Selective attention in schizophrenia: relationship to verbal working memory

Deanna M. Barch *, Cameron S. Carter

University of Pittsburgh Medical School, Department of Psychiatry, 3811 O'Hara Street, Pittsburgh, PA 15213, USA
Received 13 February 1998; accepted 19 May 1998

Abstract

In previous work using the Stroop task to examine cognitive function in schizophrenia, we have suggested that reaction time (RT) facilitation and error interference should be more sensitive measures of cognitive function than RT interference. We examined this hypothesis in 36 DSM-IV schizophrenia and schizoaffective patients, who performed both the Stroop and the Speaking Span, a measure of verbal working memory. The results supported our hypotheses, demonstrating that RT facilitation and error interference were associated more strongly with working memory performance than RT interference. The robust correlations between these measures of selective attention and Speaking Span performance has implications for understanding the nature and selectivity of cognitive dysfunction in schizophrenia. We present several different hypotheses that may explain this relationship, including: (1) a generalized deficit; (2) a common cognitive disturbance; and (3) a common neurobiological dysfunction. © 1998 Published by Elsevier Science B.V. All rights reserved.

Keywords: Selective attention; Stroop; Schizophrenia; Cognition; Working memory

1. Introduction

In recent years, a number of researchers have utilized single-trial versions (i.e., one stimulus at a time) of the Stroop color naming task to examine selective attention in schizophrenia. In the Stroop task, participants are presented with words printed in colors. They are instructed to ignore the word and to name the color in which it is printed. When the word and its color conflict (such as RED printed in blue), participants are slower than when there is no such conflict. This very robust and reliable effect is called interference, and is thought to result from the obligatory nature of word reading disrupting color naming (MacLeod, 1991). When presented with a word printed in a congruent color (e.g., RED printed in red), participants are faster than when they are presented with a neutral, color-unrelated stimulus (e.g., DOG printed in red). This effect is referred to as facilitation (MacLeod, 1991).

On the surface, one would predict that if schizophrenia patients suffer from a deficit in selective attention, they should demonstrate increased reaction time (RT) Stroop interference. In other words, one would expect patients to show a greater increase in RTs when they have to color name conflicting words, compared to color naming neut-
eral stimuli. However, the majority of researchers do not find increased RT interference on the single-trial Stroop task among schizophrenia patients. Instead, patients demonstrate greater RT facilitation (neutral–congruent RT), and greater interference in errors (incongruent–neutral errors) (Carter et al., 1992; Taylor et al., 1996; Schooler et al., 1997; Barch et al., in press, 1998; Cohen et al., 1998).

In previous work, we have suggested that a deficit in selective attention can explain the pattern of performance displayed by schizophrenia patients on the Stroop task—no increase in RT interference, but an increase in RT facilitation and error interference (Barch et al., in press). Specifically, we have hypothesized that schizophrenia patients show this particular pattern of Stroop performance for at least two reasons. First, if schizophrenia patients have difficulties in ignoring the influence of the irrelevant dimension (i.e., the word), this disturbance should influence performance in all conditions of the Stroop, not just in the incongruent condition. In the congruent condition, the word information could contribute to a relative speeding of responses. At the same time, in the neutral condition, a stronger influence of the word could cause patients to experience a greater degree of ‘interference’ from neutral stimuli (if they are words) than do controls. Relatively faster RTs for congruent stimuli and/or slower RTs for neutral stimuli would contribute to an increased Stroop facilitation effect in schizophrenia patients, given that facilitation scores are expressed as a difference between neutral and congruent RTs. At the same time, slower RTs for neutral stimuli would also effect the magnitude of Stroop RT interference. Because interference is also expressed as a difference score (incongruent–neutral RT), a slowing of RTs for neutral stimuli would tend to decrease the magnitude of RT Stroop interference, even if RTs for incongruent stimuli were increased among schizophrenia patients.

