ABSTRACT

The present study examined the effect of negative expectations on neuropsychological test performance. It was hypothesized that having attention called to a history of prior head injury and the potential effects of head injury on cognition would result in diminished neuropsychological test performance relative to individuals with a similar head injury history but who did not have their attention called to their head injury history (‘diagnosis threat’). Of 36 participants with a history of mild head injury, 17 were randomly assigned to diagnosis threat and 19 to neutral test directions. The diagnosis threat group performed significantly worse on tests measuring general intellect and memory, but were not different from the neutral group in basic attention or psychomotor speed. The diagnosis threat group rated themselves as putting forth less effort on the neuropsychological battery, and self-rated effort correlated with test performance in that group. Overall, results serve as a reminder that neuropsychological tests are measures of behavior, and thus can be influenced by nonneurological factors.
patients with chronic fatigue syndrome (van der Werf, Prins, Jongen, van der Meer, & Bleijenberg, 2000) and fibromyalgia (Gervais et al., 2001), raising concerns that researchers are reaching faulty conclusions about neurological dysfunction from neuropsychological test performance in these groups.

While recent studies provide evidence that poor effort is a major contributor to neuropsychological test performance, they do not necessarily demonstrate the reasons for poor effort. Certainly some individuals may be overtly malingering, which leads to poor effort and poor performance. However, there are likely other more subtle nonneurological factors that also lead to poor effort and diminished test performance. One such factor is the preexisting expectations for performance with which each examinee presents during testing. One example of preexisting negative expectations for performance is known as stereotype threat (Steele, 1997). The assumption behind stereotype threat is that a member of a particular group, when faced with a task that is thought to be poorly performed by members of that group, will face the threat of being judged by stereotypes about the group’s inferiority on that task. This threat is assumed to interfere with his/her performance. It is believed that this effect is seen even in those who do not personally believe the stereotype.

Evidence has accumulated that supports these assumptions. For example, there is widespread knowledge of the purported intellectual difference between Whites and African Americans. Katz, Roberts, and Robinson (1965) showed that African Americans did worse on subtests of an intellectual test when they were presented as measures of intelligence (which presumably called to mind the stereotype) rather than as measures of hand eye coordination. In fact, stereotype threat can be induced with more subtle manipulations of task instructions, such as simply informing African Americans that their performance would be compared to Caucasians (Katz, Robinson, Epps, & Waly, 1964), or asking African Americans to state their race on forms, thus priming their focus on their racial identity (Steel & Aronson, 1995).

Stereotype threat is not uniquely seen in studies of cognitive differences between racial/ethnic groups. Similar findings are seen in individuals from low socioeconomic backgrounds (Croizet & Claire, 1998). (Low SES students do worse on hard GRE problems when told that they assess intellectual abilities, but not when told they assess attention and memory.) Stereotype threat has also been demonstrated in studies of sex differences on cognitive tasks (Spencer, Steele, & Quinn, 1999; Walsh, Hickey, & Duffy, 1999). For example, Spencer and colleagues (1999) found that women underperformed on the math GRE when they were told to expect gender differences on the task, but performed at the same level as men when the test was presented as insensitive to gender differences. This is true even when women are matched with men on math skills and interest in math. Leyens, Desert, Croizet, and Darcis (2000) found that males who were told that males perform worse than females on affective processing tasks did in fact perform worse than males who were not given this information. Stereotype threat has also been suggested as a potential explanation for some of the differences in cognitive performance as we age (Levy, 1996). She demonstrated that subliminal priming of older adults with negative stereotypes about aging caused them to perform more poorly on cognitive tasks, while priming them with positive stereotypes about aging resulted in improved performance. The effect was unique to older individuals; the performance of young adults primed with negative or positive aging stereotypes was not altered.

