

# Response times for the CAGE, Short-MAST, AUDIT, and JELLINEK Alcohol Scales

RON D. HAYS, LAURAL HILL, JAMES J. GILLOGLY,  
MATTHEW W. LEWIS, and ROBERT M. BELL  
*RAND, Santa Monica, California*

and

RONALD NICHOLAS  
*Center for Counseling, Tarzana, California*

The CAGE alcohol screening tool has been touted as a good choice for clinical settings because of its brevity. We administered the CAGE and three other alcohol screening instruments (the Short-MAST, AUDIT, and JELLINEK) by microcomputer to 296 clients at a drinking driver treatment program and three of the four scales to a second sample of 270 clients from six drinking driver treatment programs. The average response times for the CAGE were 31 and 32 sec, respectively, in the first and second samples. The average response time for the JELLINEK was approximately five times longer than it was for the CAGE; response time for the AUDIT averaged four times longer; and response time for the Short-MAST was two and a half times as long. The estimated reliability of the CAGE was the lowest and its standard error of measurement was the highest of the four scales. We recommend the Short-MAST as the tool of choice if the extra minute of administration time it requires in comparison with the CAGE is not critical.

The recognition of problematic alcohol use is essential for providers of health care, but one half of the patients with diagnosable problems go undetected by medical providers (Coulehan, Zettler-Segal, Block, McClelland, & Schulberg, 1987). Practical (i.e., reliable, but brief) methods of detection are needed to help identify individuals with alcohol disorder (Hoffmann et al., 1989; Nichol & Ford, 1986).

The most popular alcohol screening tools include the CAGE (Mayfield, McLeod, & Hall, 1974) and the Michigan Alcoholism Screening Test (MAST) (Selzer, 1971). The CAGE consists of only 4 items. The short form of the MAST includes 13 items (Selzer, Vinokur, & Van Rooijan, 1975). Sensitivities and specificities of the MAST and the CAGE to alcohol problems have been found to compare fairly, with a slight edge to the MAST overall (Hays & Spickard, 1987; Magruder-Habib, Durand, & Frey, 1991). However, some reviewers of the literature have concluded that the brevity of the CAGE makes it a good choice for alcohol screening in primary care practices (Hays & Spickard, 1987; Magruder-Habib et al., 1991). "When different measures of the same dimension of drinking are about equally reliable and accurate, the more economical or efficient measure could be used"

(Tucker, Vuchinich, Harris, Gavornik, & Rudd, 1991, p. 154).

Aside from the number of items that they contain, the response burden of these alcohol screening measures has not yet been documented. Indeed, the number of items is commonly used in survey research as the barometer of burden. For example, in an approximately 1-h-long self-administered survey of chronic disease patients, Sherbourne and Meredith (1992) reported that the youngest patients completed an average of 4 items per minute, but the oldest patients completed an average of 2.7 items per minute. Sherbourne and Meredith (1992) treated all items equivalently in their analysis and relied on self-reports of the time it took to complete the entire questionnaire.

In this study, we use microcomputer-assessed response time to quantify the relative burden of the CAGE and the Short-MAST. We also examine their intercorrelation to determine whether the Short-MAST provides different information than the CAGE. In addition, we compare these popular instruments with the Alcohol Use Disorders International Test, AUDIT (Babor & Grant, 1989), and a new alcohol screening measure that we have developed, the JELLINEK (based on Jellinek, 1952).

## METHOD

### Subjects

**Sample 1.** Two hundred and ninety-six clients (216 males, 80 females) at a west coast treatment center for impaired ("drinking") drivers were sampled. The average age of the respondents was 34 years, with a range from 19 to 75. Eighty-one percent of the sample were non-Hispanic white, 90% were high school graduates, and

---

Preparation of this article was supported in part by Grant AA 07852 from the National Institute on Alcohol Abuse and Alcoholism. The opinions expressed are those of the authors and do not necessarily reflect the views of RAND or the sponsor. Thanks are due Karen Boseker for outstanding secretarial support. Correspondence should be addressed to R. D. Hays, RAND, 1700 Main Street, Santa Monica, CA 90407-2138.

23% were married and living with their spouses. The yearly household income was below \$30,000 for 53% of the sample.

