In this article, we describe results of the 5th Cognitive Neuroscience Treatment Research to Improve Cognition in Schizophrenia meeting which identified candidate imaging biomarkers for use in measuring neural activity associated with specific component processes of cognition that are targeted for treatment development in schizophrenia and other disorders. This manuscript describes the process by which measures related to executive control were selected, along with the specific measures recommended for further development. Two paradigms were recommended for measurement of the cognitive and neural mechanisms underlying 2 core component processes of executive control, rule generation and selection, and dynamic adjustments of Control. The 2 paradigms are the AX continuous performance task (letter and dot forms), implemented as an functional magnetic resonance imaging (fMRI) paradigm to engage neural systems supporting rule generation and selection, and the switching Stroop task, implemented as either fMRI or electroencephalography that may be used as a measure of both rule generation and selection as well as dynamic adjustment in control. A detailed description of each paradigm, together with a review of the relevant literature related to their cognitive and neural validity and measurement properties is provided. These 2 paradigms are recommended for further development, including further validation at the cognitive and neural level and optimization with respect to subject tolerability, psychometric, and neurometric features.

Key words: CNTRICS/executive functions/imaging biomarker/AX CPT/Stroop/schizophrenia/cognition/ treatment/cognitive neuroscience

Executive function deficits are among the most prominent cognitive impairments in schizophrenia. During the first Cognitive Neuroscience Treatment Research to Improve Cognition in Schizophrenia (CNTRICS) meeting in February 2007 in Bethesda, MD, 2 aspects of executive function, rule generation and selection, and dynamic adjustments in cognitive control were recommended to be targeted for treatment development for impaired cognition in schizophrenia. These constructs were targeted because they met a number of criteria, including being readily measured in humans, showing consistent evidence of impairment in schizophrenia, having well-established links to known neural systems at the macro circuit level and in some cases, local circuit and neurotransmitter systems and the potential availability of homologous animal model systems.

During the third CNTRICS meeting, held in March 2008 in Sacramento, California, a set of behavioral measures engaging these 2 constructs were recommended for further development. They included the 1–2 AX continuous performance task (CPT) (a version of the task described below which requires subject to maintain and switch between multiple response rules) and the intradimensional-extradimensional set-shifting task as measures of rule generation and selection and the Stroop task and the stop signal tasks as measures of dynamic adjustments in control. Since executive functions and other cognitive systems including working memory are not orthogonal but considered to have overlapping component processes, it should be noted that a working memory-related construct, goal maintenance, operationally identical to rule generation and selection was identified, and at the third meeting, the AX CPT task was recommended for further development for this construct.

The identification of these behavioral paradigms, using tasks that have been validated and widely used in basic cognitive neuroscience, was an important outcome of the consensus-based CNTRICS process. Subsequent to this meeting, a subset of paradigms has been moved along the necessary developmental pathway that involves optimizing the measure for efficient, standardized administration, validation of its sensitivity to patient control differences, effect size estimation, and characterizing and optimizing its measurement properties (see figure 1, Carter et al this volume). These paradigms are available for
In the present article, we describe the results of the 5th CNTRICS consensus-based meeting, held in Davis, California, in October 2009. At this meeting, we identified a set of neural systems based imaging biomarkers for use in the treatment development process. Such experimental medicine tools may enhance translational research, particularly in early phase studies, by providing evidence of pharmacodynamic activity in the brain, proof of mechanism in human subjects, and predictors of individual treatment response in the context of personalized medicine. During a previous meeting held in Baltimore, MD, in October 2009, we examined the known biological basis of the signals generated by widely available functional imaging methodologies and related measures, the complexities, practicalities and potential confounds associated with their use in treatment development research, and methodologies for characterizing their measurement properties. During the present meeting, we evaluated a number of paradigms nominated as potential imaging biomarkers related to executive control, using criteria developed on the basis of a web-based survey. These criteria are described in figures 2 and 3 of the overview article by Carter et al in this volume and include cognitive and neural construct validity and sensitivity to neural systems deficits as well as factors such as test retest reliability and practicality of administration. In the present article, we will describe the selection of the most promising event related potentials (ERP)/electroencephalography (EEG) or functional magnetic resonance imaging (fMRI) measures.
Potential imaging biomarkers focusing on the neural systems supporting 2 component processes supporting executive control were targeted for measurement development, following the recommendations of the first CNTRICS meeting. These component processes were (1) “rule generation and selection” and (2) “dynamic adjustments of control”.

