Journal of Occupational Health Psychology
2006, Vol. 11, No. 1, 119–134

1076-8998/06/$12.00 DOI: 10.1037/1076-8998.11.1.119

Screening Older, Blue-Collar Workers for Drinking Problems: An Assessment of the Efficacy of the Drinking Problems Index

Peter A. Bamberger
Technion Israel Institute of Technology

William J. Sonnenstuhl
Cornell University

Dana Vashdi
Technion Israel Institute of Technology

The authors examine the performance of the Drinking Problems Index (DPI) as a screening instrument for the identification of drinking problems among older blue-collar workers. Performance was assessed using a random sample of 1055 retirement-eligible blue-collar workers with the CAGE as a problem-drinking criterion. The authors also assessed the relative performance of the DPI versus the CAGE on the basis of each instrument's sensitivity, specificity, and area under the receiver operating characteristic curve (AUROC) with regard to other alcohol-related criteria. DPI was found to offer reasonable levels of sensitivity and specificity, although its sensitivity varied by age. Moreover, for older, blue-collar workers, the DPI was found to more effectively screen for problematic patterns of alcohol consumption than the CAGE.

Keywords: Drinking Problems Index, problem drinking among older workers, alcohol screening, blue-collar workers and drinking

The American workforce is aging. Older workers (i.e., age 55+) accounted for more than 16% of the civilian labor forces in 2004 (BLS, 2005) and are expected to account for nearly 20% of the labor force by the end of this decade (Rix, 2001). Retirement-eligible and officially retired workers constitute a significant portion of this older workforce. Although 50% of U.S. workers officially retire (i.e., draw retirement benefits) by age 60, only 11% have fully withdrawn from the workforce by that time (Doeringer, 1990). Indeed, many of those reporting to have fully withdrawn from the workforce (i.e., not working even on a part-time basis) may be expected to subsequently reenter the active workforce in the context of part- or full-time "bridge employment" prior to their permanent withdrawal from the labor force (Kim & Feldman, 2000).

For many of these older workers, and particularly those employed in blue-collar occupations, alcohol-related problems remain a key health concern (Bacharach, Bamberger, Sonnenstuhl, & Vashdi, 2004; Peterson & Zwerling, 1998). Community surveys estimate the prevalence rate of problem drinking among older Americans as running as high as 15% (Adams, Barry, & Fleming, 1996; Blow, 1998). However, Peterson and Zwerling (1998, p. 285) report that more than 20% of the blue collar males (ages 51–61) responding to the 1992 Health and Retirement Survey consumed three or more drinks daily,[1] putting them at risk of developing drinking problems and alcohol-related illnesses. Indeed, many of the health problems for which older individuals seek health care—harmful drug interactions, dementia, injuries, liver disease, cardiovascular disease, and

Peter A. Bamberger, Davidson Faculty of Industrial Engineering and Management, Technion: Israel Institute of Technology, Haifa, Israel, and Smithers Institute, ILR School, Cornell University, Ithaca, New York; William J. Sonnenstuhl, Department of Organizational Behavior, Smithers Institute, ILR School, Cornell University, Ithaca, New York; and Dana Vashdi, Davidson Faculty of Industrial Engineering and Management, Technion: Israel Institute of Technology, Haifa, Israel.

Authors' names appear in alphabetical order. All authors contributed equally. We gratefully acknowledge the grant support of the National Institute on Alcohol Abuse and Alcoholism (Grant 2 R01 AA011976). We also thank Samuel B. Bacharach, Paul Feigin, and Ayala Cohen for comments on an earlier version of this article.

Correspondence concerning this article should be addressed to Peter A. Bamberger, Davidson Faculty of Industrial Engineering and Management, Technion: Israel Institute of Technology, Haifa 32000, Israel. E-mail: peterb@techunix.technion.ac.il

[1] This is equivalent to approximately 33 or more grams of ethanol per day. To put this in context, in the United States, on average, nonabstaining blue-collar and white-collar male workers consume, respectively, approximately 16 and 12 grams of ethanol per day (Parker & Harford, 1992).

sleep disturbances—are influenced by drinking (Gambert & Katsoyannis, 1995). These alcohol-related health problems account for more than $60 billion a year in hospital related costs (Schonfeld & Dupree, 1995). Annual health care expenditures for alcohol-related problems amount to $22.5 billion (Rice, 1999). Over the next 25 years, the number of older people in the health care system with alcohol-related problems will nearly double (Gfroerer, Penne, Pemberton, & Folsom, 2003, 2002, 2003, 2002).

