The Recognition Memory Test
Examination of ethnic differences and norm validity

Sid E. O’Bryant\textsuperscript{a,b,∗}, Robin C. Hilsabeck\textsuperscript{a,c},
Robert J. McCaffrey\textsuperscript{b}, Wm. Drew Gouvier\textsuperscript{a}

\textsuperscript{a}Louisiana State University, Baton Rouge, LA, USA
\textsuperscript{b}Department of Psychology, University at Albany, State University of New York,
Social Sciences 112, Albany, NY 12222, USA
\textsuperscript{c}Texas Tech University Health Sciences Center, Lubbock, TX, USA

Accepted 18 October 2001

Abstract

The possibility of racial bias in neuropsychological test materials has received increasing attention in recent years. The purpose of the present study was to investigate whether an own-race recognition bias would provide an advantage for Caucasian participants over African American participants on the Faces subtest of the Recognition Memory Test (RMT). Thirty Caucasian and 30 African American undergraduates completed the RMT, Shipley Institute of Living Scale (SILS), and Symbol Digit Modalities Test (SDMT). No significant group difference was found on RMT Faces. However, mean RMT Faces scores for both groups were below the 10th percentile in spite of average scores on the SDMT and SILS. A second study was conducted to further examine the validity of the RMT norms for this age range (i.e., 18–24) and to provide 2-week test–retest reliabilities. The mean RMT Faces subtest score was 39.78 (10th percentile), and 28% of the sample scored at or below the fifth percentile. Test–retest reliabilities were .63 and .64 for RMT Words and Faces, respectively. Results of these studies suggest that re-examination of the current norms for RMT Faces is warranted for adults aged.

© 2001 National Academy of Neuropsychology. Published by Elsevier Science Ltd. All rights reserved.

Keywords: Recognition Memory Test; Face recognition; Test–retest reliability

\textsuperscript{∗} These data were presented at the 19th Annual Meeting of the National Academy of Neuropsychology, San Antonio, TX.

\textsuperscript{∗} Corresponding author. Tel.: +1-518-442-5804.
E-mail address: so2377@csc.albany.edu (S.E. O’Bryant).

0887-6177/01/$ – see front matter © 2001 National Academy of Neuropsychology.
PH: S0887-6177(01)00189-5
1. Introduction

The Recognition Memory Test (RMT; Warrington, 1984) is used to assess material-specific recognition memory deficits in adults ages 18–70. The RMT consists of two subtests: Recognition Memory for Words (RMW) and Recognition Memory for Faces (RMF). Each subtest is comprised of 50 target stimuli, which are displayed for 3 s and rated as pleasant or unpleasant by the examinee. Next, the examinee is shown each of the 50 target stimuli paired with a distracter and asked to point to (or read aloud) the target stimulus. Advantages of the RMT include its ease of administration and scoring, lack of reliance on verbal responses, minimal demands on other cognitive functions (e.g., attention, organization, and motor functioning), and relatively low vulnerability to the effects of depression and anxiety (Warrington, 1984).

Although the RMT offers these advantages and is widely used, concerns about the basic psychometric properties of the test have been raised. As noted by Kapur (1987), no information about internal consistency or test–retest reliability of the RMT was reported in the manual, interexaminer reliability has not been established, and ceiling effects, especially for younger subjects, were evident on RMW. More recently, its validity as a measure of lateralized memory dysfunction has been questioned (Hermann, Connell, Barr, & Wyler, 1995; Kelly, Johnson, & Govern, 1996; Kneebone, Chelune, & Jueders, 1997; Naugle, Chelune, Schuster, Lueders, & Comair, 1994; Sweet, Demakis, Ricker, & Millis, 2000).