“Rule generation and selection” was operationalized as “processes involved in activating task-related goals or rules based on endogenous or exogenous cues, actively representing them in a highly accessible form, and maintaining this information over an interval during which that information is needed to bias and constrain attention and response selection.” “Dynamic adjustments of control” was operationalized as “processes involved in detecting the occurrence of conflict or errors in ongoing processing, identifying the type of control adjustments needed, and recruiting additional control processes.”

Paradigms proposed as potential imaging biomarkers were nominated via an online survey and presented and discussed in the breakout groups and by the entire group at the meeting. The discussion focused on paradigms for engaging functional brain circuitry using fMRI and EEG/ERP’s with the goal of identifying 1–2 paradigms per construct.

Imaging paradigms nominated for rule generation and selection included versions of the expectancy AX CPT (fMRI), switching Stroop task (fMRI and ERP), the lateralized readiness potential (ERP), probabilistic reversal learning (fMRI), dual task performance (fMRI), and an analogue of the AX CPT, the preparing to overcome prepotency (POP) task.4

Imaging paradigms nominated to engage the neural systems supporting the dynamic adjustments on control were the switching Stroop task (fMRI/ERP), visual search (ERP), probabilistic reversal learning, (fMRI) the suppress task (fMRI), the flanker task (ERP, fMRI), and the spatial-delayed response task (fMRI).

Using the criteria outlined in figures 2 and 3 of the overview article in the present volume, with a strong weighting on “construct validity” and the demonstrated ability to engage the functional brain circuitry using fMRI or EEG/ERP, the following paradigms were recommended for further development. The primary reason for paradigms not being recommended was limited construct validity. For rule generation and selection, the AX CPT/dot pattern expectancy (DPX) was selected for use during fMRI and the switching Stroop task for use during both fMRI and EEG/ERP. For dynamic adjustments in control, the switching Stroop task was selected. The ability of this paradigm to provide behavioral and neural activity measures for both rule selection and dynamic adjustments, and its amenability to use in fMRI studies as well as ERP studies was noted as a particular strength. These 2 paradigms will be now described in detail, along with the properties that led to their being recommended.

**Expectancy AX and DPX Tasks**

**Background and Description of the Paradigm**

The A-then-X rule of the classical AX CPT, that “X is the target when preceded by an A,” is associated with Rosvold colleagues. Cohen and Servan-Schreiber increased participants’ “expectancy” of AX trials to measure participants’ capacity to use cue information to maintain a goal for the purpose of controlling a prepotent response. Whereas the classical AX CPT generally used a single button to signal the occurrence of rare AX sequences, the expectancy AX task involved 70% AX sequences and included 2 buttons to indicate targets and nontargets. The nontarget button facilitated correct rejection and reaction time analyses for rare sequences, such as BX trials. Given the expectancy manipulation, BX trials are important because they test the capacity of the participant to use their knowledge about the invalid cue (“B” or any other non-“A” cue) to overcome a habitual response to make a target response to the usually valid “X” that follows. Thus, the BX condition is sensitive to the “goal representation and maintenance” needed to make a nontarget response. Another rare sequence AY is the most challenging condition for
people with good goal maintenance because an A cue primes the target response to an expected X. A more thorough description of the logic underlying the analyses of these exception trial types, and how they can facilitate the interpretation of a specific deficits in patients with schizophrenia has been reviewed elsewhere. The DPX task is formally equivalent to the expectancy AX except that it uses underlearned dot patterns rather than letters. This can have 3 advantages: (1) the use of underlearned dot patterns makes representation of the cue challenging even over a short delay, potentially increasing the time efficiency of the task; (2) the configuration of dots allows manipulation of the similarity of stimuli thereby increasing the difficulty of AY trials for most people, potentially increasing interpretability; and (3) the proportions of critical AY and BX trials is 12.5% (rather than 10%) to provide more critical trials. For the current purposes, both these tasks will be considered exemplars of the AX paradigm. The primary dependent measures are error rates and reaction times across different conditions, as well as a signal detection measure known as d'-context. D'-context is calculated as the difference between the z-transformed proportion of AX hits and BX false positives, with a small constant used in the case of perfect performance on either condition.

