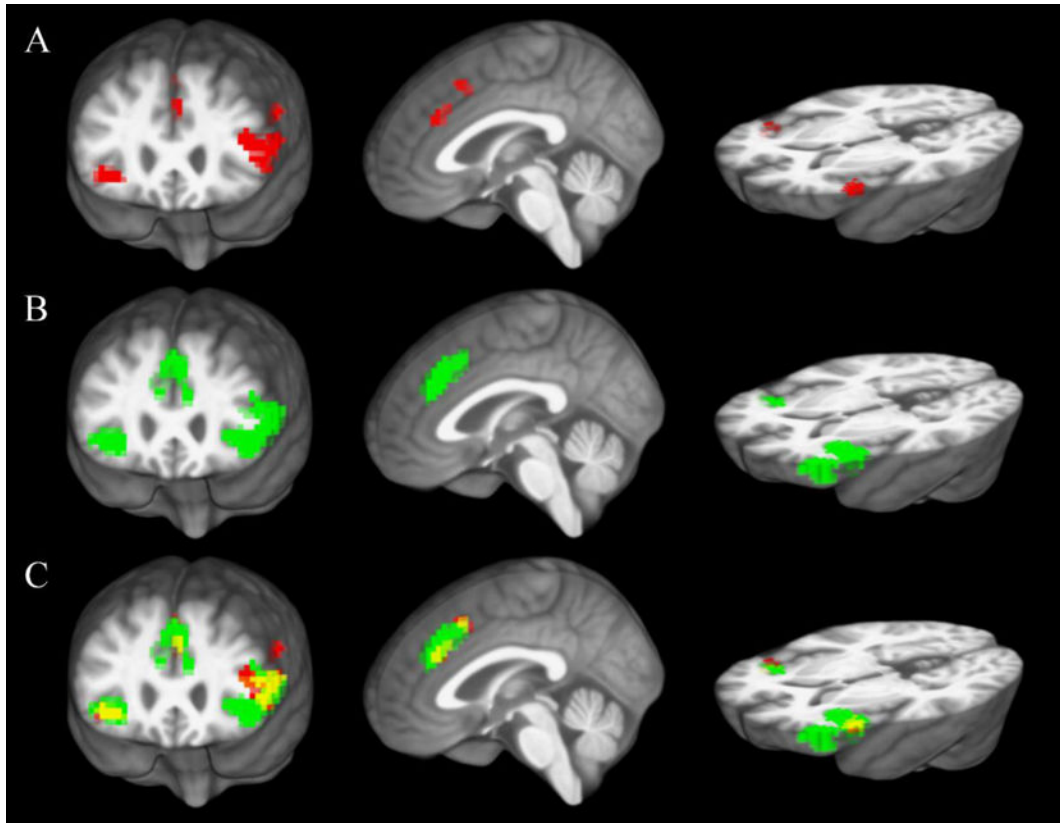
- The Wisconsin Card Sorting Test can measure persistence to overcome task difficulty
- Persistence predicts cingulo-opercular error responses and performance benefits
- The new measure can dissociate set-shifting and persistence



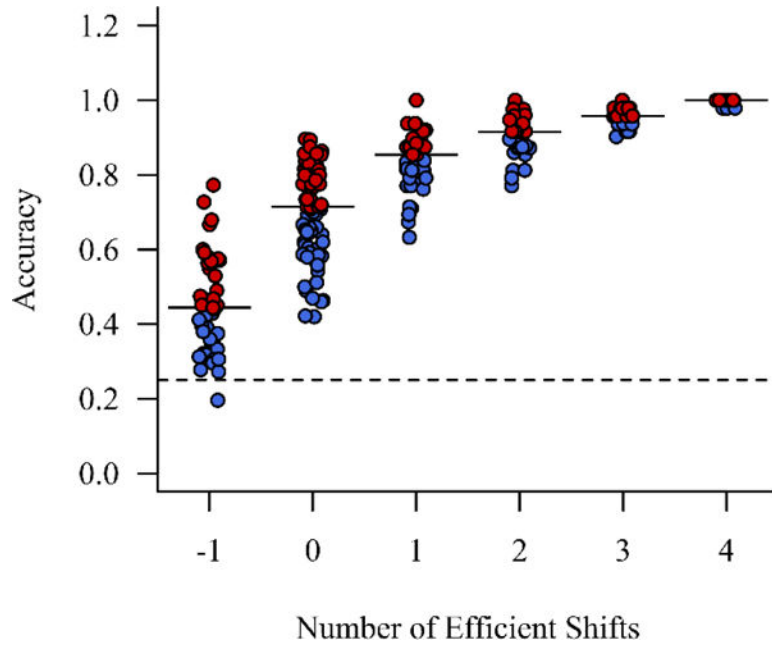
**Figure 1.** Diagram of two possible efficient shift trial sequences in which the correct sorting rule changed from color to form. ✓: correct response; ✗: incorrect response. Participants can first detect a rule change upon receiving negative feedback for using the previous sorting rule. Following detection of a rule change, participants performing optimally will switch to one of the two remaining rules, which could be the incorrect rule (Efficient Shift Pattern 1) or the correct rule (Efficient Shift Pattern 2). An efficient error occurs when the participant switches to the wrong rule but then switches to and keeps using the right rule (e.g., the “Number” response in Efficient Shift Pattern 1). Both patterns shown in the diagram are expected as part of the optimal performance strategy after the sorting rule changes.