A second factor that may influence the lack of increased RT interference in schizophrenia patients is their increased errors in the incongruent condition. Relatively slower RTs in the incongruent condition are thought to occur because the conflicting information in the word dimension interferes with the naming of the print color. If schizophrenia patients have a deficit in the ability to overcome the influence of the word, they may be more likely to actually respond to the word instead of the print color. This may occur in the incongruent condition (but not in the congruent and neutral conditions) because the incongruent condition contains the greatest amount of conflicting word information. Thus, increased ‘interference’ may manifest in errors rather than RTs among schizophrenia patients.

If our hypotheses are correct, then among schizophrenia patients, RT facilitation and error interference should be associated more strongly than RT interference with other indices of cognitive function. In previous work, we have found some support for this hypothesis, demonstrating that increased RT facilitation and error interference, but not RT interference, were associated with poor performance on a version of the AX-CPT (Cohen et al., 1998). The goal of the present study was to further empirically test these predictions, by examining the relationship between Stroop performance and performance on an independent measure of cognitive function, the Speaking Span, a measure of verbal working memory (e.g., Daneman and Green, 1986). The participants in the current report were a subset of those who participated in a previous study examining Stroop performance in schizophrenia (Barch et al., in press). The subset reported in the current study also participated in a simultaneous study examining language function in schizophrenia, during which they completed the Speaking Span.

2. Methods

2.1. Participants

Participants were 36 DSM-IV schizophrenic or schizoaffective patients, who were either in-patients (N = 17) at Mayview State Hospital or out-patients (N = 19) at Western Psychiatric Institute and Clinic. All patients were medicated and had been receiving the same medications and dosages for at least 2 weeks. Diagnoses were based on the Structured Interview for DSM-IV (Spitzer et al., 1990), an interview with a primary caretaker, and a review of the participant’s medical records.
Diagnostic interviews were completed by one of the authors (D.M.B.) or a trained research assistant. Potential participants were excluded for: (1) substance abuse within the previous 6 months; (2) neurological illness or history of head trauma; (3) mental retardation; (4) non-native English speaker; or (5) color blindness. Additional demographic and clinical characteristics of the participants are shown in Table 1. Daily oral doses of anti-psychotics were converted to chlorpromazine equivalents according to guidelines suggested by Davis et al. (1983). All participants signed informed consent forms in accordance with the University and Mayview State Hospital institutional review boards, and were paid for their participation.

The Positive and Negative Symptom Scale (PANSS) (Kay, 1991) was used to evaluate clinical state (Table 1). Ratings were completed either by a Ph.D.-level psychologist (D.M.B.) or by a trained research assistant who regularly participated in training and reliability sessions. A subset of nine patients were rated by both experimenters. Intrarater reliability, measured using intraclass correlations (Shrout and Fleiss, 1979) with raters treated as random effects and the individual rater as the unit of reliability, was 0.95 for the total PANSS score, 0.95 for the Positive Symptom subscale, 0.89 for the Negative Symptom subscale, and 0.61 for the General Psychopathology subscale.

2.2. Tasks

2.2.1. Stroop task

Each participant was administered three blocks of the Stroop task, with block order counterbalanced across participants. We varied the nature of the neutral stimulus across the three blocks to examine the influence of neutral stimulus type on the performance of schizophrenia patients (Barch et al., in press). This manipulation was not relevant for the current study. Thus, in the analyses presented below, data are collapsed across the three neutral types. None the less, here we report on the full procedures of task administration. Each Stroop block consisted of 96 trials, with 24 (25%) congruent, 24 (25%) incongruent and 48 (50%) neutral. Congruent stimuli consisted of one of the four color names presented in its own color (i.e., red, green, blue, purple). Incongruent stimuli consisted of each of the four color names presented in one of the three remaining colors. In one block, the neutral stimuli consisted of four squares printed in one of the four colors. In a second block, the neutral stimuli were one of four color-unrelated words (dog, bear, tiger, or monkey) printed in one of the four colors. These neutral words matched the color words in number of letters and frequency (Francis and Kucera, 1982), and were from a single semantic category to eliminate semantic confounds (MacLeod, 1991). In the third block, the neutral stimuli consisted of four additional color words (tan, gray, white and yellow), printed in one of the four target colors (red, green, blue, purple). These color words were also matched to the primary color words on length and frequency.