**Neural Systems**

The AX paradigm has been used extensively during fMRI, reliably activating the brain networks supporting goal maintenance. Two contrasts have been informative: the contrast of long vs short delay, wherein trials that required any cue-related information had to be held for a longer period in working memory were compared with trials when this representation could be used immediately. This contrast affords information beyond what seen in typical working memory tasks since it is more clearly linked to the maintenance and updating of rule-based information, which is typically fixed across different working memory conditions. The second contrasts trial with B vs A cues, in which trials with high demands for goal maintenance to control behavior (B-cue trials) were compared with those in which the prepotent response might be expected (A-cue trials). In a meta-analysis, Minzenberg and colleagues reported that across studies in controls, it was associated with reliable increases in activity in bilateral dorsolateral prefrontal cortex (DLPFC) (BA 9 and other regions, including BA, 2, 3, 6, and 13).

The relationships between individual differences in brain activity and individual differences in performance are generally consistent and informative with subjects who show more DLPFC activation showing evidence of better task performance. Recently, Edwards and colleagues reported that among patients who received focused attentional training, improvement in right BA 9 was associated with improvement in relative BX reaction times.

**Results in Schizophrenia**

A number of studies have demonstrated that the AX paradigm is sensitive to schizophrenia patients' specific deficits in cognitive control (summarized in ref. Pharmacology and Animal Models

Two studies have now examined the effects of pharmacologic challenges in healthy controls expectancy AX performance. The first reported that the NMDA antagonist ketamine decreased d'-context and increased BX errors in healthy volunteers in a pattern reminiscent of that seen in patients with schizophrenia. In this study, AX errors were also increased to a degree far greater than that seen in patients. A second study tested a variant of the AX that includes distracting stimuli between the cue and the probe to approximate patient performance in 12 healthy participants. Whereas the interfering stimuli impaired d'-context, a low dose amphetamine rescued AX and BX performance.

One study performed intracranial recording in macaques performing a variation of expectancy AX task in which they responded only following valid AX trials. Most recording sites in DLPFC showed greater activity following B relative to A cues in a manner consistent with the human neuroimaging literature.

**Construct Validity and Measurement Properties**

There is a reasonable literature supporting the utility of the AX and DPX based upon their psychometric properties, which are in the acceptable range and their practicality. Increasingly, the focus of this work has turned to something referred to as neurometrics or the reliability of the brain activity measure. The neurometric properties of the DPX, including retest reliability and evidence for floor and ceiling effects, is currently being evaluated in a multisite study.