**Figure 2.** Special cases arising from participant responses that match on multiple dimensions (indicated in bold). C: color response; F: form response; N: number response. Top panel: The multidimensional response shifts the detection of a rule change, and thus the start of a new sequence, by one trial. Middle panel: The multidimensional response is scored as an efficient error, because the participant subsequently switches to the correct rule. Bottom panel: The multidimensional response shifts the opportunity to determine the new sorting rule by one trial, resulting in an efficient error later in the sequence.

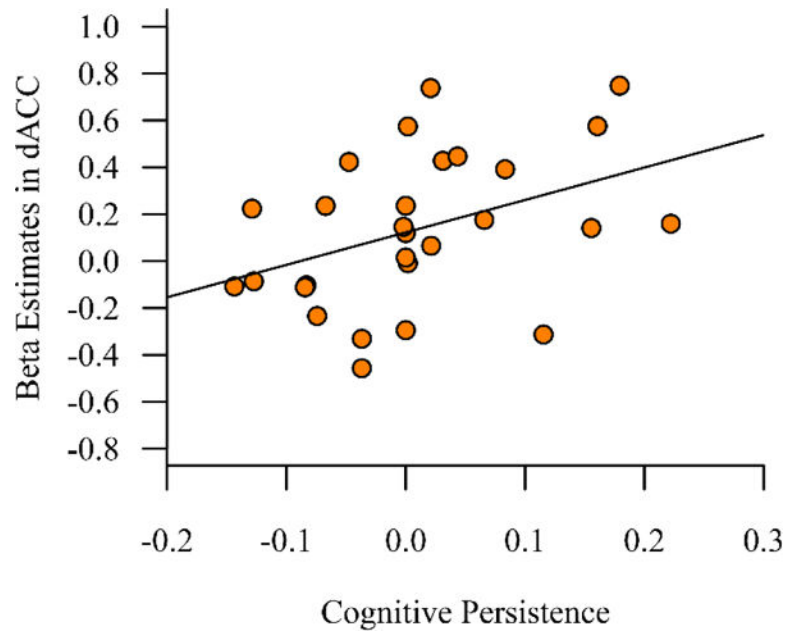


**Figure 3.** Cingulo-opercular ROIs for (A) adaptive-control effects (red) and (B) error-responses (green). Their overlap (yellow) is shown in (C). ROIs were defined from an independent sample of 18 younger adults and warped into study-specific space.



**Figure 4.** The relationship between number of efficient shifts and accuracy on the WCST. Solid horizontal lines indicate the median for each category of efficient shifts. The dashed line indicates chance accuracy (25%). Individuals who fall above the median (red) have positive accuracy differences and consequently positive persistence values. Individuals who fall below the median (blue) have negative accuracy differences and consequently negative persistence values. For visualization purposes, data have been randomly jittered  $\pm 0.1$  along the x-axis to reduce overlap. Persistence scores of exactly 0 are plotted in red.





**Figure 5.** Middle-aged to older adults with higher cognitive persistence demonstrated significantly larger adaptive-control beta estimates in the dACC.