The relatively high prevalence and serious consequences of problem drinking among older individuals, combined with the potential for effective treatment of older, problem drinkers (Barnea & Teichman, 1994; Fitzgerald & Mulford, 1992) have led a number of researchers to advocate the screening of older individuals in general (Blow, 1998; Buchsbaum, Buchanan, Welsh, Centor, & Schnoll, 1992), and older workers in high-risk (e.g., blue-collar) occupations in particular (Bacharach et al., 2004; Trice & Roman, 1972), for drinking problems. However, estimating the prevalence of problem drinking among older, high-risk workers, no less identifying those with potential alcohol problems, may present a significant challenge given questions raised in the literature regarding the validity of many of the more widely used screening instruments with regard to particular high-risk populations (Bisson, Nadeau, & Demers, 1999; Clements, 1998). In this article we examine the criterion validity of a screening instrument developed specifically for the identification of alcohol-related problems among older individuals—the Drinking Problems Index (DPI), and examine its performance relative to one of the more widely used instruments—the CAGE—in identifying alcohol-related problems among older, blue-collar workers.

Evaluating the Efficacy of Alcohol Screening Instruments

Researchers and occupational health practitioners use a variety of screening questionnaires to identify drinking problems, including the CAGE, MAST, TWEAK, and AUDIT (Cherpitel, 1997; Connors & Volk, 2004).[2] All of these instruments identify drinking problems by focusing on the adverse consequences associated with the consumption of alcohol. However, for a number of reasons, the validity of these instruments for identifying high-risk populations or for screening members of suspected high-risk groups (such as older, blue-collar workers) for alcohol-related problems may be questionable.

First, for the most part, the criterion validity of these instruments has been assessed using only clinical samples of hospitalized, emergency room/ambulatory, or primary care patients admitted or being treated for disorders other than alcoholism (Babor, Higgins-Biddle, Saunders, & Monteiro, 2001). These clinical studies have largely confirmed the validity of these measures as a screening instrument for alcoholism in clinical settings (Bush, Shaw, Cleary, Delbanco, & Aronson, 1987; Liskow, Campbell, Nickel, & Powell, 1995). However, although such clinical designs are advantageous in that they incorporate a criterion measure that is itself highly valid, namely the *clinical* (typically DSM-based) diagnosis of an alcohol problem, they are problematic in that the results generated on the basis of the clinical samples studied may have only limited ability to generalize to the broader population (Conigrave, Hall, & Saunders, 1995). Thus, although these instruments have been widely used to estimate the prevalence of problem drinkers both in the general population (Bisson et al., 1999), and among particular high-risk subgroups such as blue-collar and other types of workers (Bacharach, Bamberger, & Sonnenstuhl, 2002; Martin, Blum, & Roman, 1992), few studies have assessed their validity for such purposes using nonclinical or population-based samples. Those studies that *have* done so have typically grounded their conclusion of criterion validity on the basis of ecological correlations alone. For example, in two cases (Alvarez & Del Rio, 1994; Smart, Adlaf, & Knoke, 1991), the criterion validity of the CAGE was determined on the basis of the similarity of the prevalence estimate generated by the screening device with the prevalence of heavy drinking in the same population at approximately the same time but estimated on the basis of some other instrument in another study.

Second, the findings generated by the handful of *population* studies applying standard measures of criterion validity to assess such screening instruments suggest that the efficacy of these instruments may be quite limited. For example, using a representative sample of household members in Quebec and assessing criterion validity on the basis of the sensitivity and specificity of the instrument with respect to such criteria as heavy drinking and alcohol-related health problems, Bisson et al. (1999) found the CAGE to be

[2] CAGE refers to Cut down, Annoyed, Guilty, and Eye-opener. MAST stands for the Michigan Alcohol Screening Test. TWEAK refers to Tolerance, Worry, Eye-opener, Amnesia, "Kut down." AUDIT stands for the Alcohol Use Disorders Identification Test.

largely unable to discriminate between heavy and nonheavy drinkers.

