A Cross-Validation Study of the Victoria Revision of the Category Test

Jennifer J. Kozel
University of South Dakota

John E. Meyers
Northwest Iowa Rehab Consultants and University of South Dakota

Short form versions of the Category Test have been shown to be effective tools for assessing brain injury in relatively short amounts of time. The purpose of the current study was to compare the Victoria Revision of the Category Test with the full booklet version. The Victoria Revision short form was administered as part of a neuropsychological battery to two sets of participants matched for diagnosis. Results showed that the Victoria Revision is as effective in identifying brain injury as the full Category Test. The authors conclude that the Victoria Revision is a viable alternative to the full Category Test. In this time of managed health care, and in consideration of the fatigability of brain injured subjects, a valid short version of the Category Test is clinically relevant and useful.

Adams and Trenton (1981) and Kilpatrick (1970) have indicated that the Halstead Category Test (HCT) is nearly as effective at gleaning out the presence or absence of brain damage as the entire Halstead-Reitan battery. However, since its development many researchers and practitioners have reported that the test is too time consuming and inefficient to be useful in a clinical setting (Erickson, Calsyn, & Scheubach, 1978; Kilpatrick, 1970; Sherrill, 1987). Finlayson, Sullivan, and Alfano (1986) studied the length of time it took for groups of severely head-injured patients to complete the HCT. The HCT took 29 minutes to complete for 26% of their sample, 30 to 39 minutes for 32%, 40 to 59 minutes for 35%, and more than an hour for the remaining 7%. Although Finlayson et al. concluded that these administration times were acceptable, many disagreed with this conclusion. Sherrill (1987) for instance, pointed out that it is difficult to characterize a test as optimal when 74% of participants take 30 minutes or more to finish. This may be especially true for a test that is frequently administered to impaired individuals with lower frustration tolerance and higher fatigability than unimpaired individuals.

Because of this controversy, multiple studies have focused on developing a shortened form of the HCT. Kilpatrick (1970) began this pursuit by assessing the split-half reliability

Address correspondence to John E. Meyers, Northwest Iowa Rehab Consultants, Suite 340, 500 Jackson Street, Sioux City, IA 51102-3268.
of the HCT. He split the 208-item HCT into odd and even items, thereby creating two 104 item tests. He found that the short form error rates correlated highly \((r = .97)\), and each one also correlated highly with the whole test, \((r = .99\) and \(r = .90\), respectively). Kilpatrick concluded that a short form of the HCT was a feasible option, however, he pointed out that the use of fewer items increased the potential for errors by patients. Consequently, the error score cutoff point would need to be adjusted to accommodate the new learning curve.

Following up on this study by Kilpatrick (1970), Boyle (1975) further reduced the short form of the HCT to 84 items. Using brain-damaged and non-brain-damaged participants, he confirmed Kilpatrick’s conclusion that a short form provided a reliable measure that could be used to differentiate brain-damaged and non-brain-damaged participants. This version reduced administration time to 15 minutes and used a cutoff point of 38. Six percent of the brain-damaged participants and 22% of the non-brain-damaged participants were misclassified. However, he believed the test could be made even shorter by eliminating subtests V and VI because the two tests either separately or together, do not reliably discriminate brain-injured from non-brain-injured participants.

Gregory, Paul, and Morrison (1979) created a 120-item version of the HCT (Cat-120). This group dropped items from the end of subtests and completely removed some of the later subtests. Their test correlated highly with the HCT \((r = .95)\). Similarly, Calsyn, O’Leary, and Chaney (1980) devised a 120-item test for which they reported a correlation of .89 with the total score of the HCT. Golden, MacInnes, Kuperman, and Moses (1981) supported Cal-
syn et al.’s findings by cross-validating the use of the Cat-120 item version on four matched groups (brain-damaged, non-brain-damaged, psychiatrically diagnosed and normal outpatient participants). Once again, each of the group comparisons correlated highly with Calsyn et al.’s original score of .89.

Since Kilpatrick (1970) and Boyle (1975), other investigators have explored various meth-
ods for abbreviating the HCT. The number of items on these subsequent tests has varied as each investigator sought to establish short forms that correlated highest with the full HCT. Barker (1977) constructed a 125-item short form by using item analysis to identify items for inclusion in a short form of the HCT. To make the test more accessible, he made photographic prints of each of the items and placed them in booklet form. Barker used a cutoff point of 45 and found that his test yielded fewer misclassifications than did the HCT with its original cutoff point of 50. However, Barker’s conclusions for this form were reached after administering the test to only 30 participants. Moreover, attempts to cross validate the short form were not possible since it was difficult to ascertain from Baker’s description which items were included or excluded (Summer & Boll, 1987).