**Table 1**

## Scoring of Different Error Types

	<i>Definition</i>	<i>Scoring</i>	<i>Rationale</i>
Detection errors	First trial on which participants could detect a rule change, i.e., the first negative feedback following a rule change (Figure 1)	Detection errors were excluded from analyses	Detection errors are necessary to find out that the sorting rule has changed
Efficient errors	Incorrect trials that resulted from switching to the wrong rule immediately following the detection of a rule change and led to correct performance for all subsequent trials in the sequence (Figure 1)	Efficient errors were counted as correct	Efficient errors reflect optimal switching behavior
Other errors	All other errors, including repeating the previous rule following the detection of a rule change (i.e., perseverations) and changing rules but making other errors later in the same sequence	Other errors were counted as incorrect	Other errors reflect sub-optimal switching behavior

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 2**

## Scoring of Multidimensional Responses

	<i>Description</i>	<i>Scoring</i>	<i>Rationale</i>
Case 1	Following a rule change, participants can sort correctly by using the previous rule if the card matches both the previous and the current rule (Figure 2, top panel)	For these trials, the first incidence of negative feedback following the rule change was considered the start of the new sequence	Participants have no prior indication that the rule has changed
Case 2	Following the first negative feedback that indicates a rule change, participants can sort the card such that it matches on both the previous rule and a new rule (Figure 2, middle panel)	These responses were considered efficient if participants engaged in the optimal strategy for the remainder of the sequence, i.e., continuing to use the new rule if correct and switching to the remaining rule if incorrect	This pattern of responses is consistent with optimal switching behavior
Case 3	Participants can switch to the incorrect rule following the detection of a rule change, but receive correct feedback because the card also matches the new correct rule (Figure 2, bottom panel); this may result in a later error because the participant continues to sort on the new but incorrect rule	These later errors were considered efficient if all subsequent trials in the sequence were correct	This pattern of responses is consistent with optimal switching behavior
Case 4	Participants can receive the first negative feedback following a rule change for a response that matched on both the previous rule and the other incorrect rule	These were interpreted as responses based on the previous rule only, and efficient shifts were scored as usual for switches to either of the remaining rules	This pattern of responses is consistent with optimal switching behavior

**Table 3**

Expected accuracy (median) by number of efficient shifts

NES	Bootstrap Median Estimates			
	Expected Accuracy	Bias	SE	95% CI
-1	0.4444	-0.0046	0.0214	[0.3800, 0.4643]
0	0.7143	-0.0057	0.0176	[0.6596, 0.7347]
1	0.8541	-0.0058	0.0170	[0.8158, 0.8750]
2	0.9149	-0.0039	0.0103	[0.8750, 0.9167]
3	0.9574	-0.0018	0.0066	[0.9375, 0.9574]
4	1.0000	-0.0010	0.0045	[0.9797, 1.0000]

Note. NES = number of efficient shifts.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 4**

Correlations [95% CIs] between Demographic Variables and WCST Metrics

	1	2	3	4
1. Age	—			
2. Education <sup>a</sup>	0.01 [-0.13, 0.16]	—		
3. NES	-0.49 [-0.59, -0.37]*	0.30 [0.17, 0.41]*	—	
4. Persistence	-0.04 [-0.15, 0.07]	0.07 [-0.09, 0.23]	0.01 [-0.13, 0.09]	—
<i>M</i>				
Female	59.86	15.70	0.94	-0.01
Male	58.16	15.51	0.80	-0.00
<i>M difference</i> <sup>b</sup>	1.71 [-3.04, 6.45]	0.19 [-0.57, 0.93]	0.14 [-0.23, 0.52]	-0.01 [-0.04, 0.02]

Note.

Education: years of education; NES: number of efficient shifts.

<sup>a</sup>Education data were available for 176 participants.<sup>b</sup>Studentized 95% CIs as recommended by Canty and Ripley (2016).

\* 95% CIs estimated from non-parametric bootstrap resampling did not include 0.

**Table 5**

Correlations between WCST Metrics and Behavioral Validation Variables [95% CIs]

	Number of Efficient Shifts	Cognitive Persistence
<i>Abstract Reasoning</i>		
TCFC <sup>a</sup>	-0.33 [-0.41, -0.24] *	0.09 [-0.07, 0.25]
<i>Working Memory</i>		
WMS-aud <sup>b</sup>	0.38 [0.22, 0.51] *	0.01 [-0.15, 0.16]
WMS-vis <sup>b</sup>	0.34 [0.17, 0.48] *	0.04 [-0.13, 0.22]
<i>Performance Consistency</i>		
CPT SE <sup>c</sup>	-0.13 [-0.31, 0.08]	-0.22 [-0.40, -0.03] *

Note.

TCFC: trials to complete first category. WMS-aud: Wechsler Memory Scale auditory subtest. WMS-vis: Wechsler Memory Scale visual subtest. CPT SE: Continuous Performance Test standard error of hit reaction time.

<sup>a</sup>Higher TCFC indicates poorer abstract reasoning.

<sup>b</sup>WMS data available for 135 participants.

<sup>c</sup>CPT data available for 108 participants. Higher CPT SE indicates less consistent performance.

\* 95% CIs estimated from non-parametric bootstrap resampling did not include 0.