Finally, findings from both clinical and population-based studies suggest that the criterion validity of these instruments may vary across subgroups and that screening questionnaires with efficacy in one population may not necessarily offer screening efficacy in others (Cherpitel, 1997). For example, using a clinical design, Bradley, Boyd-Wickizer, Powell, and Burman (1998) found that the sensitivities of CAGE and AUDIT varied by race and gender. Similarly, population studies have found the screening efficacy of the CAGE to vary across subgroups. For example, several studies (Clements, 1998; Heck & Williams, 1995) found the CAGE to offer limited validity among suspected high-risk populations such as college students, and Bisson et al. (1999) found the sensitivity of the CAGE to be consistently lower for women than for men.

In this context, it is possible that the ability of many of the widely used screening instruments to identify heavy or problem drinkers may also be age contingent. Indeed, believing that other screening instruments do not sufficiently "tap adverse consequences from drinking that are more likely to occur among older adults," Finney, Moos, and Brennan (1991) designed the DPI, a 17-item questionnaire to assess drinking problems among older adults. Among the adverse consequences relevant to older adults tapped by the DPI but largely neglected by other instruments are falls, isolation as a result of drinking, and neglecting one's appearance. In addition, relative to other screening instruments, the DPI may more effectively tap the adverse consequences of particular relevance to older individuals as a result of its focus on the psychological reasons for drinking (e.g., to forget one's worries), viewed by some as particularly relevant when assessing drinking problems among older adults (Hwalek & Richter, 1983).

Finney et al. (1991) report that the DPI has high levels of internal reliability and construct validity and is highly correlated with alcohol consumption. Nevertheless, several questions remain about its use. First, no study has tested their general proposition that the DPI is more effective than other screening instruments for identifying drinking problems among older individuals in general, and older workers in particular. Second, research is needed to identify the most effective DPI cutoff for identifying older workers with drinking problems. Currently, researchers use a two-point cutoff (i.e., positive responses to 2 of the 17 questions) to identify problem drinkers (Moos, Schutte, Brennan, & Moos, 2004), but there is little evidence demonstrating than this cutoff is better than a one-, three-, or four-point cutoff. Third, no attempt has been made to compare the instrument's ability to detect drinking problems at varying cutoff levels with that of other screening instruments.

Using a population-based rather than a clinical approach, the current study examines the performance of the DPI in screening for drinking problems and problematic alcohol consumption patterns among older, blue-collar workers. Given the impracticality of using clinically based indicators of drinking problems in the context of a population-based research design, we use the CAGE as a proxy criterion. CAGE has been widely validated (Mayfield, McLeod, & Hall, 1974), correlating significantly with a clinical diagnosis of alcoholism (Ewing, 1984) and showing a strong association with other clinically based indictors of alcohol dependence such as the DSM–IV (Martin & Roman, 1996).

With the CAGE as a criterion, we first assess the sensitivity, specificity, positive predictive value, and negative predictive value of the DPI at varying cutoffs (Connors & Volk, 2004). Sensitivity is the proportion of people testing positive relative to those actually having the condition (whether correctly identified as such as well or incorrectly identified as not having the condition) who test positive. The sensitivity of a test describes its ability to correctly identify those people with the condition (i.e., true positive / [true positive + false negative]). Specificity is the proportion of those truly without the condition who in fact test negative. Specificity relates to the ability of the test to identify and leave aside those people who in fact do *not* have the condition (i.e., true negative / [true negative + false positive]). Positive predictive value is the proportion of people with positive tests (true and false) who actually have the condition. It measures how well the test rules in the condition (i.e., true positive / [true positive + false positive]). Negative predictive value refers to the proportion of people with *negative* tests who do not have the condition. It measures how well the test rules out those without the condition (i.e., true negative / [true negative + false negative]). We also assess the DPI's performance by examining its receiver operating characteristic (ROC) curve, which plots the true positive ratio (sensitivity) against the false positive ratio (1-specificity). The area under the ROC curve (AUROC) represents the instrument's ability to discriminate between disease and nondisease, with a "perfect" test generating an area of 1.0 and a worthless test giving an area of 0.5.