Subjects were tested individually. Stimuli were presented on a Macintosh computer, using PsyScope software (Cohen et al., 1993). Participants were told that they would be presented with one stimulus at a time, and that they should name the color in which the stimulus was printed as quickly and accurately as possible. Each stimulus remained on the screen until the subject responded, or until 2000 ms had elapsed, and then was replaced by a fixation cross that lasted until
the onset of the next stimulus. Regardless of RT, a new trial started 4 s after onset of the previous stimulus, so that the pace of the task was fixed for all participants. Participants’ verbal responses were coded for accuracy by the experimenter. RTs for onset of verbal response were recorded automatically by the computer using a microphone and a voice-activated relay.

2.2.2. Speaking Span test
All participants completed this task on the same day as the Stroop. The Speaking Span was administered to obtain a measure of individual differences in verbal working memory, and is an analog of the Reading Span Test (Daneman and Carpenter, 1980). The Reading Span was also designed as a measure of individual differences in working memory, and numerous studies have shown that it strongly predicts performance on real world language comprehension tasks (Daneman and Merikle, 1996). Thus, the Reading Span has good ecological validity, but was designed specifically to assess language comprehension. The Speaking Span is almost identical to the Reading Span in design and presentation format, but was designed specifically to measure working for language production (Daneman and Green, 1986). In this task, participants see a series of words appear on the computer screen, one at a time. The series start at two words, and increase by one word every three trials to a maximum of six words. The participant is asked to remember each word in a trial until the end of the trial (signaled by a blank screen). At this point, the participant is asked to use each of the words from the trial in a grammatically correct English sentence. Participants are told to use each word in a different sentence, and to not change the form of the word. After the participant produces the sentences for the previous trial, a new trial starts. Participants are presented with three trials at each series length, for a total of 15 trials. Words were presented on an MacIntosh computer, using PsyScope software (Cohen et al., 1993), and responses were recorded manually by the experimenter.

2.2.3. WAIS-R Vocabulary
The vocabulary subtest of the WAIS-R was administered to obtain a measure of general verbal intelligence. Vocabulary was chosen because it is the subtest that displays the highest correlation with Verbal IQ ($r = 0.90$) (Wechsler, 1981). Raw scores on the Vocabulary subtest were converted to standard scores based on the guidelines presented in Wechsler (1981). The mean standard score was 7.72 ($SD = 2.72$).

2.3. Data analysis
Means for correct responses were used in analyses examining Stroop RTs. Outliers were removed by eliminating RTs that were above or below 2 standard deviations of the mean for that condition for an individual subject. Measures of RT facilitation (neutral-congruent RT), RT interference (incongruent-neutral RT), and error interference (incongruent-neutral accuracy) were calculated with data collapsed across the three neutral types. For the Speaking Span, the dependent measure was the total number of words remembered and used correctly in a sentence. Thus, a positive score indicates better performance. Kolmogorov-Smirnov tests indicated that all dependent measures were normally distributed. Thus, the following correlational analyses utilized Pearson product-moment correlations.

3. Results
The means and standard deviations for the Speaking Span and each of the Stroop measures are shown in Table 2. We began by examining the first-order correlations between the Stroop measures and Speaking Span performance. RT facilitation ($r = -0.67$, $p < 0.01$, two-tailed) and error

---

1 In the study examining Stroop performance in schizophrenia (Burch et al., 1998), the full set of 40 participants displayed significantly more RT facilitation and error interference compared to a group of 20 demographically equated controls. Identical results were obtained when comparing this subset of 36 patients with the controls. This subset of 36 patients did not differ significantly from the full set of 40 participants on any of the demographic or clinical characteristics described in Table 1.
interference ($r = -0.41, p < 0.05$, two-tailed) were significantly negatively correlated with Speaking Span performance. In contrast, RT interference was not significantly associated with Speaking Span performance ($r = -0.05, p > 0.10$, two-tailed).