Fifty-four percent of the respondents were first offenders and 46% were multiple offenders. The participants reported an average of 1.6 convictions in the last 7 years for driving under the influence and an average blood alcohol concentration of 0.17 at the time of the last arrest, and 24% reported one or more automobile accidents during the last 12 months. In the sample, 27% categorized themselves as alcoholic, 18% as problem drinkers, 34% as social drinkers, and 14% as light drinkers; 7% reported drinking only on special occasions, and 1% reported that they never drank alcohol. (The few clients who reported never drinking alcohol were reporting this status since their arrest for impaired driving.)

**Sample 2.** Two hundred and seventy clients (213 males, 57 females) from six west coast treatment centers for impaired drivers were sampled. The average age of the respondents was 33, with a range from 18 to 64. Sixty percent of the sample were non-Hispanic white, 84% were high school graduates, and 20% were married

and living with their spouses. The yearly household income was below \$30,000 for 64% of the sample.

Fifty-four percent of the respondents were first offenders and 46% were multiple offenders. The clients reported an average of 1.7 arrests in the last 7 years for impaired driving and an average blood alcohol concentration of 0.16 at the time of the arrest, and 21% reported one or more automobile accidents during the last 12 months. In the sample, 23% categorized themselves as alcoholic, 8% as problem drinkers, 8% as heavy drinkers, 39% as social drinkers, and 10% as light drinkers; 10% reported drinking only on special occasions, and 2% reported that they never drank alcohol.

**Measurement**

The Microcomputer Assessment System (MAS) is a data collection program that allows for the design and computer administration of surveys by manipulating ASCII text files (Hays et al., 1992). A 125-item interview consisting of a brief reading assessment, socio-demographic items, and several questions about alcohol and drug

**Table 1**  
**New Alcohol Screening Measure Based on**  
**Jellinek Stage Model of Alcoholism**

- 
1. Which of the following statements are TRUE of you? (Select all that apply.)
    - I DRINK TO CALM MY NERVES.
    - I FEEL LIKE CONTINUING TO DRINK WHEN OTHERS HAVE STOPPED.
    - I FEEL UNCOMFORTABLE IN SITUATIONS WHEN ALCOHOL IS NOT AVAILABLE.
    - I DRINK ALCOHOL TO GET RELIEF.
    - I HAVE MEMORY LAPSSES AFTER DRINKING.
    - I THINK ABOUT OR LOOK FORWARD TO THE TIMES WHEN I CAN DRINK ALCOHOL.
    - I GET IRRITATED WHEN MY DRINKING IS DISCUSSED BY OTHERS.
    - MY TOLERANCE TO ALCOHOL HAS INCREASED.
    - NONE OF THE ABOVE
  
  2. Which of the following statements are TRUE of you? (Select all that apply.)
    - I HIDE LIQUOR OR SNEAK DRINKS.
    - I HAVE BECOME MORE DEPENDENT ON ALCOHOL.
    - I FEEL GUILTY ABOUT MY DRINKING.
    - I HAVE MEMORY BLACKOUTS AFTER DRINKING.
    - I HAVE TREMORS IN THE MORNING.
    - I FAIL TO KEEP MY PROMISES AND RESOLUTIONS.
    - I HAVE BECOME MORE DISHONEST WITH OTHERS.
    - I BECOME MORE AGGRESSIVE AND GRANDIOSE WHEN I DRINK.
    - I HAVE LOST INTEREST IN THINGS OTHER THAN DRINKING.
    - I HAVE BEEN UNABLE TO DISCUSS PROBLEMS WITH OTHERS.
    - I AM HAVING FAMILY, WORK OR MONEY PROBLEMS.
    - I NEGLECT EATING.
    - I AVOID FAMILY OR FRIENDS.
    - I DRINK ALONE.
    - I AM IN TROUBLE AT WORK OR HAVE LOST MY JOB.
    - I LIE ABOUT MY DRINKING.
    - NONE OF THE ABOVE
  