Aronson et al. (1999) suggest that stereotype threat is a situational, not characterological, effect; that is, anyone can be made to underperform on tasks if exposed to a stereotype that suggests personal inferiority. They gave white males with high math SAT scores a challenging math test. When told that Asians outperform Caucasians in math, white males performed worse than when they were not given that information. The stereotype threat effect was most strongly seen in those who premorbidly identified highly with math skills. Thus, stereotype threat may be seen in any individual who is exposed to beliefs that suggest a personal inferiority, even when the individual is unlikely to have premorbid feelings of inferiority with regard to his/her cognitive skills. The premise of
this paper is to test whether stereotype threat will be seen in those who have a history of head injury. We believe that calling attention to a personal history of head injury and its potential effects on cognition might lead to worse cognitive performance than that seen in individuals with similar head injury history, but who do not have attention called to either their head injury history or the possible consequences of head injury. We call this ‘diagnosis threat.’

Why is the stereotype threat effect seen? It has been suggested that stereotype threat calls up negative expectations for individual performance, which then increases anxiety and/or reduces effort, leading to worse performance. However, evidence for the relation of anxiety and effort to stereotype threat is mixed at best. Steele and Aronson (1995) showed that African Americans subject to the stereotype threat also rated themselves as having less confidence and performing less well, and they also endorsed more negative stereotype beliefs on a stem completion task. However, stereotype threat did not appear to affect their self-reported effort on the GRE-like task or their self-rated anxiety about the task. Similarly, anxiety, effort, perceived performance, and self-confidence did not appear to explain group differences in males subject to the stereotype threat (Aronson et al., 1999). However, in a study of gender-based stereotype threat (Spencer et al., 1999), posttest anxiety was related to performance. In addition, women who were subject to stereotype threat spent less time on individual test items, suggesting that decreased effort may have explained the difference in performance.

The present study examined the effect of ‘diagnosis threat’ on cognitive performance in persons with a history of mild head injury. We expected that those who were given test directions that made salient the possibility of neurocognitive effects of head injury would perform worse than those with neutral test directions. We also examined whether those exposed to diagnosis threat would be different than those with neutral directions on ratings of self-confidence about the cognitive tests, perceived difficulty in the tasks, and effort in completing the tasks, to attempt to explain any group differences.

METHOD

Participants
Over 2000 undergraduates at a medium-sized Midwestern university completed a large screening evaluation. All participants received extra credit points in their psychology classes in exchange for their participation in the study. The screening evaluation (which was approximately 1 hr long) included the Inventory of Depressive Symptomatology (IDS; Rush, Gullion, Basco, Jarrett, & Trivedi, 1996), a 30-item pen-and-paper measure of depressive symptoms, to screen for depression, and a head injury questionnaire (see Appendix A). Participants with a history of self-reported mild head injury (with loss of consciousness of more than 1 min but less than 30 min) but no symptoms suggestive of current depression (IDS < 18) and no self-reported neurological history were selected from the larger sample (approximately 11% of individuals from the larger sample met these criteria). A random sample of individuals meeting study criteria were then contacted by phone with an invitation to participate in the present study (72% agreed to participate). Of the 36 undergraduates who participated, 17 were randomly assigned to the diagnosis threat group and 19 to the neutral group.

Procedure
The protocol was approved by the University’s Institutional Review Board. Following informed consent, each participant was given an envelope with instructions inside. Those randomly assigned to the neutral group read the following instructions:

When the experimenter returns to the room, s/he will ask you to complete a brief collection of common neuropsychological tests. These tests will assess skills such as attention, memory, speed of information processing, problem solving skills, etc. Some of the tests are easy, some are more difficult. Please give your best effort. Questions about individual tests will be answered following the testing.

