Assessing Postoperative Cognitive Change After Cardiopulmonary Bypass Surgery

Julian R. Keith, Antonio E. Puente, Keri L. Malcolmson, Shane Tartt, and Amy E. Coleman
University of North Carolina at Wilmington

Cognitive decline after cardiopulmonary bypass (CPB) surgery has been a concern since the advent of CPB procedures. A primary focus of many studies on this topic has been to quantify the incidence of post-CPB cognitive impairment. However, studies that have used traditional parametric statistics have generally failed to confirm that long-lasting (≥ 1 month) cognitive declines occur reliably after CPB surgery. For the present study, the authors used a split-plot analysis of variance model that revealed preoperative memory impairments in the CPB patients and new postoperative impairments of attention. The authors discuss the assumptions of, and problems associated with, analysis methods that are often used to quantify the incidence of cognitive impairment following CPB surgery.

Cardiopulmonary bypass (CPB) surgeries are performed on over 500,000 individuals annually in the United States (American Heart Association, 2002). It is now clear that the surgical and extracorporeal perfusion procedures involved in CPB expose patients’ brains to a variety of abnormal physiological conditions including showers of microemboli, hypoxia, hyperglycemia, and inflammation (Barbut et al., 1997; Baris, Israel, Amory, & Benni, 1995; Blauth, Arnold, Schulenberg, McCartney, & Taylor, 1988; Harris et al., 1998; Lehot, Piriz, Villard, Cohen, & Guidollet, 1992; Moody et al., 1995; Mutch et al., 1997). CPB also triggers a cytokine response (Asimakopoulos & Taylor, 1998) that can interfere with cerebral function in its own right (Gibertini, Newton, Klein, & Friedman, 1995; Jankowsky & Patterson, 1999). Remarkably, frank neurological impairment occurs in only about 3% of patients (Eagle et al., 1999). However, postoperative declines in cognitive performance may be far more prevalent.

Cognitive impairment after CPB surgery has been reported since the advent of the procedure (Ehrenhaft & Claman, 1961; Gilberstadt & Sako, 1967; Gilman, 1965). The reported incidence of cognitive decline during the 1st postoperative week has consistently been quite high, with around 75% of CPB patients showing some evidence of cognitive decrements (for a review, see Benedict, 1994). Between 10% and 57% of CPB patients may continue to experience cognitive declines that persist beyond the 1st postoperative month (Grote et al., 1992; Patel, Turtle, Chambers, Newman, & Venn, 1993). The nearly sixfold difference across studies in the reported incidence of persistent postoperative cognitive decline is likely due, at least in part, to between-study differences in the criteria used to define significant cognitive decline (Mahanna et al., 1996).

In the CPB literature, the discussion of how to quantify the incidence of postoperative cognitive decline has focused on setting the threshold that a performance decrement must exceed in order for a patient to be classified as cognitively impaired (Kneebone, Andrew, Baker, & Knight, 1998; S. P. Newman, 1995; Slade, Sanchez, Townes, & Aldea, 2001; Stump, 1995; Stump, James, & Murkin, 2000). Two particular methods have been most widely used—the standard deviation (SD) criterion and the reliable change index (RCI; Dumas, Dupuis, Searle, & Cartier, 1999; Gugino et al., 1997; McLean et al., 1994; Nathan et al., 1995; M. F. Newman et al., 2001; Sylivris et al., 1998; Toner, Hamid, Peden, Taylor, & Smith, 1993; Toner et al., 1994; Townes et al., 1989; Vingerhoets, Van Nooten, & Constantin, 1997).
Patients are typically tested prior to surgery to obtain a baseline measure of cognitive performance and tested again at least once postoperatively. Investigators who use the SD criterion define impairment as a decline of one SD in performance on any of the individual cognitive tests included in the assessment battery (cf. M. F. Newman et al., 2001; Vanninen et al., 1998).¹ The SD is based on the mean of the preoperative baseline scores of the study sample. Several articles have drawn attention to some of the flaws inherent in the SD method (Andrew, Baker, Bennetts, Kneebone, & Knight, 2001; Kneebone et al., 1998; Slade et al., 2001). One problem is that this criterion is quite liberal in that one SD decrement would be expected to occur by chance alone with about a 14% likelihood (i.e., $\alpha = .14$).