However, recognizing the possible limitations of

the CAGE as a valid criterion (Bisson et al., 1999), we also assess the DPI *relative* to the CAGE using standard measures of criterion validity (i.e., sensitivity, specificity, and AUROC scores) along the two sets of criteria recommended by Bisson et al. (1999, p. 716), namely problematic patterns of alcohol consumption (e.g., heavy drinking and periodic heavy consumption) and alcohol-related problems (e.g., health, work, and family problems). To determine whether the relative performance of these two instruments varies by age cohort, we conduct this assessment for two cohorts of retirement-eligible, blue-collar workers, namely for those under age 60 and those 60 and older.

Finally, drawing from the sensitivity and specificity data for the DPI, we seek to identify the cutoff score providing the DPI with the maximum potential to discriminate between those older workers with and without a drinking problem.

Method

Subjects

Subjects were identified through the membership files of nine national and local unions representing workers employed in three blue-collar sectors in the United States: transportation, manufacturing, and construction. Workers in these sectors have been identified as being at high risk of becoming heavy drinkers and developing alcohol problems (Parker & Harford, 1992). The survey data were collected from a sample of workers soon to be eligible for retirement in each of the nine unions at three times (T1 approximately 6 months prior to retirement eligibility, and T2 and T3 6 and 18 months subsequent to retirement eligibility, respectively). Retirement-eligible workers were defined as those individuals who met their union's criteria for full retirement benefits. Because these criteria vary from union to union, some retirement-eligible members in the sample are relatively young. The mean age of workers at T1 was 57 with a range of 43 to 70 years of age.

Each union gave us the names, phone numbers, and retirement eligibility dates of all of its members eligible for retirement between May 2001 and February 2002. T1 data were collected on the basis of computer-assisted telephone interviewing beginning in November 2000 with all participants interviewed approximately 6 months ($\pm$ 2 weeks) prior to their retirement eligibility date. Although we utilized the entire list of names given to us by the local unions, in the case of the national unions, we drew a random sample of names provided. Overall, the total number of respondents in T1 was 1279 (out of a target sample of 2812; overall response rate of 46%). The number of respondents from each employment sector are as follows: 933 respondents were members of three unions in transportation (including railroad workers, flight attendants, and urban transport workers); 178 respondents were members of two unions in manufacturing (including unskilled assembly-line operators, semiskilled machine operators, and skilled-trades workers); and 168 respondents were members of four unions in construction (including electricians, steamfitters, and painters).

Of the 1279 T1 respondents, 1122 participated in the T2 survey (dropout rate of 12%) and 1055 participated in the T3 survey (dropout rate of 6% from T2). Those dropping out of the sample between times T1 and T3 were not significantly different on the following characteristics from those remaining in the sample: education, race, gender, mean number of drinking problems reported at T1, mean quantity of alcohol consumed per drinking occasion, mean frequency of drinking in the past month. Moreover, there was no substantial change in the composition of the sample from T1 to T3. At T1, 882 (69%) were males and 397 (31%) were females, 76% of respondents were married, and 20% reported a DPI score of one or greater. Of the respondents, 79.5% were Caucasians, 11% were African Americans, and 3% were Hispanic. At T3, 715 respondents (68%) were males and 340 were females (32%), 77% of them were married, and 23% reported a DPI score of one or greater. Eighty-two percent of the respondents were Caucasians, 9% were African Americans and 3% were Hispanic.

Notably, despite their eligibility for retirement, rather than dropping out of the labor force, 62% of the T3 respondents remained in the active workforce, most in their T1 jobs. Indeed, as suggested by Kim and Feldman (2000), most of those 410 respondents reporting to be "fully retired" (i.e., not working even on a part-time basis) at T3 indicated that they were still seeking alternative part- or full-time "bridge employment." Most indicated that they were seeking bridge employment for one or more of the following reasons: (a) they wished to maximize their postretirement income, (b) they had taken early retirement, and (c) their spouse was still working. Only 10% of those fully retired at T3 indicated that a return to work was unlikely, primarily because of poor health or family obligations.