Those randomly assigned to the diagnosis threat group read the following:

You have been invited to participate in this study because of your responses to one of the questionnaires included in the mass screening at the beginning of the quarter. Your responses indicated a history of head injury/concussion. A growing number of neuropsychological studies find that many individuals with head injuries/concussions show cognitive deficits on neuropsychological tests. Deficits in areas such as
attention, memory, and speed of information processing are common – though other deficits sometimes emerge. This study examines the role that head injury may play in these cognitive areas to better understand the nature of the disorder. When the experimenter returns to the room, s/he will ask you to complete a brief collection of common neuropsychological tests. These tests will assess skills such as attention, memory, speed of information processing, problem solving skills, etc. Some of the tests are easy, some are more difficult. Please give your best effort. Questions about individual tests will be answered following the testing.

Participants were instructed to read the contents of their envelope, return them to the envelope, and to keep the examiners unaware of what the instructions said. Thus, all examiners were unaware of group assignment at the time of testing.

After reading the instructions, participants participated in a brief neuropsychological battery that included tests of memory, intellect, attention, and psychomotor speed. Memory tests were the Auditory Verbal Learning Test (AVLT; Lezak, 1995) immediate recall, delayed recall, and delayed recognition, and the Complex Figure Test (CFT; Rey, 1941) delayed recall. Intellect was assessed by the Wechsler Adult Intelligence Scale-III (WAIS-III; Wechsler, 1997) Information and Block Design subtests. Attention was assessed with the WAIS-III Digit Span and Letter Number Sequencing subtests. Finally, psychomotor speed was measured using Controlled Oral Word Association (COWA; Benton, Sivan, & Hamsher, 1994) total number words and Trailmaking Test (TMT; Reitan, 1971) time to complete part A and time to complete part B.

Following neuropsychological testing, participants rated, using a 9-point scale, how much effort they put into completing the tasks, how difficult they thought the tasks were, how confident they were in their performance, and how well they thought they did on the tasks (Appendix B). This measure was adapted from Aronson et al. (1999).

Participants were then debriefed. It was explained to each participant that a random set of participants received directions suggesting that head injury groups perform more poorly on cognitive tasks. It was further explained that some studies do find head injury effects on cognition, but that these studies tend to include participants with more severe head injuries than they reported experiencing in their past. In addition, participants were informed that many individuals who have a history of head injury don’t experience cognitive difficulties at all, and that, given their high functioning (college), they should not be worried about their cognitive skills. However, they were also offered feedback about their individual performance if interested (3 participants requested feedback, 2 from the diagnosis threat condition and 1 from the neutral condition).

RESULTS

Groups were not different in age, \( t = .46, p = NS \), or years of schooling, \( t = .14, p = NS \). Sex distribution was not different among groups, \( \chi^2(1) = 1.99, p = NS \). (See Table 1.) All participants reported a history of one mild head injury, with greater than 1 but less than 10 min loss of consciousness, and greater than 1 but less than 30 min posttraumatic amnesia.

Performance on Neuropsychological Tests

It was expected that the groups would differ in memory performance. A MANOVA, with group status as the between groups measure and AVLT immediate recall, AVLT delayed recall, AVLT delayed recognition, and CFT delayed recall as the within groups measures was significant, \( F(4, 31) = 5.29, p < .005 \). Due to the directional nature of the hypotheses, one-tailed tests were used in the post hoc ANOVAS. As expected, the diagnosis threat group performed significantly worse on AVLT immediate recall, \( F = 6.68, p < .005 \), AVLT delayed recall, \( F = 2.86, p < .05 \), AVLT delayed recognition, \( F = 4.18, p < .05 \), and CFT delayed recall, \( F = 3.00, p < .05 \). (See Table 2.)