The so-called RCI has been proposed as an alternative to the SD method (Andrew et al., 2001; Kneebone et al., 1998; Slade et al., 2001). The RCI differs from the SD method in three respects. First, the RCI uses the 90% confidence interval as the criterion for defining reliable change. Second, the RCI method includes a parameter that takes into account the test–retest reliability of the cognitive measure—the standard error of the measurement based on a control group’s (typically consisting of non-CPB volunteers) raw data. Third, the control group’s mean change scores across the two test sessions serve as an estimate of the control group’s expected practice effect (i.e., practice-related improvement on the test). So, $RCI = (SE_{diff}) \times (\pm 1.64) \pm$ practice effect, where $SE_{diff}$ is equal to $\sqrt{2(SE_{m})^2}$ and $SE_m = (SD_1)\sqrt{1-r_{xx}}$, where $(SD_1)$ is the SD of the baseline score, and $(r_{xx})$ is the test–retest reliability coefficient. Patients whose assessment scores fall outside of the RCI are defined as improved or impaired, depending on the direction of the difference.

If a researcher wanted to determine whether an individual patient’s score on a single cognitive task differed from the mean score of a comparison group by more than what would be expected by chance alone, then the RCI method would be superior to the SD method, though some of the assumptions of the RCI method have been questioned (Bruggemans, van de Vijver, & Huysmans, 1999). However, it is important to recognize that both the SD and the RCI methods retain or reject the null hypothesis (i.e., postoperative performance equals preoperative performance) for each research participant on a case-by-case basis. Therefore, when applying a fixed criterion to more than 1 participant, each of whom is evaluated using several different cognitive tests, special care is required to control the risk of committing Type I errors (i.e., rejecting the null hypothesis when it should be retained or, in the case of CPB patients, misclassifying an individual as impaired).

In other words, when the RCI method is used, each CPB patient’s postoperative change scores on each cognitive measure are judged relative to an RCI that is computed for a given task. The null hypothesis is rejected for a patient if his or her score falls outside of the RCI. Because the 90% confidence interval is used to compute the RCI, the likelihood of making a Type I error when classifying a single patient on a single task is 5%. When the test battery consists of several cognitive measures, however, familywise Type I errors increase proportionally to the number of tests included in the assessment battery. Moreover, the number of Type I errors committed using this analysis method also grows as the size of the study sample increases because the null hypothesis is tested individually for each research participant. To our knowledge, the problem of familywise error has not been taken into consideration when, as in CPB studies, the RCI and SD methods were used to classify cognitive impairment with the use of multiple cognitive tests to evaluate groups of patients. Thus, studies that have used the SD or RCI methods to classify postoperative impairment most likely overestimated incidence rates. Furthermore, it is not known whether surgery reliably caused persistent (i.e., beyond the first 2 postoperative weeks) cognitive declines in CPB patients relative to controls because inferential statistics were not reported in these studies (Andrew et al., 2001; Kneebone et al., 1998; M. F. Newman et al., 2001).

Traditional parametric, inferential statistics have been used to confirm the reliability of cognitive declines in CPB patients that occur during the first 10 postoperative days (cf. Blumenthal et al., 1991; Jacobs et al., 1998; M. F. Newman et al., 1994). However, studies that used parametric inferential statistics to compare between group performance changes (CPB vs. controls) or changes over time (pre- vs. postoperative) have failed to detect statistically significant declines in cognitive performance after the first few postoperative weeks ( Jacobs et al., 1998; Townes et al., 1989). In fact, in some cases, significant performance improvements on several cognitive tasks were reported for groups of CPB patients (O’Brien et al., 1992; Townes et al., 1989).

Recently, Stump et al. (2000) argued that parametric statistics were not appropriate for the analysis of CPB effects on cognitive performance because, in their words, “The ‘up goers’ and the ‘down goers’ offset each other so that the mean stays the same but the SD increases due to the greater dispersion of scores (variance). This increase in the SD decreases our ability to detect group differences using parametric statistics. (p.1783)

Essentially, then, because group comparisons that were analyzed using parametric statistics failed to detect reliable postoperative decrements in CPB patients, researchers in this area have abandoned conventional data analysis approaches and pursued methods that would support the researchers’ beliefs that at least some patients experienced enduring postoperative cognitive declines after CPB.

However, the movement away from parametric statistics by researchers in this area may have been the wrong decision. Failures of early studies that used parametric statistics to detect reliable cognitive declines beyond the 1st postoperative week may be due to the fact that researchers used statistical models (one-way analyses of variance [ANOVAs] and $t$ tests) that did not properly account for practice-related performance improvements that commonly occur when se-

¹ Some investigators required patients to decline by one SD on two or more measures of cognitive performance; other investigators used a 20% decline to define a significant change.
rial assessments are carried out. The problem is that practice effects and effects of other factors (e.g., cerebral damage caused by CPB) are superimposed. To appreciate this point, imagine that CPB produced a 10% decline in cognitive processes that support performance of a task but other cognitive processes that influence performance on the task improved by 15% because of practice, so a net postoperative improvement of 5% would be observed. Thus, practice-related performance improvements may conceal actual cognitive decline.