Measures

In all three time periods, respondents were asked questions regarding the frequency of alcohol consumed in the last month. Drawing from the measures used by Martin et al. (1992) in their research on employee drinking behavior, we asked respondents (a) on how many days in the previous month they consumed an alcoholic beverage such as beer, wine, or liquor and (b) on those occasions when they did drink alcoholic beverages, about how many drinks were consumed each time (noting that one drink equals 12 oz. of beer, 4 oz. of wine, and 1 oz. of liquor). Based on Dawson (2003), we used the product of responses to these two items to calculate the *average number of drinks consumed on a monthly basis*. However, because frequency-based measures of consumption are likely to miss unusually high drinking days, which are likely to pose a significant risk for older individuals (Dawson, 2003), we also evaluated *periodic heavy drinking* by asking respondents on about how many days in the previous month they drank six or more drinks in a given day. Individuals who reported consuming six or more drinks on at least a single day in the past month were coded as periodic heavy drinkers. We used the WHO AUDIT (Babor et al., 2001) six-drink cutoff for our definition of periodic heavy drinking because a five-drink cutoff has been noted by some to be so low as to include close to

half of all drinkers in certain high-risk groups (Wechsler & Austin, 1998). Dawson (2003, p. 20) notes that this consumption-based approach (i.e., frequency with average serving amounts specified, together with frequency of high-risk drinking) offers a high degree of parsimony and adequacy in assessing alcohol intake.

In addition to these consumption measures, we also asked respondents to report on possible consequences of problem drinking. Specifically, we asked about three health consequences. All respondents were asked to indicate whether they had been diagnosed with any type of *liver condition such as cirrhosis* and whether they had been *hospitalized in the past year*. Given that only 17 respondents were positive for a liver condition, this health consequence was dropped from the analysis. We also assessed the number of emerging health problems by comparing T3 to T1 responses to questions regarding 11 common health problems of older individuals (e.g., heart condition, diabetes). These comparative responses were used as a basis for determining whether the individual *experienced two or more health problems* for the first time over the previous two years. We also asked about family and work consequences. We used the life events scale (LES) (La Greca, Akers, & Dwyer, 1988) to determine whether respondents experienced either a stressful family problem or a problem at work over the past year. Using a six-point response format ranging from 0 (*not troubled at all*) to 5 (*most troubled I've ever been in my life*), the LES asks respondents to indicate the degree to which their life has been troubled by any one of 14 different life events over the past 12 months. Individuals responding with a score of 4 or more on any one of three family problems (i.e., worsening relationship with spouse, divorce/marital separation, worsening relationship with child) were deemed to be positive for a *family problem*. Individuals responding with a score of 4 or more on the LES item "troubles with boss or coworkers" were deemed to be positive for a *problem at work*.

In all three time periods we assessed *problem drinking* behaviors on the basis of the DPI(Finney et al., 1991). (See Appendix 1 for the list of DPI items.) Respondents were read the following statement: People often report that a variety of things happen as a result of their drinking. Then, they were asked to respond to the 17 items and how often each happened to them in the past 12 months, with response categories ranging from *never* (1) to *often* (5). Sample items are as follows: (1) In the past 12 months, how often have you had a family member worry or complain about your drinking? (2) In the past 12 months, how often have you felt you were spending too much money on drinking? (3) In the past 12 months, how often have you lost friends because of your drinking? For the actual DPI scale, which is calculated by summing all the responses to all the 17 items, to have a minimum response of zero, indicating *no* drinking problems, we transformed all the items from a 1 to 5 scale to a 0 to 4 scale. Additionally, as recommended by Finney et al. (1991, p. 398) to decrease the influence of two "liberal" items (i.e., "got a buzz/high" and "drink to forget worries"), responses were dichotomized (responses 0 to 2 were recoded as 0, and responses 3 and 4 were recoded as 1 prior to their inclusion in the summed, scale score. Indeed, consistent with Finney et al.'s findings, in the current study 30% of the participants responded that they had "gotten a buzz" "occasionally" or more often in the past year, whereas 6% said that they engage in drinking to forget their worries at least occasionally. On all other items, the response "occasionally" or more often was less than 1%, providing substantial justification for the dichotomization of these items and the exclusion of "occasionally" as criterion for a positive score (i.e., score of 1) for these items. As such, the potential scoring range for the DPI was 0 to 62 (top score of 4 for 15 items plus top score of 2 for 2 items), with Finney et al. (1991, p. 399) reporting a mean DPI score of 2.92. With this in mind, in the results section we report sensitivity, specificity, and positive predictive values for DPI scores of 1 or more, 2 or more, 3 or more, and 4 or more (1+, 2+, 3 +, and 4+ respectively).