**Switching Stroop Task**

**Background and Description of the Paradigm**

This paradigm, which may be implemented as a measure of neural activity using either ERP’s or fMRI, was selected as a valid measure of both constructs, rule generation and selection, and the dynamic adjustment of control. The switching Stroop task represents an extension of the classic color-word Stroop task. In the task, color-words (eg, RED presented in red or BLUE
presented in green) are presented on each trial and individuals are asked to identify either the color or the word based upon a cue that precedes stimulus onset (ie, exogenous cueing)\textsuperscript{24} or a rule that is established at the beginning of a block of trials (ie, endogenous cueing). In the individual stimulus version of the Stroop task, 2 or 3 types of trials are typically presented. For congruent trials, a color-word is presented in a matching color (eg, RED presented in red); for neutral trials, a letter string (eg, XXX) or noncolor-word (eg, DOG) is presented in a task relevant color; and for incongruent trials, a color-word is presented in an opposing color (eg, RED presented in blue). Continued debate over what type of stimuli represent a “neutral” trial has led some investigators to focus exclusively upon differences in behavior or neural recruitment between congruent and incongruent. Within the current discussion, this approach also has the advantage of reducing the total number of trials required to obtain stable measures of performance or neural activity.

In most studies utilizing the switching Stroop, a cue is presented before the onset of the color-word indicating which of the 2 dimensions is relevant to task performance on that trial which engages processes related to rule generation and selection\textsuperscript{24,25} The task may include also 2 types of blocks (ie, pure and mixed) that allow the investigator to localize the contribution of various cognitive processes including rule generation and selection and dynamic control, respectively. In pure blocks, one of the stimulus dimensions is relevant for the entire block and the same cue is presented for each trial. This version lends itself to investigating dynamic control. In mixed blocks, the cue indicating the relevant stimulus may change from trial-to-trial resulting allowing investigation of rule generation and task switching. Repetitions represent trials where the same cue occurs on consecutive trials, therefore, these trials do not require a task switch; alternations represent trials were the cue differs from that presented for the previous trial, therefore, these trials require a task switch. Differences in behavior or neural recruitment between pure trials and repetitions are generally described as mixing costs, and differences between repetitions and alternations are described as switching costs. In studies where multiple cues are associated with each task, alternations can be further divided into cue alternations (ie, trials where the cue but not the task changes across trials) and task alternations (ie, trials were the cue and task change across trials). This manipulation allows one to determine whether differences between repetitions and alternations reflect processes related to retrieval of the cue-task association from memory (ie, repetitions vs cue alternations) or processes related to task set configuration (ie, cue alternations vs task alternations).

The assessment of dynamic adjustments in control has generally used pure or nonmixed blocks of Stroop trials in which subjects color name throughout the block and, as in the mixed version of the task, trials may be incongruent or congruent. Dynamic adjustments in control, reflected by faster and more accurate responding for incongruent trials following incongruent trials and slower and less accurate responding for incongruent trials follow congruent trials (with comparable effects for congruent trial performance also), are reliably seen with this task in the absence of repetition priming effects.\textsuperscript{26,27} While these adjustments are widely believed to be driven by conflict monitoring mechanisms in the brain, alternative theories regarding the computational basis of this monitoring function have been proposed including reinforcement learning and error-likelihood prediction. Most contemporary models assume that a single common mechanism underlies post conflict adjustments in control as well as the slower and more accurate responding that is typically seen following errors during speeded forced choice responding. These theoretical issues are discussed in ref.\textsuperscript{28}

**Neural Systems**

**Electrophysiology.** The switching Stroop task has been used in a number of studies incorporating the ERP method to examine the neural correlates of cognitive processes related to rule generation and selection during task preparation associated with presentation of a cue.\textsuperscript{25,29,30} Converging with the larger literature using ERPs to study task switching, each of these studies has revealed components of the ERPs over the parietal region of the scalp that distinguish pure trials from repetitions—reflecting mixing costs—and repetitions from alternations—reflecting switching costs. Differences in ERP amplitude between the various types of trials tend to emerge between 200 and 400 ms after cue onset and persist until between 1000 and 1200 ms after cue onset. The amplitude of the P3b component is greater for mixed trials than for pure trials.\textsuperscript{30,31} This effect has been interpreted as reflecting the need to encode the cue in mixed blocks relative to pure blocks.