In addition to the above psychometric concerns, it is possible that a racial bias exists on the RMT because the RMF subtest consists of White male faces only. Studies of facial recognition have routinely demonstrated an own-race recognition bias (Brigham & Barkowitz, 1978; Chance, Goldstein, & McBride, 1975), with few exceptions (Lindsay & Wells, 1983). An own-race recognition bias is suggested when a racial group exhibits a superior recognition for own-race faces than for other-race faces (Barkowitz & Brigham, 1982). This bias has been shown to occur equally in Caucasian and African American subjects, accounting for 10 and 11% of the variance, respectively, in recognition ability (Bothwell, Brigham, & Malpass, 1989). Therefore, the purpose of this study was to investigate whether an own-race recognition bias would provide an advantage for Caucasian participants over African American participants on the RMF subtest of the RMT. We hypothesized that Caucasian examinees would score higher than African American examinees on the RMF but that there would be no significant group difference on RMW.

2. Study 1

2.1. Method

2.1.1. Participants

Sixty undergraduates at a large southern university participated in this experiment in exchange for extra credit in their psychology courses. Every participant signed written informed consent to participate in the study. Thirty of the participants were Caucasian (5 males and 25 females) and 30 were African American (3 males and 27 females). The average age of the Caucasian group was 19.77 years (S.D. = 1.87), and the average age of the African American
group was 20.07 years (S.D. = 2.63). Average education was 13.13 years (S.D. = 1.20) and 13.17 years (S.D. = 1.05) for Caucasian and African American participants, respectively. There were no significant group differences in age [t(58) = -0.51, \( P = .61 \)], education [t(58) = -0.12, \( P = .91 \)], or gender (\( \chi^2 = 0.58, P = .45 \)). No participant had a history of neurological or psychological problems, and all denied significant histories of alcohol and drug abuse.

2.1.2. Measures
2.1.2.1. RMT (Warrington, 1984). As noted above, the RMT assesses RMW and RMF. Internal consistency of the RMT was recently established in a sample of patients with traumatic brain injuries (Malinda, Bowers, Millis, & Uekert, 1998). Cronbach’s \( \alpha \)'s for RMW and RMF were .86 and .77, respectively. Administration time of the RMT is approximately 15 min.

2.1.2.2. Shipley Institute of Living Scale (SILS; Zachary, 1986). This test is commonly used as a quick method of obtaining an estimated WAIS-R Full Scale IQ score. The SILS consists of two subtests, Vocabulary and Abstraction, and typically requires approximately 15 min to administer. Correlations with WAIS-R Full Scale IQ have been found to be as high as .79, with the highest correlations obtained from samples with at least 10 years of education (Lezak, 1995). The SILS was administered in this study to establish the overall cognitive ability of our sample and to ensure the comparability of the groups.

2.1.2.3. Symbol Digit Modalities Test (SDMT; Smith, 1982). The SDMT provides a brief measure of attention, visual scanning and tracking, and psychomotor functioning. It is one of the most sensitive measures of organic dysfunction (Lezak, 1995) and was administered in this study as an additional means of establishing cognitive functioning and group comparability. The SDMT requires about 5 min to administer.

2.1.3. Procedure
Participants were administered the RMT, SILS, and SDMT in a counterbalanced order. All measures were administered according to their standardized instructions.

2.2. Results
Means and standard deviations of each measure by group are presented in Table 1. On the SILS, both groups obtained estimated WAIS-R Full Scale IQ scores in the average range. There was no significant group difference in estimated IQ [t(58) = 1.09, \( P = .28 \)]. Both groups performed within normal limits for their age and education on the SDMT, and there was no significant group difference [t(58) = 1.60, \( P = .12 \)]. These results suggest that both groups were cognitively intact at the time of testing.