According to Finney et al. (1991), the DPI has excellent reliability and validity. They report an internal consistency reliability estimate of .94, a cross-temporal correlation over a 1-year period of .66, and cross-sectional correlations with alcohol consumption at two points a year apart of .37 and .42. In the current study, Cronbach's alpha for the DPI was 0.9 at T1, 0.85 at T2, and 0.82 at T3. Among the drinking problems identified most frequently were (a) "became intoxicated/drunk" (19% of drinkers at T1 and 23% at T3 reported this as a problem), (b) "family member worry/complain" (9% of drinkers at both T1 and T3 reported this as a problem), and (c) "spending too much money" (9% of drinkers at both T1 and T3 reported this as at least an occasional problem). The cross-temporal correlations of the DPI were 0.54 between T1 and T2, 0.59 between T1 and T3, and 0.59 between T2 and T3.

Finally, at T3, respondents were asked to complete the four-item CAGE instrument (Ewing, 1984). As noted previously, the CAGE has been widely used by researchers and practitioners and validated in numerous clinical studies (Buchsbaum et al., 1992; Bisson et al., 1999). The CAGE poses four questions about the problematic consequences of respondents' drinking behavior in the past 12 months. (See Appendix 1 for the CAGE items.)

Analysis

Because CAGE data were collected at T3 only (unavailable for T1 and T2), all analyses were conducted on the basis of the T3 data ($n = 1055$). The relative performance of a given screening instrument was assessed on the basis of that instrument's sensitivity, specificity, positive predictive value, negative predictive value, and ROC. In general, researchers and practitioners tend to emphasize the maximization of sensitivity over specificity (Connors & Volk, 2003). As noted by Connors and Volk (p. 24), "this logic is apparent when the purpose of screening is considered. Screening is done on unselected groups (e.g., asymptomatic primary care patients) for the purpose of identifying cases where there is a heightened suspicion of a disorder."

We used a number of criteria to assess the sensitivity and specificity of the DPI. First, consistent with current practice for the CAGE, we used a positive score of two or more on the CAGE as a criterion for assessing the DPI's ability to identify alcohol problems (Ewing, 1984; Martin & Roman, 1996). Second, we used four alcohol consumption criteria drawn from Bisson et al. (1999) to assess DPI's ability to identify heavy drinking: (1) an average of 60 or more drinks per month, (2) an average of 88 or more drinks per month, (3) one or more days of heavy drinking in the past month, and (4) two or more days of heavy drinking in the past

month. As noted by Bisson et al. (1999, p. 717), these criteria are based on the current guidelines regarding hazardous consumption levels proposed by a number of national health councils (Royal College of Physicians & Surgeons of Canada, 1991). Furthermore, the use of multiple heavy drinking criteria provides the ability to assess the extent to which instrument validity may be affected by the stringency of the criteria.

Based on the sensitivity and specificity calculations, we then estimated a ROC curve. Because the performance of the DPI may vary depending upon respondent age, we estimated the sensitivity, specificity, and the area under the ROC (AUROC) for two subsets of the sample, drinkers who were older than 60 years of age at T3 ($n = 654$) and those who were younger than 60 at T3 ($n = 623$). We used two separate approaches to estimate the statistical significance of the differences in the AUROCs, depending on whether the samples compared were dependent or independent. For two dependent samples (such as comparing the AUROC of DPI vs. CAGE for the older than 60 sample), the statistical differences between the areas under two ROC curves was compared using nonparametric analysis of correlated ROC curves (DeLong, DeLong, & Clarke-Pearson, 1988). For two independent samples (such as comparing the AUROC of the DPI for the people older than 60 vs. those younger than 60), the statistical differences between the areas under two ROC curves was compared using a Chi Square test with one degree of freedom.

Finally, to determine an appropriate cutoff for the DPI, we examined the prevalence of health, family, and work problems at different cutting points for respondents younger than 60 and those 60 and older.