To determine whether either error interference or RT facilitation was significantly more negatively correlated with Speaking Span performance than RT interference, we utilized the methods for comparing correlated correlation coefficients suggested by Meng et al. (1992). Both RT facilitation ($Z = -3.29, p < 0.001$, one-tailed) and error interference ($Z = -2.44, p < 0.01$, one-tailed) were significantly more negatively correlated with Speaking Span performance than RT interference. In addition, RT facilitation was significantly more negatively correlated with Speaking Span Performance than error interference ($Z = -1.72, p < 0.05$, one-tailed).2

Correlations can be reduced if a measure has low variance or low reliability. To be sure that such potential psychometric differences among the Stroop measures did not account for their differential correlations with the Speaking Span, we examined their variance and reliability. As can be seen in Table 2, the standard deviation for RT interference was greater than the standard deviation for RT facilitation. Thus, reduced range of the RT interference measure cannot explain its failure to correlate with the Speaking Span. We examined the reliability of the Stroop measures by calculating split-half correlations. This was done by calculating RT facilitation, RT interference and error interference separately for even-numbered and odd-numbered trials. As can be seen in Table 2, the split-half reliability was higher for RT interference than for either RT facilitation or error interference. Thus, lower reliability for the RT interference measure cannot explain its lack of correlation with the Speaking Span. However, RT facilitation was more reliable than error interference. Thus, the stronger correlation between RT facilitation and Speaking Span than error interference and Speaking Span may be attributable, at least in part, to the better reliability of RT facilitation.

It is possible that error interference and RT facilitation were associated with Speaking Span performance because they all simply reflect lower verbal intelligence, rather than some more specific cognitive dysfunction. Consistent with this hypothesis, Speaking Span performance and WAIS-R Vocabulary scores were significantly positively correlated ($r = 0.47, p < 0.01$, two-tailed). In addition, Vocabulary scores were significantly correlated with RT facilitation ($r = -0.34, p < 0.05$, two-tailed) and marginally correlated with error interference ($r = -0.31, p = 0.07$, two-tailed), but were not associated with RT interference ($r = 0.02, p > 0.10$, two-tailed). To explore this hypothesis in more detail, we examined whether the Stroop performance measures would still be associated

<table>
<thead>
<tr>
<th>Measure</th>
<th>Mean</th>
<th>SD</th>
<th>Split-half reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stroop</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reaction time facilitation</td>
<td>100.17</td>
<td>(66.11)</td>
<td>0.52</td>
</tr>
<tr>
<td>Error interference</td>
<td>0.08</td>
<td>(0.064)</td>
<td>0.26</td>
</tr>
<tr>
<td>Reaction time interference</td>
<td>120.00</td>
<td>(89.74)</td>
<td>0.70</td>
</tr>
<tr>
<td>Speaking span</td>
<td>26.31</td>
<td>(6.55)</td>
<td></td>
</tr>
</tbody>
</table>

---

2 It is possible that the correlations of the difference scores with Speaking Span are secondary to correlations with performance in specific conditions (e.g., neutral RT). Thus, we computed the correlations between Speaking Span and RT from the three Stroop conditions, only one of which was significant (neutral RT: $r = -0.36, p < 0.05$, two-tailed). A partial correlation between RT facilitation and Speaking Span performance, controlling for neutral RT remained highly significant ($r = -0.61, p < 0.01$, two-tailed). In contrast, a partial correlation between neutral RT and Speaking Span was no longer significant when controlling for RT facilitation ($r = -0.10, p > 0.10$, two-tailed). Thus, the correlation between Speaking Span performance and RT facilitation does not appear to be secondary to a correlation between Speaking Span performance and neutral RT.
with Speaking Span performance after accounting for the relationship between verbal intelligence and Speaking Span performance. We conducted three hierarchical multiple regressions, using the WAIS-R Vocabulary scores and the three Stroop measures to predict Speaking Span performance. The Vocabulary scores were forced to enter the equations in the first step, and the Stroop measures were entered in the second step. After entering the Vocabulary scores, RT facilitation was still a significant predictor of Speaking Span performance ($R^2$ change = 0.29, $p < 0.001$). Error interference was only a marginally significant predictor of Speaking Span performance after entering the Vocabulary scores ($R^2$ change = 0.08, $p = 0.06$). RT interference was not a significant predictor of Speaking Span performance after entering the Vocabulary scores ($R^2$ change = 0.003, $p > 0.10$).