As hypothesized, there was no significant difference between the two groups on RMW [t(58) = 1.49, \( P = .29 \)]. However, contrary to our hypothesis, there was no significant difference between Caucasian and African American participants on RMF [t(58) = 0.18, \( P = .67 \)]. Unexpectedly, both groups obtained mean RMF scores below the 10th percentile, with 13 of the 30 participants (43%) in both the Caucasian and African American groups scoring at or
Table 1
Means and standard deviations of test scores by group

<table>
<thead>
<tr>
<th>Test</th>
<th>Caucasian group</th>
<th>African American group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>S.D.</td>
</tr>
<tr>
<td>SILS</td>
<td>102.77</td>
<td>5.76</td>
</tr>
<tr>
<td>SDMT</td>
<td>63.97</td>
<td>10.70</td>
</tr>
<tr>
<td>RMW</td>
<td>48.40</td>
<td>1.38</td>
</tr>
<tr>
<td>RMF</td>
<td>37.97</td>
<td>4.34</td>
</tr>
</tbody>
</table>

below the fifth percentile. The RMF means of the Caucasian and African American groups were significantly lower than the normative data provided for their age group \(t(29) = -6.49, P < .001\) and \(t(29) = -6.10, P < .001\), respectively, and the combined RMF mean score of 38.2 (\(n = 60\)) was also significantly lower than the normative score provided in the manual \(t(59) = -8.97, P < .001\). In contrast, both groups scored near the ceiling on the RMW subtest.

2.3. Discussion

There was no significant difference in performance between Caucasian and African American participants on the RMF subtest, failing to support the hypothesis that an own-race recognition bias would give Caucasian participants a performance advantage over African American participants. One explanation for the lack of an own-race recognition advantage in our study is offered by the contact hypothesis (Ng & Lindsay, 1994). The contact hypothesis states that increased exposure to other races will increase the individual’s ability to discriminate between other-race individuals (Chiroro & Valentine, 1995; Goldstein & Chance, 1985; Lavarakas, Buri, & Mayzner, 1976). On a large university campus, such as the university attended by these participants, it is likely that students obtain frequent exposure to individuals of different ethnic groups and races. This exposure could have decreased, or even eliminated, the own-race recognition bias in our African American participants.

Another explanation for the lack of an own-race recognition advantage could be due to the nature of the RMF task. Recent research by Levin (1996, 2000) has suggested that deficits in cross-race facial recognition are not due to an inability to code details of cross-race faces but to a tendency to place emphasis on the broad category of race rather than details of individual faces (i.e., the feature selection hypothesis). In the case of a memory task, examinees are instructed to remember each stimulus, which involves focusing on the details of that stimulus. Therefore, on the RMF, it is likely that African American participants paid more attention than usual to the individuating details of the faces because they knew they would be asked to remember them, later, thus eliminating a cross-race recognition deficit.

A third possibility for the present study’s finding of no own-race recognition advantage involves the nature of the RMF stimuli. The photographs in the RMF are black-and-white photographs, which may eliminate some distinguishing characteristics of the faces (e.g., eye color, hair color, and skin tone). This neutralizing of the stimuli may make remembering the photographs equally difficult for both Caucasian and African American examinees.
An unexpected finding in the present study was that average RMF scores of both the Caucasian and African American groups were below the 10th percentile for persons their age, and 43% percent of both groups scored at or below the fifth percentile. These poor performances on the RMF were unexpected given that our sample was screened for past and present psychological and neurological problems, and their performances on the SILS and SDMT were indicative of intact cognitive functioning. Further, both groups scored within normal limits on the RMW subtest. Given that the present sample size was nearly twice that of the original normative sample (i.e., \( n = 60 \) and 31, respectively), and there was no significant difference between our cognitively intact Caucasian and African American groups, this finding suggests that a re-examination of the RMF norms for persons aged 18–24 is warranted. Thus, the primary purpose of the second study was to replicate the findings of Study 1.

3. Study 2

To investigate the possibility that our finding of poor RMF performances by American undergraduates was due to an unknown artifact, we targeted this same age group in our second study. A secondary goal was to examine the 2-week test–retest reliability of the RMT, which was not provided in the manual for any age group. Currently, the only test–retest data available in the literature are from a diverse group of 40 neurologically impaired outpatients (Soukup, Bimbela, & Schiess, 1999). A test–retest reliability coefficient of .81 on RMF was obtained in this group over an average interval of 7 months. These patients were not administered RMW, so no test–retest information for this subtest was available.