As mentioned in the earlier discussion of the RCI method, one method for estimating the magnitude of the practice effect involves testing non-CPB patients’ performances on cognitive tests using the same test—retest interval as that for the CPB group. A statistical model that is naturally suited for studying different groups of participants across time in a longitudinal fashion is the mixed factor (also called split-plot design) ANOVA. When this analysis model is used, the F test for the between- and within-subjects interaction compares the separate effects of group (CPB vs. controls) and repeated measures to each other, effectively controlling for practice effects. To determine whether postoperative differences in cognitive performance between CPB patients and controls were reliable, we adopted a split-plot data analysis model for the present study.

The primary goal of the present study was to determine whether CPB produces reliable cognitive declines that persist beyond the first few postoperative weeks. Additionally, we sought to determine whether certain aspects of cognitive performance were disrupted more than others. Previous studies have noted relatively high incidences of postoperative performance disruptions on instruments such as the trail-making, digit-symbol, and grooved-pegboard tasks (Hammeke & Hastings, 1988; Malone, Prior, & Scholtz, 1981; Raymond et al., 1984; Savageau, Stanton, Jenkins, & Klein, 1982). Impairments on these tasks, however, do not reveal much about which aspects of cognitive performance are disturbed by CPB because these impairments may reflect general psychomotor slowing or disruptions of decision making, visual field scanning, fine motor dexterity, short-term verbal memory, and visual-spatial memory. Therefore, we designed a test battery for the present study that decomposed the demands of these tasks and analyzed post-CPB cognitive performance changes on each element of the test battery in a stepwise manner.

Method

Participants

CPB Participants

All participants in the CPB group were patients at New Hanover Regional Medical Center, Wilmington, North Carolina, a 520-bed community and teaching hospital. Fifty-seven CPB patients were recruited into the study. Of the original 57, 39 CPB participants completed both testing sessions. Only data from participants who completed both sessions were included in the analyses presented below. The patients included in the present study underwent isolated elective coronary artery bypass graft surgery with CPB, and surgeries were performed by a single surgeon (Howard F. Marks, Jr.), who performs approximately 200 CPB surgeries annually. Data collection took place between May 1996 and May 1999. Patients were ineligible for the study if they had a previous cardiac operation, were visually impaired and/or could not read text, or were not fluent English speakers. Patients who met the inclusion criteria were either contacted by telephone or were met in their hospital room and invited to participate. All participants completed institutional review board consent (initially approved on May 24, 1996) and received a copy of the Mayo Clinic Heart Book (McGoon, 1993).

Control Participants

Fifty-five control participants were recruited through a senior citizen wellness program sponsored by the hospital. Forty-nine completed both sessions. The same exclusion criteria that applied to CPB patients were applied to control participants.

Table 1 shows demographic data for the patients and controls that completed both phases of the study. The groups were similar on every demographic variable except gender, with 69% women in the control group and 27% women in the CPB group.

Surgical Technique

General endotracheal anesthesia was administered, and the patient was taken into the operating room and had a Swan-Ganz catheter and radial arterial line placed. The patient was prepped and draped in a sterile field. The sternum was opened with a Stryker saw, and the patient was systemically heparinized. The pericardium was opened in an inverted T fashion from the dia-

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Demographic and Surgical Data for CPB Patients and Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
<td>Control</td>
</tr>
<tr>
<td>Sample size</td>
<td>49</td>
</tr>
<tr>
<td>Age</td>
<td>M</td>
</tr>
<tr>
<td>Male</td>
<td>61.8</td>
</tr>
<tr>
<td>Female</td>
<td>62.8</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>15</td>
</tr>
<tr>
<td>Female</td>
<td>34</td>
</tr>
<tr>
<td>Race</td>
<td>White</td>
</tr>
<tr>
<td>Male</td>
<td>45</td>
</tr>
<tr>
<td>Female</td>
<td>37</td>
</tr>
<tr>
<td>Handiess</td>
<td>Right</td>
</tr>
<tr>
<td>Male</td>
<td>43</td>
</tr>
<tr>
<td>Female</td>
<td>35</td>
</tr>
<tr>
<td>Education</td>
<td></td>
</tr>
<tr>
<td>&lt; high school</td>
<td>2</td>
</tr>
<tr>
<td>High school</td>
<td>20</td>
</tr>
<tr>
<td>College</td>
<td>21</td>
</tr>
<tr>
<td>&gt; college</td>
<td>6</td>
</tr>
<tr>
<td>CPB time (in minutes)</td>
<td>M</td>
</tr>
<tr>
<td>Male</td>
<td>119</td>
</tr>
<tr>
<td>Female</td>
<td>67</td>
</tr>
</tbody>
</table>

Note. CPB = cardiopulmonary bypass.