Results

Table 1 presents the means, standard deviations, and tetrachoric correlations (necessary given the dichotomous nature of the variables included in this table) for all the variables included in our analyses. The tetrachoric correlation between two dichotomous variables is calculated on the basis of a 2×2 matrix. Consequently, no correlation is presented in the case of those variable dyads in which one or more of the cells in the matrix had zero observations. Although most of the bivariate relationships presented in Table 1 are consistent with previous research, several are not. For example, for those older than 60, a significant negative correlation was found between the health problems variable and both DPI of 3+ ($r = -0.22$) and 4+ ($r = -0.32$). Similarly, for this same group, a negative correlation was found between the health problems variable and consumption of 60+ drinks monthly ($r = -0.32$). Such results are not so unreasonable for this age group when one considers the fact that older individuals with health problems are usually urged to cut down on their alcohol consumption, particularly if their condition requires that they take medication on a daily basis.

Table 2 presents the demographic, screening, and outcome characteristics of the DPI respondents at T3. Relatively few respondents (i.e., less than 6%) had a DPI score of 4 or more (i.e., 4+). Nearly one quarter of the full sample had a DPI score of 1 or more (i.e., 1+), although the prevalence of drinking problems was greater among respondents under age 60 than among respondents ages 60+ (27.2 vs. 19%). Similarly, twice the proportion of respondents under age 60 met the CAGE criteria for problem drinking (two or more problems) than did respondents ages 60+ (7.9 vs. 3.9%, respectively). Interestingly, these dif-

Table 1
Means, Standard Deviations, and Tetrachoric Correlations

	M	*SD*	1	2	3	4	5
DPI 1+	0.23	0.42	1	—	—	—	0.52***
DPI 2+	0.13	0.34	—	1	—	—	0.68***
DPI 3+	0.08	0.28	—	—	1	—	0.69***
DPI 4+	0.06	0.23	—	—	—	1	0.65***
CAGE 2+	0.06	0.24	0.69***	0.73***	0.74***	0.77***	1
60+ drinks monthly	0.10	0.30	0.62***	0.71***	0.71***	0.76***	0.52***
88+ drinks monthly	0.04	0.19	0.59***	0.68***	0.72***	0.77***	0.43***
1+ episode per month	0.18	0.38	0.77***	0.76***	0.77***	0.78***	0.54***
2+ episode per month	0.13	0.33	0.79***	0.83***	0.85***	0.85***	0.56***
Family problems	0.10	0.30	0.10	0.06	0.09	0.13	0.21
Problems at work	0.12	0.33	0.07	010	0.17	0.23	0.36
2+ health problems	0.28	0.45	−0.07	−0.12	−0.22*	−0.32**	−0.15
Hospitalization	0.11	0.31	0.02	0.03	0.11	0.05	0.27
1+ problems	0.60	0.49	0.13	0.10	0.02	0.13	0.14

Note. Means and standard deviations are for the full sample. Bold and below diagonal are for older than 60.
* $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$.

6	7	8	9	10	11	12	13	14
0.60***	0.62***	0.51***	0.59***	−0.04	0.05	−0.04	0.14	0.02
0.62***	0.63***	0.59***	0.55***	0.03	0.09	−0.07	0.07	−0.01
0.65***	0.64***	0.71***	0.65***	0.12	0.13	−0.00	−0.03	0.09
0.66***	0.73***	0.67***	0.65***	0.20	0.17	−0.08	0.08	0.04
0.42***	0.44***	0.56***	0.49***	0.30	0.06	0.11	0.19	0.30***
1	—	0.68***	0.75***	−0.37***	−0.01	−0.15	0.24*	−0.15
—	1	0.80***	0.86***	−0.13	0.06	−0.09	0.22	0.03
0.59***	0.60***	1	—	−0.03	0.225	−0.05	0.10	0.08
0.65***	0.66***	—	1	−0.01	0.23	−0.09	0.15	0.07
0.09	0.25	0.04	0.02	1	0.03	0.19	0.13	—
−0.34	—	0.40**	0.25	0.44***	1	0.27***	0.38***	—
−0.30**	−0.21	−0.10	−0.19	−0.02	0.129	1	0.40***	—
0.00	−0.05	−0.01	0.08	−0.03	0.41**	0.36***	1	—
−0.10	0.14	0.17	0.03	—	—	—	—	1

ferences were not reflected in respondents' actual consumption of alcohol. Similar proportions of under 60 and 60+ respondents consumed an average of 60 or more and 88 or more drinks per month, indicating heavy drinking among both groups. Indeed, periodic heavy drinking at the rate of once a month was significantly more prevalent among 60+ respondents than respondents under age 60 (22 vs. 15%, $\chi^2 = 5.01$, $p < .05$). Although family problems were more prevalent among respondents under the age of 60, health problems were more prevalent among older respondents.