3.1. Method

3.1.1. Participants

Eighteen individuals between the ages of 18 and 24 participated in Study 2. Every participant provided written informed consent prior to participation. Fifteen were undergraduates who participated in exchange for extra credit in their psychology courses, and three were individuals from the community who participated so a friend or family member could receive extra credit in a psychology course. Three of the participants were male and 15 were female. Thirteen were Caucasian, four were African American, and one was Asian American. The average age of this sample was 20.33 years (S.D. = 2.14), and the average education was 13.39 years (S.D. = 1.09). All participants denied current or past neurological or psychological problems and significant histories of alcohol or drug abuse.

3.1.2. Measures

3.1.2.1. Kaufman Brief Intelligence Test (KBIT; Kaufman & Kaufman, 1990). The KBIT is a measure of intelligence comprised of two subtests, Vocabulary and Matrices. The Vocabulary subtest is a composite of Definitions and Expressive Vocabulary (i.e., naming). Standard scores are generated for Vocabulary and Matrices, and the sum of these two scores results in a composite IQ score. The KBIT has demonstrated good internal consistency and test–retest reliability and has been shown to be a valid measure of intelligence (Spreen & Strauss, 1998).
Administration time is approximately 20–30 min. As with the SILS in the first study, the KBIT was administered to help establish the overall cognitive ability of our sample.

3.1.3. Procedure

All participants were administered the RMT and KBIT in a counterbalanced order on two occasions, 2 weeks apart. Both measures were administered according to standardized instructions.

3.2. Results

Average KBIT standard scores on the first administration were 101.06 (S.D. = 8.65) for Vocabulary, 100.56 (S.D. = 11.38) for Matrices, and 100.83 (S.D. = 8.49) for Composite IQ. All of these scores were in the average range and suggestive of intact general cognitive ability in this sample. On the second KBIT administration, average standard scores were 106.78 (S.D. = 8.43), 102.61 (S.D. = 8.43), and 105.17 (S.D. = 8.42) on Vocabulary, Matrices, and Composite IQ, respectively. Participants demonstrated a significant improvement in standard scores on Vocabulary and Composite IQ \[t(17) = -5.72 \text{ and } -4.33, \text{ respectively, } P < .001\] but not on Matrices \[t(17) = -2.06, P = .27\]. The 2-week test–retest reliabilities of Vocabulary, Matrices, and Composite IQ were .88, .73, and .83 (\(P < .001\) for all three).

On the first administration of the RMT, mean raw scores were 47.56 (S.D. = 2.83) and 39.78 (S.D. = 3.84) on RMW and RMF, respectively. Five of the 18 participants (28%) scored below the fifth percentile on RMF. The RMF score was significantly lower than that of the normative group \[t(17) = -3.67, P = .002\]. The distribution of RMF raw scores is presented in Figure 1. On the second administration of the RMT, average raw scores for RMW and RMF were 48.72 (S.D. = 1.67) and 43.78 (S.D. = 3.99), respectively. A significant improvement in raw scores was seen on both RMW and RMF \[t(17) = -1.17 \text{ and } -4.00, P < .05\] and

![Figure 1](image-url)
.001, respectively]. Two-week test–retest correlation coefficients were .63 for RMW and .64 for RMF ($P < .01$ for both)

3.3. General discussion

Results of Study 1 failed to reveal that an own-race recognition bias led to better performance by Caucasian participants on the RMF subtest of the RMT compared to African American participants. Potential reasons for failing to find this typically robust effect may be explained by the contact hypothesis, feature selection hypothesis, and/or neutralizing effects of black-and-white photographs. An own-race recognition advantage may have been eliminated in the present study since our African American participants were frequently exposed to Caucasians as students at a large, ethnically diverse university, supporting the contact hypothesis (Ng & Lindsay, 1994). Also, since examinees were instructed to remember the faces, they may have focused on individuating details more than they would have in other situations, which is consistent with the feature selection hypothesis (Levin, 2000). Alternatively, the black-and-white photographs used in the RMF may have eliminated distinguishing facial characteristics, making it equally difficult for Caucasian and African American participants to remember individual faces.