*Amount of time the ascending aorta is clamped shut during the bypass procedure.
phragm to the innominate vein and then tacked to each chest wall. The ascending aorta was cannulated with a curved plastic 21 French drainage line pump cannula. The right atrium was cannulated with a two-stage Sarns venous cannula, the tip of which was passed into the inferior vena cava. A needle cardioplegia catheter was placed in the ascending aorta, and the patient was placed on partial bypass. A left ventricular vent was placed through the right superior pulmonary vein, and the patient was placed on full bypass and cooled to 30 °C. The aorta was cross-clamped, and the heart was arrested with warm, followed by cold, sanguineous antegrade cardioplegia. Topical ice saline was used as an adjuvant to myocardial protection in order to drop the myocardial temperature to between 10 and 15 °C. Additional doses of cold retrograde cardioplegia were given every 20 min throughout the cross clamp as grafting was performed. A terminal dose of warm retrograde cardioplegia was given when grafting was almost complete. The cross clamp was removed, a partial occlusion clamp was placed, and grafts were then completed and de-aired. Chest tubes and ventricular and atrial epicardial pacing wires were placed. At 36.5 °C and on epinephrine, the patient was weaned from bypass. Protamine was given to reverse the heparinization, and the patient was decannulated. The chest was then closed and all incisions were dressed with Betadine ointment. The patient was then taken to the Cardiac Surgery Recovery Unit at New Hanover Regional Medical Center.

Psychological and Cognitive Performance Evaluations

After giving their written informed consent, participants completed a battery of psychological and cognitive performance evaluations. Each CPB patient completed the test battery in 24–48 hr prior to surgery and again in 3–4 weeks postoperatively. Control participants were tested at the same intervals as CPB patients. Testing was carried out in a small private office on the campus of either the hospital or the university.

Test Battery

A questionnaire was used to obtain demographic (age, gender, race, education, and handedness) and health-related information including a detailed medical history as well as current medical problems and/or medications. The state component of the Spielberger State-Trait Anxiety Inventory (STAI; Spielberger, Gorsuch, & Lushene, 1970) was administered to quantify anxiety at the time of testing.

The battery of cognitive performance evaluations consisted of seven instruments that included five computerized tests and two manually administered tests of auditory verbal memory. The computerized tests were administered using a Macintosh PowerBook 150. Most participants completed the entire evaluation in under 1 hr.

Simple reaction time. Simple reaction times were measured using commercially available software (MacLaboratory for Psychology; Research; Chute, 1994). Participants were instructed to press a designated key on the computer keyboard as quickly as possible each time the target stimulus appeared in the center of the computer screen. The target stimulus was a white circle, 34-cm in diameter. Each participant completed 20 total trials, 10 trials with each hand, alternating in blocks of 5 trials. The computer recorded response latencies with better than 10-ms resolution. The first 10 trials were practice trials. During these practice trials, reaction times were often relatively slow and quite variable but soon stabilized as participants became familiar with the task. Mean reaction time over the final 10 trials was used as the dependent measure.

Choice reaction time. Choice reaction times were measured using the same procedures and software described above. However, during the choice reaction time task, the computer was programmed to present the stimulus 30 cm either to the left or to the right of the center of the screen. The two screen locations used for stimulus presentations were marked in outline continuously throughout the task. The order of presentation locations was random. Participants were instructed to press the d key when stimuli were presented on the left and the k key when stimuli were presented on the right. Each participant completed 20 total trials, and reaction time was used as the dependent measure.

Visual attention. Visual attention was assessed using a commercially available program (Visual Field Attention; Brainmetric, 1991b). A black-and-white oval target (10 mm in height, 5 mm in diameter) appeared at random screen locations at an average rate of one target per second for 80 trials, with 20 targets appearing in each of the four screen quadrants. Participants were instructed to stare at a fixation point (a solid black dot 10 mm in diameter) located in the center of the screen and to press a designated button on the keyboard each time the target appeared. Reaction times were measured with better than 10-ms resolution, and participants’ mean reaction times were the dependent variable.

Visuomotor tracking. Visuomotor tracking skills were measured using a commercially available program (Chute, 1994). Participants were instructed to keep the cursor within a circular target (10 mm in diameter) while the target moved at a constant speed (15 mm/s) in a circle (88 mm in diameter) around the computer screen. Participants controlled the cursor using the trackball component of the computer keyboard. A tone provided auditory feedback when the cursor was off target. Each individual test trial consisted of one complete circuit around the track. Participants completed six trials with a 30-s interval between trials, and the computer recorded the total time off target during each trial. Visuomotor tracking task performances were characterized by steep acquisition functions over the first two or three trials. On the basis of inspections of each participant’s raw data, it was clear that the last three trials best represented participants’ asymptotic performances. Therefore, the mean time off target for the final three trials was computed for each participant and used as the dependent measure.