It was expected that groups would differ in performance on WAIS subtests of general intellect. A MANOVA, with group status as the between groups measure and WAIS Information

<table>
<thead>
<tr>
<th>Variable</th>
<th>Diagnosis threat</th>
<th>Neutral</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( N = 17 )</td>
<td>( N = 19 )</td>
</tr>
<tr>
<td>Age</td>
<td>18.6 (0.7)</td>
<td>18.5 (0.8)</td>
</tr>
<tr>
<td>Education</td>
<td>13.2 (0.6)</td>
<td>13.2 (0.5)</td>
</tr>
<tr>
<td>Male</td>
<td>35</td>
<td>53</td>
</tr>
</tbody>
</table>

Table 1. Demographic Characteristics of the Study Groups.
and Block Design as the within groups measures was significant, $F(2, 33) = 4.85, p < .01$. Due to the directional nature of the hypotheses, one-tailed tests were used in the post hoc ANOVAs. As expected, the diagnosis threat group performed significantly worse on both Information, $F(2, 33) < 1$, and Block Design, $F = 5.11, p < .01$.

It was also expected that groups would differ on attention and speed of information processing. MANOVA for attention (Digit Span, Letter Number Sequencing) was not significant, $F(2, 33) < 1$. MANOVA for psychomotor speed (Trails A, Trails B, and COWA) was also not significant, $F(3, 32) < 1$.

To examine whether there were differences between groups in clinically relevant findings, all participants were classified as ‘impaired’ on neuropsychological tests if their performance fell greater than 1 SD below the mean based on standard normative data on any of the clinical tests utilized in the study. Groups were different in the percentage of participants who were impaired by this classification: 37.5% of the diagnosis threat group and 6.6% of the neutral group had at least 1 score that was considered ‘impaired,’ $\chi^2(1) = 4.21, p < .05$.

### Ratings of Performance

$T$-tests revealed that the diagnosis threat group rated themselves as putting forth less effort on tasks, $t = 2.06, p < .05$. They also found the tests more difficult, $t = 1.73, p < .05$, had less confidence in their performance, $t = 1.76, p < .05$, and perceived themselves as doing less on the tests, $t = 2.50, p < .01$. They did not rate themselves as different in terms of the pressure they experienced during testing, $t = 1.18, p = NS$. (See Table 3.)

### Relationship of Effort to Test Performance

Pearson correlations revealed that self-ratings of test difficulty, confidence in performance, pressure to perform, and performance success were all highly correlated, all $p < .05$. However, none of these ratings was significantly correlated with self-rated effort on tasks, $r = .02$ to .27, all $p = NS$. (See Table 4.) Thus, self-rated effort appeared to be measuring a construct unique to the other self-ratings. We examined the ratings of

---

**Table 2. Performance on Neuropsychological Tests by Study Groups.**

<table>
<thead>
<tr>
<th>Neuropsychological Tests by Domain</th>
<th>Diagnosis threat $N = 17$</th>
<th>Neutral $N = 19$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M (SD)$</td>
<td>$M (SD)$</td>
</tr>
<tr>
<td>Memory</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AVLT immediate recall (of 15)***</td>
<td>11.5 (2.5)</td>
<td>12.1 (1.8)</td>
</tr>
<tr>
<td>AVLT delayed recall (of 15)*</td>
<td>11.1 (3.0)</td>
<td>12.4 (1.8)</td>
</tr>
<tr>
<td>CFT delayed recall (of 26)*</td>
<td>21.2 (5.2)</td>
<td>23.6 (3.4)</td>
</tr>
<tr>
<td>Intellect</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WAIS-III Information ACSS**</td>
<td>10.4 (1.8)</td>
<td>11.8 (2.1)</td>
</tr>
<tr>
<td>WAIS-III Block Design ACSS**</td>
<td>11.2 (2.7)</td>
<td>13.7 (3.2)</td>
</tr>
<tr>
<td>Attention</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WAIS-III Digit Span ACSS</td>
<td>11.1 (2.6)</td>
<td>11.5 (2.9)</td>
</tr>
<tr>
<td>WAIS-III Letter Number Sequencing ACSS</td>
<td>11.3 (2.5)</td>
<td>11.8 (2.7)</td>
</tr>
<tr>
<td>Psychomotor speed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TMT speed to complete part A</td>
<td>23.7 (7.9)</td>
<td>21.9 (7.0)</td>
</tr>
<tr>
<td>TMT speed to complete part B</td>
<td>53.1 (15.8)</td>
<td>48.3 (15.3)</td>
</tr>
<tr>
<td>COWA number of words</td>
<td>39.5 (6.9)</td>
<td>39.3 (8.9)</td>
</tr>
</tbody>
</table>