What Study 1 did reveal, however, was unexpectedly low group means for both Caucasian and African American participants on RMF (37.97 and 38.43, respectively) in spite of overall intact cognitive functioning on the SILS and SDMT and normal RMW scores. These poor RMF performances led to a replication study (i.e., Study 2), which revealed a similarly low mean RMF score (i.e., 39.78), with over a quarter of the sample scoring at or below the fifth percentile. The low RMF scores in two samples of neurologically intact adults aged 18–24 ($n = 78$) calls into question the validity of the RMF norms for this age group. The current normative information provided in the RMT manual for this age range may be too high, which would help explain the limited floor and unexpected impaired RMF performances of patients with normal CT scans found by Kelly et al. (1996). Results of other studies failing to employ a control group may also need reinterpretation if the sample was comprised predominantly of young adults aged 18–24. The low RMF scores of our two samples when compared to the normative data may be due to the nature of the RMF stimuli.

Study 2 also examined the 2-week test–retest reliabilities of the RMW and RMF, which were not provided in the RMT manual. Test–retest coefficients were found to be disappointingly low at .63 and .64 for RMW and RMF, respectively. Although the sample size yielding these coefficients was small ($n = 18$), the 2-week test–retest reliabilities for the KBIT were strong (i.e., .88 for Vocabulary, .73 for Matrices, and .83 for composite IQ), indicating adequate power for this purpose. Thus, the moderate test–retest coefficients found for the RMW and RMF may be due to unstable test properties.

In contrast to the current test–retest results, Soukup et al. (1999) found a 7-month test–retest coefficient of .81 for the RMF in a heterogeneous sample of 40 neurological outpatients. While the sample and test–retest interval used by Soukup et al. is more clinically relevant than those used in the current study, it is important to establish the test–retest reliabilities produced by neurologically intact individuals so practice effects can be calculated and taken into account when interpreting serial RMT performance. Therefore, more research examining test–retest reliabilities of the RMT is recommended.
One possible limitation to this study should be noted. In both studies, the samples were predominately female, which prevented us from examining the influence of gender on RMF performance. However, previous research with undergraduate students, patients with traumatic brain injuries, neurological outpatients, and elderly community volunteers have consistently found no gender differences on the RMF portion of the RMT (Diesfeldt & Vink, 1989; Iverson & Franzen, 1994; Millis, 1992; Soukup et al., 1999). Therefore, it is unlikely that a gender effect is responsible for the present findings.

In conclusion, an own-race recognition bias did not provide an advantage for Caucasian participants over African American participants on the RMF subtest of the RMT. However, there did appear to be a discrepancy between the RMF performances of neurologically intact adults aged 18–24 in the United States and the hospital inpatients in London used in the standardization sample. The specific reasons for this performance discrepancy are unclear but are likely due to differences in the samples. It is possible that the cues offered by the RMF stimuli (i.e., hair cuts, clothing, etc.) are more relevant to London participants than to American participants, making it easier for the London participants to remember the faces on this portion of the RMT. Results of the current studies suggest that a re-examination of the norms using a larger, more representative sample of persons in this age group is warranted. A final caveat should be noted when using the RMT as a forced-choice measure of effort or malingering (Iverson & Franzen, 1994; Millis, 1992, 1994). The unexpectedly poor RMF performance of our sample, in the presence of intact intellectual and neurocognitive functioning, cautions against interpreting poor scores as necessarily indicative of poor effort.

References