  3. Which of the following statements are TRUE of you? (Select all that apply.)
    - I HAVE DETERIORATED PHYSICALLY BECAUSE OF MY DRINKING.
    - I HAVE LOST MY WILL POWER AS A RESULT OF MY DRINKING.
    - I HAVE FELT A NEED TO DRINK ALCOHOL IN THE MORNING.
    - I EXPERIENCE REMORSE AFTER DRINKING.
    - I HAVE HAD MEMORY LOSS AND IMPAIRED THINKING.
    - I HAVE HAD FAMILY PROBLEMS BECAUSE OF MY DRINKING.
    - I HAVE LESS TOLERANCE TO ALCOHOL THAN BEFORE.
    - I HAVE GOTTEN DRUNK SEVERAL TIMES.
    - I HAVE RECEIVED HOSPITAL TREATMENT FOR A DRINKING PROBLEM.
    - I HAVE UNSPECIFIC FEARS ABOUT LIFE.
    - I HAVE BEEN UNABLE TO GET GOING AND CHANGE MY LIFE FOR THE BETTER.
    - I HAVE BEEN UNABLE TO WORK.
    - I FEEL RESENTMENTS TOWARD OTHERS.
    - THERE ARE MANY REASONS I DRINK.
    - I HAVE GIVEN UP ON DRINKING ALCOHOL.
    - NONE OF THE ABOVE
-

use was administered with the use of the MAS to clients in Sample 1; a 112-item interview was administered to clients in Sample 2. The average time to complete the microcomputer interview was 30 min ( $SD = 11$ ; range, 9-96) in Sample 1 and 26 min in Sample 2 ( $SD = 7$ ; range, 17-56). The reading level of the 125-item interview (item stems and responses) was judged to be approximately 6th-grade level by the UNIX style program (according to Kincaid, Coleman-Liau, and Flesch readability indices).

### Alcohol Screening Scales

The CAGE acronym stands for the 4 yes/no items that constitute this screener (Cut down, Annoyed, Guilty, Eye-Opener). Item responses are scored 0 or 1, with a higher score an indication of alcohol problems. A total score of 2 or greater is considered clinically significant. The 13 items in the short-form MAST are also presented with yes/no response choices and are often scored with unitary weighting of the items (Skinner, 1979). The MAST items focus on symptoms of problematic drinking and negative consequences of alcohol use. A score of 3 or greater is the clinical cutpoint for this measure. The AUDIT is a 10-item instrument with varying number of response options per item, ranging from three to six choices. The responses are differentially weighted, so that 0-4 points are possible per item. A score of 8 or greater is suggestive of alcohol problems. Finally, the JELLINEK consists of 3 "select all that apply" items with a different number of possible selections for each item (8, 15, and 16). The JELLINEK items, shown in Table 1, are unit-weighted, and a score of 3 or greater is the clinical cutpoint. The JELLINEK was administered to Sample 1 clients only; the other scales were administered to both samples.

## RESULTS

Raw and 0-100 linear transformed means, standard deviations, observed ranges, reliability estimates, and response times for the alcohol screening scales are provided in Table 2. The linear transformation of raw scores to 0-100 scales facilitates comparisons across scales and does not influence interval-level properties of the measures. Internal consistency reliability estimates for the four alcohol scales were 0.64 or higher. These reliabilities are comparable to those reported in previous studies for iden-

**Table 3**  
Intercorrelations Among Alcohol Scales

| Scale      | CAGE | Short-MAST | AUDIT | JELLINEK |
|------------|------|------------|-------|----------|
| CAGE       | 1.00 | 0.68       | 0.57  | 0.52     |
| Short-MAST | 0.73 | 1.00       | 0.65  | 0.66     |
| AUDIT      | 0.64 | 0.64       | 1.00  | 0.60     |

Note—Correlations from Sample 1 are provided above the diagonal ( $n = 296$ ). Correlations from Sample 2 are given below the diagonal ( $n = 270$ ). JELLINEK was administered only to Sample 1. All correlations are statistically significant ( $p < .001$ ).

tical or similar measures when administered by paper-and-pencil (Meier & Lambert, 1991; Selzer et al., 1975; Zung, 1979), although the estimated internal consistency reliabilities of both the CAGE (Fleming & Barry, 1991) and the Short-MAST (Hays & Revetto, 1992) were notably higher in some previous studies.

The standard error of measurement was largest for the CAGE, with the other three scales having similar values that were each smaller than that of the CAGE. In contrast, response times for the 4-item CAGE were shortest (8 sec per item), followed by the 13-item Short-MAST (6-7 sec per item), the 10-item AUDIT (12-13 sec per item), and the "3-item" JELLINEK (49 sec per item).