4. Discussion

The results of this study indicated that, as predicted, RT facilitation on the Stroop was strongly associated with poor performance on the Speaking Span. This was true even after accounting for the relationship between verbal intelligence and Speaking Span performance. In contrast, RT interference was not associated with performance on the Speaking Span task. Further, RT facilitation was significantly more strongly related to poor Span performance than was RT interference. It should be noted that this later finding cannot be accounted for by psychometric differences between the two measures. RT facilitation was also significantly more strongly related to Span performance that error interference. However, it is possible that this difference reflects, at least in part, the fact that the RT facilitation measures had higher reliability than the error interference measure.

We had also predicted that error interference on the Stroop task would be associated with poorer Speaking Span performance. Consistent with this hypothesis, error interference was significantly negatively correlated with Speaking Span performance, a correlation that was significantly stronger than the correlation between RT interference and Speaking Span performance. However, error interference was only a marginal predictor of Span performance after accounting for verbal intelligence.

These findings raise interesting questions about the selectivity and nature of cognitive deficits in schizophrenia. The Stroop is traditionally thought of as a paradigmatic measure of selective attention (MacLeod, 1991), and entails very little memory load, except for maintaining the task instructions. In contrast, the Speaking Span makes substantial demands upon phonological rehearsal mechanisms within working memory (Daneman and Green, 1986). As such, there are at least three possible explanations for the relationship between Stroop and Speaking Span performance in schizophrenia patients. The first is that these correlations simply reflect a non-specific generalized deficit that may or may not have pathophysiological importance. This explanation would be consistent with the fact that schizophrenia patients tend to show deficits on many different tasks that appear to tap a range of cognitive functions. However, if the relationship between Stroop and Speaking Span performance in schizophrenia reflects a generalized deficit, one would have expected a different pattern of Stroop performance in patients. A generalized deficit implies that patients should show the worst performance in the conditions that are also the hardest for controls (Chapman and Chapman, 1978). On the Stroop, this would imply that patients should show increased RT interference which, in turn, should be correlated with Span performance. Thus, the fact that schizophrenia patients instead show increased RT facilitation and error interference (Carter et al., 1992; Taylor et al., 1996; Schooler et al., 1997; Barch et al., in press, 1998; Cohen et al., 1998), which are correlated with Span performance, argues against the generalized deficit hypothesis.

A second explanation is that performance on the Stroop and Speaking Span are correlated in schizophrenia patients because a specific common underlying cognitive mechanism supports performance on both tasks. For example, in Baddeley's (Baddeley and Della Sala, 1996) classical formulation, working memory is a complex system comprising buffer systems and a central executive. A critical function of the central executive is the...
allocation of attention during task performance to support both the maintenance and manipulation of information in the buffer systems. Performance on the Speaking Span would require the operation of both the central executive and a buffer system (i.e., phonological rehearsal in the articulatory loop). However, performance on the Stroop requires the central executive, but does not clearly require the operation of a buffer system such as the articulatory loop. Thus, the correlation between impairments on the Stroop and the Speaking Span suggests that a common cognitive dysfunction might be a deficit in central executive function, rather than a disturbance specific to a buffer system. In a connectionist formulation, Cohen and colleagues (Cohen and Servan-Schreiber, 1992; Cohen et al., 1996) have argued that a single underlying mechanism, the ability to represent and maintain context information, supports both selective attention and active memory. In their models, context representations accomplish selective attention by providing support for task-relevant processes (e.g., color naming), allowing these to compete against irrelevant ones (e.g., word reading). Context representations also accomplish active maintenance by providing support for task-relevant information against the interfering effects of noise over time. If Cohen et al.'s hypotheses are correct, then deficits on the Stroop task (selective attention) and the Speaking Span (active memory) may both reflect a common underlying mechanism in schizophrenia.