Since the mode of item presentation differed in the booklet version of the short form, it needed to be evaluated for its comparability to the traditional form. Summers and Boll (1987) investigated the comparability of a smaller card short form of the HCT. Both tests were administered to 50 normal college students 1 week apart and in counterbalanced order (Sum-
mers & Boll, 1987). Results showed high split-half reliability for both the short form and the traditional form (.81 and .89, respectively). This led the authors to conclude that there were no significant differences between the two versions.

Delfilippis and McCambell (1991) generated and published a booklet version of the original HCT. The original HCT and the Booklet format showed a .91 correlation \((p < .001)\). In a study comparing six different versions of the HCT, Ivins and Gillman (1995) found no differences between the original HCT, the Booklet Category Test, and three different computer versions of the original HCT.

Russell and Levy (1987) devised a 95 item test called the Revised Category Test which correlated well with the HCT \((r = .97)\). This version proved to be time efficient, and has kept the nature of the task consistent with the traditional form. Taylor, Hunt, and Glaser
Russell and Levy \((r = .96)\). Researchers have also been concerned with the short form’s ability to generalize across groups. Pierce et al. (1990) used groups of participants varying in age, gender, and educational level. They found that none of these factors affected the short form’s ability to predict actual scores on the HCT. There has also been concern with the short form’s ability to discriminate laterality of brain injury. Taylor, Goldman, Leavitt, and Kleinmen (1984) unveiled what appeared to be the first drawback of the short forms—the inability to differentiate right and left hemisphere injuries and to assess the degree of damage. They acknowledged that the short form was still a valuable tool because it saves time and frustration, but they advised practitioners using the instrument to be aware of the possibility of misdiagnosis. Conversely, Sherrill (1987) proposed that the short forms were, in fact, clinically efficient in their ability to judge degrees of impairment. Moreover, he demonstrated that the skepticism about whether or not the short form was testing the same cognitive abilities as the HCT (i.e., problem solving) was unfounded.

Sherrill (1987) set out to establish the accuracy of his proposals by using a short form developed by Labreche (1983). LaBreche’s short form discarded all memory items from the HCT (subtest VII), discarded the formal scoring of subtest I and II (because the only purpose of these subtests was to orient patients to the test procedures), and took out all items that were redundant in content. This resulted in a short form of 81 items for scoring and 5 for orienting. Labreche named the test the Victoria Revision and demonstrated that the test correlated very well with the HCT \((r = .96)\). Moreover, the Victoria Revision yielded a classification rate of 84%, which was comparable to the 83% classification rate of the HCT. Sherrill (1987) also cross-validated the Victoria Revision on a sample of neuropsychological referrals and determined that “having the smallest number of scored items of any of the attractive short forms, the Victoria Revision probably offers the best combination of relatively short administration time and relatively good predictive accuracy for routine clinical use” (p. 350). From this study, Sherrill developed a formula for predicting HCT performance \(4.335 + (1.839 \times (\text{Victoria Revision errors})) = \text{HCT error score}\).

The culmination of these studies suggests that it is the content of the HCT, and not its presentation, that is sensitive to brain injury. These studies also support the use of the short versions and booklet formats of the HCT. The purpose of the current study was to cross-validate the booklet format of the Victoria Revision of the HCT. In addition, Sherrill’s (1987) regression equation will be used to predict scores of this short form with the HCT. The hypothesis is that there will be no difference between the Victoria Revision of the HCT and the full booklet version of the HCT.

**METHODS**

**Participants and Procedures**

Four groups of individuals referred for neuropsychological evaluation participated in this study. Two groups \((ns = 43)\) consisting of individuals with a diagnosis of closed head injury due to a motor vehicle accident (MVA). Two other groups \((ns = 15)\) consisted of individuals with a clinical diagnosis of probable Alzheimer’s disease as determined by board certified neurologists. Participants in the Alzheimer’s groups were diagnosed as having probable Alzheimer’s if they had impaired neuropsychological functioning (excluding the Category Test), no cerebrovascular accident (CVA) indicated on their computed tomography (CT)/magnetic resonance imaging (MRI), had scores on the Blessed Dementia Scale (Lezak, 1995) of >4,
and had scores on the Ischemic Scale (Lezak, 1995) of <4. Two groups (1 Alzheimer and 1 MVA) were randomly assigned to receive the Victoria Revision and the other two groups were assigned to receive the full version. Both versions were given as part of a neuropsychological battery. All but one of the participants were caucasian, the remaining demographic and clinical features of the participants are summarized in Table 1.

Each of the scores for the Victoria Revision of the HCT were derived using Sherrill’s (1987) regression equation. Scores were converted to T Scores using normative data from Heaton, Grant, and Mathews (1991). For participants between 15 and 19 years of age, the Heaton et al. (1991) norms for 20-year-olds were used. All other participants scores were converted to T scores using the procedure identified in Heaton et al. (1991). This allowed for consistency in using the same normative set for all data comparisons.