In relationship to dynamic adjustments in control, a reliable ERP consisting of a modulation of the second negative deflection following stimulus onset in stimulus aligned data has been reported. This conflict-related negativity is seen on a variety of conflict eliciting tasks including the Stroop and the Eriksen Flanker paradigm.\textsuperscript{32–34} It is widely considered to be computationally related to the error-related negativity (ERN), another ERP-based marker of dynamic adjustments in control that may be identified in response aligned data comparing correct and incorrect trials.\textsuperscript{28,32} The Stroop task has been one of the more widely used measures for eliciting error-related brain activity during ERP recordings as well as during fMRI.\textsuperscript{35,36}

**Functional Neuroanatomy.** MRI studies using the switching Stroop and analogous paradigms consistently
Results in Schizophrenia

Using the switching Stroop task, Cohen et al. found that individuals with schizophrenia were much more susceptible to intrusion errors, particularly for the color task, than were healthy individuals or psychiatric control patients. This finding leads to the suggestion that schizophrenia is associated with a deficit in rule generation and selection.

A study using ERPs found that attention switching in response to a task cue was associated with increases in the amplitude of the P2 and P3b components for alternations relative to repetitions, however, the magnitude of this difference did not differ significantly between patients and controls. This pattern was replicated in a study comparing informative and noninformative cues. In this study, the magnitude of the effects of cue informativeness and task switching on ERPs over the central and parietal regions appeared to be similar in patients and controls, although the topographic maps isolating the effect of switching reveals some differences in the distribution of the effect between these groups for the parietal slow wave. This finding may converge with behavioral evidence indicating that schizophrenia is associated with a deficit in task set maintenance. Snitz et al. used an analogue of the switching Stroop, the POP task, and showed reduced cue-related activity in the DLPFC for high control trials in never medicated first-episode schizophrenia patients, consistent with disruption of this element of the cognitive control circuitry associated with impaired rule generation and selection.

The Stroop task has been reliably used to elicit brain activity associated with dynamic adjustments in control. ERP studies have shown reductions in both the ERN and the conflict-related negativity/N450 in schizophrenia, summarized in ref. Similarly, studies using fMRI during Stroop performance have consistently revealed reduced activity in the ACC associated with both errors and conflict. In the case of both methodologies, the results parallel those observed using related methodologies such as the POP task and the Eriksen Flanker paradigm. It is important to note that while fMRI and ERP studies have almost uniformly indicated reductions in error- and conflict-related neural signatures in schizophrenia, the results of behavioral measures of dynamic adjustments have been more mixed. While a number of studies have reported decreased behavioral adjustments to errors and conflict in the presence of reduced neural activity, others have reported intact adjustments in the presence of reduced neural signals. The reasons for these discrepancies, which might be related to the experimental design and instructions, patient factors or medication effects remain unclear and this should clearly be the basis of further investigation.

Pharmacology and Animal Models

Related to rule generation and selection, 2 studies have examined the neural systems underpinning task switching in the macaque using single unit recording and pharmacological challenge. Study examining single unit activity in the posterior parietal cortex revealed that neurons in this region demonstrated sustained firing in the cue-to-stimulus interval that was selective for 1 of the 2 tasks. The data bear some resemblance to the finding from the switching Stroop literature demonstrating that the amplitude of the parietal slow wave differentiating alternations from repetitions is greater for the color than word task. In a second study, the investigators examined the impact of ketamine challenge on task switching. Ketamine administration led to an increase in the interference effect for response time and accuracy in both animals. In contrast, the effect on task switching was more variable, with one monkey demonstrating increased switch costs for response time and the other demonstrating a trend toward greater switch costs in accuracy.