A third possible explanation for the correlation between Speaking Span and Stroop deficits in schizophrenia is that it reflects a common neurobiological dysfunction underlying different cognitive deficits. Thus, Speaking Span and Stroop deficits may reflect disturbances in different cognitive functions (i.e., active memory and selective attention). However, the same neurobiological dysfunction could lead to deficits in both active memory and selective attention. For example, one hypothesis is that a reduction in dopamine modulation of frontal cortex in schizophrenia contributes to deficits in both active memory and selective attention. Numerous studies in human and non-human primates have shown dorsolateral prefrontal cortex (DLPFC) to be active when participants are required to maintain information over time (e.g., as is required during Speaking Span Performance). Schizophrenia patients show a failure to activate DLPFC during such tasks (e.g., Weinberger et al., 1986; Berman et al., 1986). However, administration of a dopamine agonist increases DLPFC blood flow during task performance in patients with schizophrenia (Daniel et al., 1991), consistent with the hypothesis that a disturbance in dopamine modulation may contribute to DLPFC activation deficits. In addition, recent studies using PET have shown that the anterior cingulate (AC) is reliably activated when subjects perform variations of the Stroop task (Pardo et al., 1990; Bench et al., 1993; Taylor et al., 1994; Carter et al., 1995). Patients with schizophrenia show reduced AC blood-flow responses during Stroop performance (Carter et al., 1997), as well as during performance of related tasks (e.g., Andreasen et al., 1992). The AC has dense dopaminergic innervation (Benes et al., 1997) and its blood-flow and metabolism are modulated by antipsychotic drugs (Miller et al., 1997; Holcomb et al., 1996). Further, in schizophrenia, AC blood-flow response during verbal fluency is increased by apomorphine (Dolan et al., 1995). Thus, it is possible that a common neurobiological substrate, such as altered dopaminergic modulation of local circuit function in various regions of frontal cortex, could account for the correlations seen in the current study. These latter two hypotheses are not mutually exclusive.

In summary, the results of this study provide strong support for the hypothesis that on the Stroop, RT facilitation is correlated more strongly with other measures of cognitive function than is RT interference. This study also provided some support for the hypothesis that error interference is correlated more strongly with other measures of cognitive function than is RT interference, but the results were not as strong as those for RT facilitation. These findings highlight the importance of examining measures of RT facilitation and error interference in future studies of Stroop performance in schizophrenia, and in studies of the relationship between cognitive deficits and symptomatology in this illness. We have proposed several hypotheses that could account for the relationship between Stroop and Speaking Span
deficits in schizophrenia, including: (1) a generalized deficit; (2) a common cognitive mechanism; and (3) a common neurobiological mechanism. Further research examining the convergent and divergent relationships among tasks that measure putatively different cognitive functions, as well as research identifying the neurobiological substrates of these cognitive processes, will be necessary in order to establish the precise nature and selectivity of cognitive dysfunction in schizophrenia. For example, large-scale behavioral and neuroimaging studies examining the interrelationships among tasks measuring constructs such as selective attention, inhibition and working memory, as well as the brain regions activated by such tasks, would help further this endeavor.

Acknowledgment

This work was supported by NARSAD Young Investigator’s awards to both authors. We wish to thank the staff and patients of Western Psychiatric Institute and Clinic and Mayview State Hospital for their kind cooperation and support.

References


Barch, D., Carter, C., Perlstein, W., Baird, J., Cohen, J., Schooler, N., 1998. Increased Stroop facilitation effects in schizophrenia are not due to increased automatic spreading activation. Submitted.