*Note. AVLT = Auditory Verbal Learning Test. CFT = Complex Figure Test. WAIS-III = Wechsler Adult Intelligence Test-III. ACSS = age corrected scaled score. TMT = Trailmaking Test. COWA = Controlled Oral Word Association

*p < .05 (one-tailed).

**p < .01 (one-tailed).

***p < .005 (one-tailed).
effort relative to test performance for each group individually, using only those neuropsychological tests for which there were significant group differences. In the diagnosis threat group, self-rated effort was significantly correlated with AVLT immediate recall, $r = -.53$, $p < .05$, CFT delayed recall, $r = -.53$, $p < .05$, and Information, $r = -.48$, $p < .05$. It was not significantly correlated with AVLT delayed recall or recognition or Block Design. In the neutral group, self-rated effort was not significantly correlated with any of the outcome measures. (See Table 5.)

**DISCUSSION**

As hypothesized, participants who were exposed to diagnosis threat performed worse on tests of general intellect and immediate and delayed memory relative to age-, education-, gender- and head injury history-matched participants who were given neutral test instructions. The two groups were not different in basic attention or psychomotor speed. Thus, exposure to information that highlighted their ‘diagnosis’ and emphasized the potential for poor cognitive performance in their diagnostic group led to poorer memory and intellectual test performance.

An exploration of factors that may have been related to their poorer performance suggests that effort may well be a major contributing factor. Individuals subject to diagnosis threat rated themselves as putting forth less effort on the tasks than those with neutral directions. They also rated the tasks as more difficult overall, felt they performed less well, and had less confidence in their performance. In the diagnosis threat group, self-rated effort correlated highly with immediate recall of the AVLT, delayed recall of the CFT, and Information subtest, while it did not correlate significantly with any neuropsychological variables in the neutral group.

Although it seems likely that effort was a major explanatory factor for group differences, it...
is unlikely that poor effort in the diagnosis threat group was related to overt malingering. No individuals in either group were involved in any litigation or disability claims for their relatively mild injuries. All were performing successfully in college. All were volunteers for a research project and were not participating for clinical reasons. It is unlikely that they were consciously performing poorly for any secondary gain. However, future studies in this area might examine the effects of diagnosis threat on performance on standard effort tasks, to determine whether diagnosis threat is significant enough to invalidate test results.

By virtue of the sample selection methods, it is unlikely that the participants in the present study were identifying themselves strongly as head injured prior to completing screening questionnaires that asked about their head injury history. None were involved in litigation related to their injury, none were in treatment related to their injury or any other psychological problem, and none reported symptoms of significant depression. In addition, as all were successful college students, it is less likely that they had strong feelings of cognitive inferiority prior to entering the study. Thus, it is even more surprising that group differences emerged based on minor changes to the instructions they received. One can imagine the effect of diagnosis threat in those who are already vulnerable, who already feel inferior in cognitive abilities, and who are actively experiencing medical and or legal procedures related to their injury. It might be of interest to examine the effects of diagnosis threat in a more clinically relevant sample, given these issues. But it is unethical to lead persons in a clinically important situation to have expectations that increase likelihood of invalid performance, even for a brief time. Given the small sample size, however, the present study should be replicated, perhaps using a community based sample, to determine whether our results are generalizable.