Visual spatial working memory. Visual spatial working memory was evaluated using commercially available software (Memory Match; Brainmetric, 1991a). Eighteen different computer-generated images were randomly assigned positions in a 6 × 6 matrix on the computer screen; each image was assigned two different positions. Initially, the cells of the matrix appeared as empty squares. Participants revealed the images by positioning the cursor over a cell in the matrix (using the computer trackball) and pressing a designated response button on the computer keyboard. On each trial, participants could select two cells in the matrix. If the images in the two selected cells were the same, then participants received credit for a hit, and the cells were removed from the matrix (i.e., the cells turned black and were inactivated). If the images in the two selected cells were different, then a miss was recorded, and the images disappeared (but remained in fixed locations within the matrix). Participants were instructed to maximize their hit–miss ratio, and they were told that accuracy was more important than speed on this task. The dependent variable was the ratio of hits to misses. Performance on this task, therefore, was related to participants’ abilities to remember the spatial positions of images.

Auditory verbal working memory. Auditory verbal working memory was evaluated using the Digit Span Backward subscale from the Wechsler Adult Intelligence Scale—Revised (Wechsler,
The researcher presented digits verbally, and participants were instructed to recall the digits and repeat them back in reverse order. The shortest digit series included two digits and the longest included seven digits. The two digit series was presented first, and two samples of each series length were administered until participants failed to recall at least one of the samples of a given series length in the correct order. The total number of correct sequences recalled was the dependent variable.

Auditory verbal paired-associate memory. Auditory verbal paired-associate memory was evaluated using the Word Paired-Associates subscale from the Wechsler Memory Scale—Revised (Wechsler, 1984). The test involves two components, an auditory presentation of eight word pairs that the participant was instructed to remember, and a cued-recall component in which one member of each pair is presented by the experimenter, and the participant’s task is to report the other member of the pair. Each correctly recalled pair added one point to a participant’s overall score on this task.

During the computerized task, the experimenter carefully monitored participants’ performances and permitted participants to start a task again if the participants misunderstood instructions or showed any evidence that they were performing poorly because of unfamiliarity with the task.

Data Analyses

Participants’ raw scores on each individual task were transformed to \( z \) scores on the basis of the control group’s first session performance. Only data from participants who completed both sessions were included. The SD used to compute all participants’ \( z \) scores was based on the mean of the control group during the first test session.

Statistical analyses were performed using the Statview statistical software package (Caldera et al., 1993). Separate repeated measures ANOVAs were performed on each dependent variable. Alphas were adjusted to control for multiple analyses using the Holm–Hochberg method (see Legendre & Legendre, 1998) in which the adjusted \( p \) values (\( p' \)) are computed from \( p' = (k - I + 1)p \), where \( k \) is the number of \( p \) values and \( I \) is the rank of a \( p \) value.

Results

Seventeen CPB patients and 6 control participants who completed the first test session did not return for the 1-month follow-up test session. Two CPB patients declined for reasons related to postoperative complications. The others gave no reason for declining. One CPB participant’s data were not included in the analyses because his postoperative scores on all of the cognitive measures were at least four SDs worse than the mean of the CPB group. Again, only data from participants who completed both sessions were used in the analyses presented below (CPB, \( n = 39 \); control, \( n = 49 \)).

Each panel in Figure 1 presents the results from one of the cognitive tasks and shows the mean of each group’s \( z \) scores as a function of session. Simple and choice reaction times were unaffected by CPB surgery (Panels A and B). Panel B appears to show main effects of both group and session on choice reaction times but these effects failed to reach statistical significance (\( ps = .16 \) and .04, respectively; the Holms-Hochberg adjusted alpha level was < .01). On the visual field attention test (Panel C), however, the control group improved during the second session relative to the first. Such practice effects were not observed in the CPB group. An ANOVA confirmed that this Group × Session interaction was significant, \( F(1, 82) = 7.49, p < .01 \). Pursuit rotor performance (Panel D) improved in both groups from the first session to the second. The main effect of session was significant, \( F(1, 84) = 9.60, p < .01 \). Visual spatial memory was unaffected by CPB surgery (Panel E). Panel F shows that Digit Span Backward scores were affected by CPB, with the control group outperforming the CPB group during the second session. An ANOVA revealed a significant Group × Session interaction that approached statistical significance, \( F(1, 83) = 5.80, p < .05 \). Finally, the control group outperformed the CPB group both pre- and postoperatively on the Word Paired-Associates task, and this observation was confirmed by a main effect of treatment, \( F(1, 83) = 8.70, p < .01 \).