RESULTS

Independent samples t-tests were used to compare the two versions of the test for the groups with matching diagnoses. The mean T Scores for the MVA Victoria Revision group was 41.88 (SD = 13.03) and the mean T score for the full form group was 41.79 (SD = 10.34). The two groups were not significantly different (p > .05). The two MVA groups did not differ significantly (p > .05) on age, education, or WAIS-R FSIQ. The average length of unconsciousness (for those who had a loss of consciousness) for the Victoria Revision was 5.14 (SD = 11.89) days and for the full test group it was 3.83 (SD = 8.06) days; this difference was not significant (p > .05). Data on Glasgow Coma Scales or Post Traumatic Amnesia was not available on all subjects and so was not included in the analysis. The two groups did differ on length of time since injury. The group that received the full version of the HCT was assessed 20.79 (SD = 32.81) months postinjury. The Victoria Revision Group was assessed 30.37 (SD = 76.79) months postinjury. This difference in the groups was significant (p = .046), but did not appear to adversely influence performance on the other measures. Therefore, the finding is noted but not considered a significant factor.

For the Alzheimer’s groups, the mean score for the Victoria Revision was 30.13 (SD = 5.28) and the mean score for the full form group was 27.93 (SD = 8.07). There were no significant difference between the two comparison groups in the category scores, age, education, or WAIS-R FSIQ scores (all p’s > .05).

DISCUSSION

The review of the literature shows consistently that the original HCT and its derived alternate formats do not differ significantly in effectiveness for assessing brain injury. Additionally, the short forms developed from the original version, or the booklet version, do not appear significantly different in effectiveness in identifying brain-injured individuals.

In the current study, the booklet format of the Victoria Revision was given to two groups of matched subjects, and compared with matched groups who received the full-length booklet version of the HCT. Results show that the Victoria Revision when given as part of a neuropsychological battery, is as effective at identifying brain injury as the full length test.

The advantage of using the Victoria Revision is the shortened administration time (15 to 20 minutes). In this time of managed health care, time savings without loss of sensitivity is important in improving the cost effectiveness of neuropsychological services. Therefore, when the need arises to shorten the battery of neuropsychological testing, whether due to time constraints by insurers or to the fatigue of patients, the use of the Victoria Revision is an appropriate alternative to the full HCT.
### TABLE 1

Description of Participants

<table>
<thead>
<tr>
<th>Group</th>
<th>Number of</th>
<th>Males</th>
<th>Females</th>
<th>Age M</th>
<th>Age SD</th>
<th>Age Range</th>
<th>Years of Education M</th>
<th>Years of Education SD</th>
<th>FSIQ M</th>
<th>FSIQ SD</th>
<th>Number of Right-Handed</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>MVA Victoria Revision</td>
<td>43</td>
<td>29</td>
<td>14</td>
<td>30.47</td>
<td>13.02</td>
<td>15–58</td>
<td>13.16</td>
<td>3.10</td>
<td>93.09</td>
<td>14.61</td>
<td>38</td>
<td>40.88</td>
<td>13.03</td>
</tr>
<tr>
<td>MVA Full Form</td>
<td>43</td>
<td>32</td>
<td>11</td>
<td>35.23</td>
<td>17.61</td>
<td>17–79</td>
<td>12.56</td>
<td>2.24</td>
<td>92.26</td>
<td>11.91</td>
<td>40</td>
<td>41.79</td>
<td>10.34</td>
</tr>
<tr>
<td>Alzheimer’s Victoria Revision</td>
<td>15</td>
<td>8</td>
<td>7</td>
<td>74.40</td>
<td>8.63</td>
<td>54–87</td>
<td>12.07</td>
<td>2.49</td>
<td>79.53</td>
<td>11.58</td>
<td>14</td>
<td>30.13</td>
<td>5.28</td>
</tr>
<tr>
<td>Alzheimer’s Full Form</td>
<td>15</td>
<td>8</td>
<td>7</td>
<td>77.87</td>
<td>6.93</td>
<td>67–88</td>
<td>11.40</td>
<td>1.55</td>
<td>77.93</td>
<td>14.91</td>
<td>13</td>
<td>27.93</td>
<td>8.07</td>
</tr>
</tbody>
</table>

$M =$ mean; $n =$ number; $SD =$ standard deviation.
Anecdotally, subjects in the Alzheimer’s groups who received the Victoria Revision compared to the full version appeared to the evaluator as less “upset” during the testing and fewer complained of the length of the testing. Although this is not quantifiable data, it is believed that the Victoria Revision may be more easily accepted by older participants than is the full-length test.

REFERENCES