The effects demonstrated in the present study are certainly consistent with stereotype threat. Diagnosis threat and stereotype threat appear to be specific instances of the effects of negative response expectancies (Kirsch, 1999). For example, the nocebo effect is a well-documented finding that having expectations for medical symptoms because of something an examiner does to you or gives you can cause the symptoms in an expectant individual (Kennedy, 1961) and has been demonstrated to cause pain symptoms, allergy symptoms, asthmatic attacks, pseudoseizures, and the side effects of medications (Jewett, Fein, & Greenberg, 1990; Lancman, Asconape, Craven, Howard, & Perry, 1994; Luparello, Lyons, Bleecker, & McFadden, 1968; Myers, Cairns, & Singer, 1987, Schweiger & Parducci, 1981). A recent study by Foerster and Strack (1998) demonstrates the effects of negative response expectancies on memory. They asked students to learn a list of words that they were later to recall with or without music. Some were told that music would enhance learning, and others that it would inhibit it. Those who were told that music would inhibit learning did in fact do worse on the recognition task, consistent with a negative expectation for performance. While these studies illustrate the effects of being told that negative effects are expected following a particular event, stereotype threat and diagnosis threat illustrate that simply calling attention to group membership and beliefs about that group can also elicit expected outcomes (i.e., worse cognitive performance).

Applying the concept of negative response expectancies to head injury, it is possible that information provided to individuals following mild head injury may in fact increase their experience of symptoms that they then attribute to the injury itself. Several studies have demonstrated the effect of negative expectancies on the self-report of postconcussive symptoms (Ferguson, Mittenberg, Barone, & Schneider, 1999; Gunstad & Suhr, 2001; Mittenberg, DiGuilio, Perrin, & Bass, 1992). Further, Mittenberg, Tremont, Zielinski, Fichera, and Rayls (1996) demonstrated the efficacy of a brief cognitive-behavioral intervention based on expectancy theory given to individuals seen in an emergency room for head trauma. Those who received the therapy reported less frequent, less severe postconcussive syndromes and rated them as having shorter duration at 6-month follow-up. Perhaps the effect of the therapy was to minimize the expected negative outcomes induced by standard head injury information typically provided at
hospital discharge following any head injury. The present study provides evidence that, in addition to its effects on self-reported cognitive abilities, negative expectancies can actually influence performance on neuropsychological tasks.

The present results certainly do not lead to the conclusion that neuropsychological measures are invalid ways to assess brain dysfunction. There is a wealth of evidence demonstrating that brain damage leads to poor performance on neuropsychological measures in ways consistent with what we know about brain structure and function. However, the results do suggest that both clinicians and researchers must remain attuned to the effect that nonneurological variables can have on test performance and how this might affect the interpretation of clinical and research results. Further, both clinicians and researchers should attempt to measure the influence of well-known variables, such as effort, and account for its effects before reaching clinical or scientific conclusions about brain dysfunction.

REFERENCES


Clinical Pharmacology and Therapeutics, 42, 250–253.
APPENDIX A

Head Injury Questions Asked During Screening

Have you ever experienced a concussion or a head injury? YES NO

Were you knocked unconscious? YES NO
If yes, how long were you unconscious?
1. a few sec 2. < 5 min 3. < 30 min 4. > 30 min

Do you remember the events before and after your head injury? YES NO
If so, how long of a time period are you unable to remember?
1. a few sec 2. < 5 min 3. < 30 min 4. 30–60 min 5. > 60 min

APPENDIX B

Posttest Questions

1. How hard did you try on the tests?
   Not hard at all Very hard
   1 2 3 4 5 6 7 8 9

2. How difficult did you find the tests?
   Not difficult at all Very difficult
   1 2 3 4 5 6 7 8 9

3. How much pressure did you feel during testing?
   No pressure at all Very pressured
   1 2 3 4 5 6 7 8 9

4. How confident are you in your performance?
   Not confident at all Very confident
   1 2 3 4 5 6 7 8 9

5. How well did you do on the tests?
   Very poorly Very well
   1 2 3 4 5 6 7 8 9