The preoperative difference between the control and the CPB groups on the Word Paired-Associates memory task led to the question of whether that effect was due to the group differences in gender or education (see Table 1). To examine the influences of gender and education on paired-associates memory, we analyzed control participants’ preoperative scores as a function of these variables. Gender had no influence on paired-associates memory performance (\( p = .67 \)). Table 2 shows the group means on each test as a function of session presented in terms of their original units of measurement.

Anxiety and Cognitive Performance Self-Appraisal

Figure 2 shows data from the STAI. As would be expected, CPB patients were significantly more anxious than control participants during the preoperative, but not the postoperative, test session. An ANOVA revealed a significant Treatment × Session interaction, \( F(1, 84) = 10.16, p < .01 \).

Group Differences in Medication Use

Table 3 presents information regarding medication use by each group. Control participants reported no relevant changes in medication status between Session 1 and Session 2. Medications taken by control participants are shown in the first column of Table 3 and are expressed in terms of the percentage of the group that used each drug class. Medications taken by CPB participants are shown for the pre- and postoperative sessions separately because some potentially important changes in medication status occurred. During the first test session, 22% of the CPB participants were taking beta-adrenergic blockers and none of the control patients were. This raises the possibility that the preoperative impairment observed on the paired-associates learning task may have been due to effects of the beta-adrenergic blockers. To evaluate this possibility, we performed a t test (independent) comparing preoperative scores of CPB patients who took beta-adrenergic blockers with those of patients who did not. The t test failed to detect a statistically significant difference between the two groups, \( t(36) = 0.17 \).
We also analyzed the effects of beta-adrenergic blockers, antiarrhythmics, angiotensin converting enzyme (ACE) inhibitors, and narcotic analgesics on postoperative performance of CPB participants on postoperative performance on the visual attention task and the Digit Span Backward task using t tests (independent). No statistically significant differences were found (p = .13, .61, .30, and .91 for beta-adrenergic blockers, antiarrhythmics, ACE inhibitors, and narcotic analgesics, respectively, on the visual attention task; and p = .47, .99, .65, and .10, respectively, on the Digit Span Backward task). Interesting to note, in the two cases in which the effects approached statistical significance (beta-adrenergic blocker effects on visual attention and narcotic effects on the Digit Span Backward task performance), the trends favored participants who were taking the drugs over those who were not. Overall, these analyses show that the effects of these drug classes do not explain the pattern of cognitive impairment observed in the CPB group.

**Discussion**

Significant group differences in neuropsychological performance were present both pre- and postoperatively. Control participants outperformed CPB candidates on the Word Paired-Associates task during the first session. This finding was consistent with previous reports of auditory memory impairments in CPB candidates (O’Brien et al., 1992; Vingerhoets et al., 1997). Relatively worse performance on measures of verbal short-term memory implies that functionally significant cerebral damage may be present preoperatively in CPB candidates. Consistent with this hypothesis, magnetic resonance imaging studies have reported evidence of prior ischemic cerebral damage (i.e., cortical infarct, lacuna, or moderate-to-severe patchy or confluent white-matter hyperintensity) in 48% to 93% of CPB candidates (Moody et al., 1995; Simonson et al., 1994; Toner et al., 1993, 1994).

Although the CPB and control groups were not ideally matched with respect to gender, this variable did not seem to strongly influence performance on the Word Paired-Associates subscale. We were also unable to locate reports in the literature on aging and memory that documented the existence of gender differences in verbal memory in older individuals. However, preoperative state anxiety scores were higher in CPB patients than in control participants.
These group differences in anxiety levels should be taken into consideration when interpreting the relatively poor verbal memory scores of the CPB patients and may suggest a nonorganic basis for the preoperative memory impairments observed. Regardless of the source of the effect, the presence of preoperative cognitive impairment in CPB patients may have important implications for how postoperative group differences are interpreted, an issue that we revisit in the final part of this article.

One of the main findings of the present study was that control participants reliably outperformed CPB patients on the visual field attention and Digit Span Backward tasks during the second (i.e., postoperative) session. It is important to note that the cognitive performance differences between the two groups could not be attributed to general cognitive slowing or to general sensorimotor impairment in the CPB patients because performances on the simple reaction time and rotor pursuit tasks were not affected by CPB.

Postoperative performances on the visual field attention and Digit Span Backward tasks were also not associated with postoperative medication use by CPB patients. Finally, it is also important to note that these group differences were not due to a small subset of the patients that had relatively large postoperative performance declines. Recall that Stump et al. (2000) suggested that the CPB group’s SD increased postoperatively because of the greater dispersion of scores. With respect to this point, it should be noted that postoperative variances on the cognitive tests were approximately the same as those obtained during the preoperative session (see Table 2).