Correlations among the four alcohol screening measures, given in Table 3, reveal from 27% to 53% shared variance between pairs of measures. The CAGE and Short-MAST were the most strongly intercorrelated, and the CAGE and JELLINEK were the most distinct pair. As expected, first offenders had significantly lower CAGE ( $r_s = -.24$  and  $-.27$ ,  $p < .01$ ), Short-MAST ( $r_s = -.35$  and  $-.32$ ,  $p < .01$ ), AUDIT ( $r = -.16$ ,  $p < .01$ ; and  $r = -.14$ ,  $p < .05$ ), and JELLINEK ( $r = -.18$ ,  $p < .01$ ) scores than did clients in the multiple offender programs. In addition, the number of convictions for driving under the influence of alcohol during the past 7 years

**Table 2**  
Descriptive Statistics, Reliabilities, and Response Times (in Seconds)  
for Alcohol Scales

| Measure           | Raw |     | Observed Range | Alpha Reliability | 0-100 Transformed |      |      | Response Time |
|-------------------|-----|-----|----------------|-------------------|-------------------|------|------|---------------|
|                   | M   | SD  |                |                   | M                 | SD   | SEM  |               |
| <b>CAGE</b>       |     |     |                |                   |                   |      |      |               |
| Sample 1          | 1.8 | 1.3 | 0-4            | 0.66              | 46.1              | 32.6 | 19.0 | 31            |
| Sample 2          | 1.9 | 1.3 | 0-4            | 0.63              | 48.6              | 31.6 | 19.2 | 32            |
| <b>Short-MAST</b> |     |     |                |                   |                   |      |      |               |
| Sample 1          | 4.5 | 3.2 | 0-13           | 0.84              | 34.4              | 24.6 | 9.8  | 81            |
| Sample 2          | 4.3 | 3.2 | 0-13           | 0.84              | 33.2              | 24.7 | 9.9  | 95            |
| <b>AUDIT</b>      |     |     |                |                   |                   |      |      |               |
| Sample 1          | 9.9 | 7.7 | 0-35           | 0.83              | 24.8              | 19.1 | 7.9  | 123           |
| Sample 2          | 9.6 | 7.4 | 0-31           | 0.82              | 23.9              | 18.5 | 7.8  | 133           |
| <b>JELLINEK</b>   |     |     |                |                   |                   |      |      |               |
| Sample 1          | 5.3 | 6.3 | 0-30           | 0.86              | 13.6              | 16.3 | 6.1  | 148           |

Note—SD, standard deviation; SEM, standard error of measurement. Possible range and clinical cutpoints (in parentheses) for screener scales are as follows: CAGE is 0-4 (>1), Short-MAST is 0-13 (>2), AUDIT is 0-40 (>7), and JELLINEK is 0-39 (>2).

was significantly positively correlated with CAGE ( $r_s = .33$  and  $.31, p < .01$ ), Short-MAST ( $r_s = .40$  and  $.38, p < .01$ ), AUDIT ( $r_s = .19$  and  $.18, p < .01$ ), and JELLINEK ( $r = .30, p < .01$ ) scores. Self-reported blood alcohol concentration at the time of the arrest for impaired driving was also positively related to CAGE ( $r_s = .20$  and  $.19, p < .01$ ), Short-MAST ( $r = .28, p < .01$ ; and  $r = .17, p < .05$ ), AUDIT ( $r_s = .20$  and  $.21, p < .01$ ), and JELLINEK ( $r = .17, p < .05$ ) scores.

## DISCUSSION

The average response time for the Short-MAST was 2.5 times longer than that for the CAGE. Response times for the AUDIT and the JELLINEK were longer than those for both the Short-MAST and the CAGE. Moderate correlations were observed among all four alcohol screener instruments. The Short-MAST and CAGE were the most strongly intercorrelated, sharing about half their variance in common. The estimated reliability of the CAGE was notably lower and its standard error of measurement much higher than those of the other three alcohol screening measures. Correlations of the Short-MAST with indicators of drinking status (whether the client was in a first offender or multiple offender program, number of convictions for driving under the influence of alcohol, and blood alcohol concentration at the time of the impaired driving arrest) were all statistically significant and tended to be higher than corresponding correlations for the CAGE.

The results of this study indicate that different alcohol screening scales yield similar, but clearly not identical, information. The higher reliability of the Short-MAST than of the CAGE is consistent with the fact that the former has three times as many items as the latter. The somewhat stronger correlations of the Short-MAST with other indicators of drinking status accord with previous research favoring the Short-MAST over the CAGE for detecting alcohol disorder (Magruder et al., 1991). When selecting an alcohol screening measure for clinical applications, the user needs to decide whether the 50–63 sec in administration time saved by using the CAGE as opposed to the Short-MAST offsets the cost of lower reliability, greater error of measurement, and possibly diminished sensitivity to alcohol problems.