Studies of the neural systems in animals underlying dynamic adjustments in control have been limited to nonhuman primates and have generated some controversy. While neuronal recordings from the dACC in humans have clearly demonstrated conflict-related activity studies in nonhuman primates have consistently failed to show this activity. This has led some investigators to suggest that there may be important species differences between macaques and humans, possibly driven by unique cytoarchitectural features of the dACC that are only present in great apes.
A number of studies have shown that conflict- or error-related evoked potentials are positively modulated by dopaminergic neurotransmission (reviewed in ref. 51). There are also reports that these potentials are modulated by a common serotonin transporter polymorphism— to anxiety-related individual differences and clinical phenotypes of anxiety and depression, however, these results have been more mixed. 52

Construct Validity and Measurement Properties

Construct Validity. Many of the findings considered in the section on neural systems lend support to the construct validity of ERP and fMRI measures of the neural correlates of task switching. Studies using ERPs reveal that distinct components of the physiology can be associated with different cognitive processes thought to underpin task switching (eg, P3b— cue encoding, parietal positivity— cue retrieval, parietal slow wave— task set configuration). Physiological measures taken from both ERP and fMRI studies can discriminate between different types of switches (eg, attention, effector, intention) and processes that are recruited with different types of cues (eg, task cues vs transition cues that require the same switch of task set). Together, this evidence lends support to the suggestion that ERP and fMRI indices can provide biomarkers of distinct cognitive processes related to rule selection— associated with executive function— and goal maintenance— associated with working memory— that may be the target of behavioral and pharmacological interventions in schizophrenia.

Practicality and Tolerability. A critical consideration in utilizing ERP and fMRI measures as biomarkers represents the tolerability and practicality of the paradigms. As is common in the ERP literature, studies using this method to examine the neural correlates of task switching have tended to use a relatively large number of trials (eg, 750–1000) to obtain robust estimates of components of the ERPs. This large number of trials may lead to concern related to the issue of tolerability, especially if the switching Stroop task were being considered with other measures. However, evidence from a recent study reveals that reliable differences in components of the ERPs associated with task switching can be obtained with as few as 25–50 trials per condition. These data lead to the suggestion that robust estimates of components of the ERPs related to specific cognitive processes underpinning task switching could be obtained with a design including a moderate number of trials.

Conflict adaptation effects can be examined in the EEG using similar numbers of trials while studies of error-related activity have often used fewer trials with success, presumably due to the robustness of this measure, which can usually be observed in single subject data. Similar levels of sampling are typical in fMRI studies examining conflict- and error-related brain activity.

Psychometrics. The psychometric properties of behavioral and physiological measures taken from the switching Stroop (or the regular blocked version) have not been examined.

Summary and Conclusions

The selection of imaging biomarkers that engage the neural systems supporting component processes of executive control is a first step toward addressing a growing need in treatment development and personalized medicine. A number of additional steps will be needed to complete this process. In the case of the AX CPT, many of these steps are currently underway and ongoing progress, including their psychometric properties and the reliability of their performance as fMRI paradigms in healthy subjects and people with schizophrenia, can be obtained at the CNTRICS web site. Versions of this and several other paradigms can be downloaded from this site. It is important to note that many others, beyond those recommended at this meeting, will also have the potential for serving as valid and reliable imaging biomarker measures. CNTRICS sought, through its consensus-based process, to recommend a limited set of tasks with strong construct validity as well as meeting additional criteria outlined above. Additional measures surely can and will be developed, in each case requiring the developmental processes outlined in figure 1 of the overview article in order to be ready for use during treatment development for impaired cognition in schizophrenia and other disorders affecting executive functions.

Funding

Funded by grant (R13078710) from NIMH to C.S.C.

Acknowledgments

The authors have declared that there are no conflicts of interest in relation to the subject of this study.

References

2. Carter CS, Barch DM, Bullmore E. Cognitive neuroscience treatment research to improve cognition in schizophrenia (CNTRICS) II: developing imaging biomarkers to enhance treatment development for schizophrenia and related disorders. Biol Psychiatry. 2011;70:7–12.


18. Edwards BG, Barch DM, Braver TS. Improving prefrontal cortex function in schizophrenia through focused training of cognitive control. *Front Hum Neurosci*. 2010;5:12.