Two sources of evidence suggest that visual attention and Digit Span Backward performance share a subset of cognitive resources. First, impairments on the Digit Span Backward task have been identified as a concomitant of visual attention deficits in brain damaged individuals (Robertson, 1990). Second, factor analysis studies have reported strong correlations between the Digit Span Backward task and several measures of attention (Malec, Ivnik, & Smith, 1993). Taken together, these findings converge with those of the present study to suggest that brain systems that support attention may be particularly vulnerable to CPB surgery.

Both sustained visual attention and performance on the Digit Span Backward task depend on the intact functioning of frontal cortical regions (Gronwall, 1987). Visual-spatial attention also involves regions of parietal cortex (Posner, Cohen, & Rafal, 1982; Posner, Walker, Friedrich, & Rafal, 1987). In this regard, it is interesting that postoperative declines in glucose metabolism have been noted in the frontal and parietal (as well as inferior temporal and occipital) regions of the cerebral cortex 8 to 12 days after CPB surgery.

Table 2

<table>
<thead>
<tr>
<th>Test component</th>
<th>Control First session</th>
<th>Control Second session</th>
<th>CPB First session</th>
<th>CPB Second session</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Reaction time (in milliseconds)</td>
<td>369 14</td>
<td>365 14</td>
<td>361 12</td>
<td>364 38</td>
</tr>
<tr>
<td>Choice reaction time (in milliseconds)</td>
<td>441 14</td>
<td>417 12</td>
<td>464 21</td>
<td>454 16</td>
</tr>
<tr>
<td>Visual attention (in milliseconds)</td>
<td>468 21</td>
<td>416 16</td>
<td>447 18</td>
<td>454 21</td>
</tr>
<tr>
<td>Pursuit rotor (seconds off target)</td>
<td>6.96 0.60</td>
<td>5.89 0.60</td>
<td>8.00 0.80</td>
<td>6.88 0.70</td>
</tr>
<tr>
<td>Visual spatial memory (% correct)</td>
<td>32.3 12.0</td>
<td>33.0 13.0</td>
<td>31.6 15.0</td>
<td>32.6 17.0</td>
</tr>
<tr>
<td>Digit Backward (No. correct)</td>
<td>5.80 0.30</td>
<td>6.60 0.30</td>
<td>6.03 0.40</td>
<td>6.00 0.40</td>
</tr>
<tr>
<td>Paired-Associates (No. correct)</td>
<td>4.1 0.2</td>
<td>4.4 0.2</td>
<td>3.4 0.3</td>
<td>3.5 0.2</td>
</tr>
</tbody>
</table>

Note. All data are expressed in terms of the original units of measurement. CPB = cardiopulmonary bypass; Digit Backward = Digit Span Backward subscale; Paired-Associates = Word Paired-Associates subscale.

Figure 2. Mean Spielberger State-Trait Anxiety Inventory (STAI) scores for each group as a function of test session. Error bars represent standard error of the mean. CPB = cardiopulmonary bypass.
Because attention itself can be decomposed into several specialized subcomponents (Parasuraman, 1984; Posner & Cohen, 1984; Treisman & Gelade, 1980), the results of the present study suggest that a detailed analysis of the particular aspects of attention that are affected by CPB would be valuable.

One of the most important general findings from the present study is that traditional inferential statistical methods confirmed the reliabilities of persistent group differences between CPB patients and controls on individual measures of cognitive performance. Still, there are arguments that may be presented in favor of reporting incidences of postoperative decline rather than quantifying the overall reliabilities of group differences. One such argument holds that investigators should distinguish clinically meaningful performance declines from those that are not clinically significant (S. P. Newman, 1995; Stump, 1995). The distinction between clinical and statistical significance is a valid one and has an important place in all clinical outcomes research. However, it is also important to acknowledge that the criteria currently used to classify postoperative cognitive impairments (i.e., the SD and RCI methods) actually do not make such a distinction. Indeed, the task of specifying which cognitive changes are clinically meaningful does not have a simple quantitative solution. Small changes in performance in one cognitive domain may produce greater functional impairment in everyday life than do larger changes in another domain. The clinical significance of the cognitive changes that occur after CPB can only be determined by studies that analyze relationships between postoperative changes on particular measures of cognitive performance and corresponding changes on tasks the patients perform in their everyday lives.

With the goal of understanding the basic nature of the cognitive declines that occur after CPB, there is much to be gained from quantitative approaches that provide accurate estimates of effect sizes. When performances on the cognitive tests are expressed as continuous variables, both subtle and large performance changes contribute to the overall effect size. In contrast, when performances are categorized (e.g., impaired vs. improved), information about the effect size is lost and postoperative decrements that fail to reach the threshold to be classified as changed are essentially discarded.