Future evaluations of response times for the Short-MAST, CAGE, AUDIT, and other alcohol screening scales in other samples will help to quantify the relative response burden of these alternative instruments in a wider range of respondents than those studied here. The results of this study lead us to recommend the Short-MAST as the tool of choice if the extra minute of administration

time it requires in comparison with the CAGE is not critical.

## REFERENCES

- BABOR, T. F., & GRANT, M. G. (1989). From clinical research to secondary prevention: International collaboration in the development of the Alcohol Use Disorders Identification Test (AUDIT). *Alcohol Health & Research World*, *13*, 371-374.
- COULEHAN, J. L., ZETTLER-SEGAL, M., BLOCK, M., MCCLELLAND, M., & SCHULBERG, H. C. (1987). Recognition of alcoholism and substance abuse in primary care patients. *Archives of Internal Medicine*, *147*, 349-352.
- FLEMING, M. F., & BARRY, K. L. (1991). A three-sample test of a masked alcohol screening questionnaire. *Alcohol & Alcoholism*, *26*, 81-91.
- HAYS, R. D., GILLOGLY, J. J., HILL, L., LEWIS, M. W., BELL, R. M., & NICHOLAS, R. (1992). A Microcomputer Assessment System (MAS) for administering computer-based surveys: Preliminary results from administration to clients at an impaired-driver treatment program. *Behavior Research Methods, Instruments, & Computers*, *24*, 358-365.
- HAYS, R. D., & REVETTO, J. P. (1992). Old and new MMPI-derived scales and the Short-MAST as screening tools for alcohol disorder. *Alcohol & Alcoholism*, *27*, 685-695.
- HAYS, J. T., & SPICKARD, W. A. (1987). Alcoholism: Early diagnosis and intervention. *Journal of General Internal Medicine*, *2*, 420-427.
- HOFFMANN, N. G., HARRISON, P. A., HALL, S. W., GUST, S. W., MABLE, R. J., & CABLE, E. P. (1989). Pragmatic procedures for detecting and documenting alcoholism in medical patients. *Advances in Alcohol & Substance Abuse*, *8*, 119-131.
- JELLINEK, E. M. (1952). Phases of alcohol addiction. *Quarterly Journal of Studies on Alcohol*, *13*, 673-684.
- MAGRUDER-HABIB, K., DURAND, M., & FREY, K. A. (1991). Alcohol abuse and alcoholism in primary health care settings. *Journal of Family Practice*, *32*, 406-413.
- MAYFIELD, D., MCLEOD, G., & HALL, P. (1974). The CAGE questionnaire: Validation of a new alcoholism screening instrument. *American Journal of Psychiatry*, *131*, 1121-1123.
- MEIER, S. T., & LAMBERT, M. E. (1991). Psychometric properties and correlates of three computer aversion scales. *Behavior Research Methods, Instruments, & Computers*, *23*, 9-15.
- NICHOL, E. F., & FORD, M. J. (1986). Use of the Michigan Alcoholism Screening Test in general practice. *Journal of the Royal College of General Practitioners*, *36*, 409-410.
- SELZER, M. L. (1971). The Michigan Alcoholism Screening Test: The quest for a new diagnostic instrument. *American Journal of Psychiatry*, *127*, 1653-1658.
- SELZER, M. L., VINOKUR, A., & VAN ROOIJAN, L. (1975). A self-administered Short Michigan Alcoholism Screening Test (SMAST). *Journal of Studies on Alcohol*, *36*, 117-126.
- SHERBOURNE, C. D., & MEREDITH, L. S. (1992). Quality of self-report data: A comparison of older and younger chronically ill patients. *Journal of Gerontology*, *47*, S204-S211.
- SKINNER, H. A. (1979). A multivariate evaluation of the MAST. *Journal of Studies on Alcohol*, *40*, 831-844.
- TUCKER, J. A., VUCHINICH, R. E., HARRIS, C. V., GAVORNIK, M. G., & RUDD, E. J. (1991). Agreement between subject and collateral verbal reports of alcohol consumption in older adults. *Journal of Studies on Alcohol*, *52*, 148-155.
- ZUNG, B. J. (1979). Psychometric properties of the MAST and two briefer versions. *Journal of Studies on Alcohol*, *40*, 845-859.