At least two aspects of the present study limit the conclusions that can be reached, one of which should be of general concern to investigators who use serial assessments to evaluate cognitive performance changes over time. The first limitation of this study is that we cannot attribute postoperative declines in cognitive performance observed in the present study to factors that are unique to CPB surgery. Indeed, previous studies that have compared the effects of CPB surgery with effects caused by non-CPB thoracic surgeries have found similar changes in these two groups 8 weeks postoperatively (Smith et al., 1986).

A second limitation of the present study, which should be of more general concern to neuropsychologists and other investigators involved in neuropsychological outcome research, is related to the practice effect. As has typically been the case in studies on the effects of CPB on cognitive performance, in the present study, CPB patients’ scores on the Digit Span Backward and visual field attention tests did not actually decline postoperatively. They just did not improve as much as the control patients did (see Benedict, 1994; Stump et al., 2000, for further discussion of this issue). As mentioned earlier, control participants typically perform better the second time that they are assessed than they do during their initial baseline session.

Now, it has been assumed that CPB patients’ scores on the assessments would improve as a function of practice

---

**Table 3**

**Medications Used by Participants Expressed in Terms of Percentage of Group Taking the Drug Class**

<table>
<thead>
<tr>
<th>Drug class</th>
<th>Control</th>
<th>Preoperative</th>
<th>Postoperative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heart-related medications</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antiarrhythmic</td>
<td>5</td>
<td>6</td>
<td>31</td>
</tr>
<tr>
<td>ACE inhibitors</td>
<td>7</td>
<td>14</td>
<td>20</td>
</tr>
<tr>
<td>Beta-adrenergic blockers</td>
<td>0</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>Ca ions2 channel blocker</td>
<td>12</td>
<td>14</td>
<td>20</td>
</tr>
<tr>
<td>Anticoagulant</td>
<td>0</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Antihyperlipidemic</td>
<td>10</td>
<td>17</td>
<td>23</td>
</tr>
<tr>
<td>Diuretic</td>
<td>5</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>Vasodilator (anti-angina)</td>
<td>0</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>Other medications</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analgesics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Narcotic</td>
<td>0</td>
<td>3</td>
<td>17</td>
</tr>
<tr>
<td>Nonnarcotic</td>
<td>7</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Antibiotic</td>
<td>2</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>Benzodiazepine</td>
<td>2</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Hormone replacement</td>
<td>30</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>SSRI antidepressant</td>
<td>7</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Miscellaneous</td>
<td>30</td>
<td>23</td>
<td>30</td>
</tr>
</tbody>
</table>

*Note.* CPB = cardiopulmonary bypass; ACE = angiotensin converting enzyme; SSRI = selective serotonin reuptake inhibitor.
alone, like those of the healthier controls, if it were not for the loss in cognitive functioning caused by CPB (Kneebone et al., 1998; Slade et al., 2001). The present study, and prior studies (O’Brien et al., 1992; Vingerhoets et al., 1997), however, demonstrated that CPB candidates have preoperative learning impairments. Therefore, researchers should consider the possibility that cardiac surgery patients cannot learn, or retain until the follow-up session, the skills and knowledge that are responsible for the practice-related performance improvements that are typically observed in healthier controls. According to this account, the postoperative differences in cognitive performance between CPB patients and controls may just reflect preexisting learning impairment in CPB patients (or poor retention of skills learned during the preoperative session) rather than effects of CPB-related neuropsychological impairment. The preoperative learning deficits present in CPB patients should also raise further concerns about quantitative methods, such as the RCI, that assume that practice-related performance improvements in CPB patients are equivalent to those measured in a group of healthy control participants. In view of the available facts, the validity of such an assumption appears to be highly questionable. Clearly, the challenges associated with serial assessment continue to present significant problems for researchers studying CPB effects on cognitive performance. Fortunately, experiments could be designed to test the hypothesis that postoperative differences in cognitive performances between CPB patients and controls are due to preoperative learning impairments in CPB patients. One such experiment would include CPB and non-CPB participants, with a subset of participants from each group being tested both pre- and postoperatively, whereas the other participants would be tested only postoperatively. The proposed hypothesis predicts that the control participants who are familiar with the tests (i.e., tested two times) would outperform CPB patients during the second (i.e., postoperative) test session, as occurred in the present experiment. According to this hypothesis, however, control participants who are exposed to the test battery only once should not perform better than CPB patients would postoperatively.

References


Received January 3, 2001
Revision received August 23, 2001
Accepted December 5, 2001