Table 3 shows the DPI's performance when using a two-point cutoff on the CAGE as the performance criterion. As can be seen in Table 3, for the whole sample, DPI offers reasonable levels of sensitivity and specificity regardless of whether its cutoff is set at one or two positive responses. As would be expected given Finney et al.'s (1991) argument for the need to create the DPI in the first place, specificity

Table 2
Percent of Respondents Responding in the Affirmative

	Full sample (*n*)	Full sample (%)	Under 60 (*n*)	Under 60 (%)	60+ (*n*)	60+ (%)	χ^2
DPI of 1+	1050	22.88	514	27.24	535	18.69	10.85**
DPI of 2+	1050	13.44	514	15.95	535	11.03	5.47*
DPI of 3+	1050	8.48	514	9.34	535	7.66	0.95
DPI of 4+	1050	5.81	514	6.61	535	5.05	1.18
CAGE of 2+	687	6.11	378	7.94	309	3.88	4.87*
60+ drinks monthly	1052	10.28	514	10.12	537	10.43	0.03
88+ drinks monthly	1052	3.81	514	4.09	537	3.54	0.21
1+ episodes heavy/month	687	18.05	378	15.08	309	21.68	5.01*
2+ episodes heavy/month	687	12.81	378	11.11	309	14.89	2.17
Family problems	837	10.17	393	12.21	443	8.35	3.4
Problems at work[a]	664	12.07	413	10.90	250	14.00	1.41
2 or more health problems	1055	0.08	516	0.07	538	0.08	3.66
Hospitalization	1055	10.91	516	9.30	538	12.45	2.69
No problems	730	0.54	390	0.59	339	0.46	24.68***
1+ problem	730	0.46	390	0.41	339	0.54	24.68***

[a] In the case of 382 of the 391 cases for which there are missing data on this variable, the respondent was fully retired.
* $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$.

Table 3
Performance of DPI with the CAGE serving as criterion

	DPI												
	The sample as a whole				Under 60				60+				χ^2 (p value) Difference in AUROC
Criteria (Cage 2+)	1+	2+	3+	4+	1+	2+	3+	4+	1+	2+	3+	4+	
Sensitivity	0.81	0.76	0.62	0.5	0.76	0.73	0.57	0.43	0.92	0.83	0.75	0.66	
Specificity	0.68	0.83	0.9	0.94	0.66	0.83	0.91	0.94	0.7	0.83	0.89	0.94	
Difference in sensitivity given a common specificity across age groups	—	—	—	—	—	—	—	—	—	χ^2 = 5.5*	—	χ^2 = 1.18	
Positive Predictive Value	0.14	0.23	0.29	0.26	0.16	0.27	0.35	0.38	0.11	0.17	0.22	0.3	
Negative Predictive Value	0.98	0.98	0.97	0.97	0.97	0.97	0.96	0.95	0.99	0.99	0.99	0.98	
AUROC (SE)		0.824				0.8 (0.049)				0.89 (0.054)			1.53 (0.21)

* $p < 0.05$.

appears to remain largely invariant across age groups. However, the DPI's sensitivity does appear to vary by age, with sensitivities substantially higher across cutoff levels for respondents ages 60+ than for respondents younger than 60. Indeed, comparing the sensitivity of the DPI for those 60+ and those younger than 60 (possible at those cutoff levels with identical levels of specificity), we find that the sensitivity of the DPI with a cutoff of 2+ is significantly greater for those ages 60+ (0.83) than for those younger than 60 (0.73) ($\chi^2 = 5.5$, $p < .05$). In contrast, positive predictive values tend to be greater for respondents younger than 60 than for those 60 and older. Moreover, although the AUROC curve is higher for 60+ than for respondents younger than 60 (AUROC = .89 vs. .80), this difference is not statistically significant.

However, a comparison of AUROCs (see Table 4) for both the DPI and CAGE across a variety of other drinking-related criteria suggests that the performance of the DPI is generally superior to that of the CAGE. For both younger than 60 and 60+ respondents, the DPI generated significantly higher validities (as indicated by the AUROC) than the CAGE along three out of the four criteria. More specifically, the results in Table 4 indicate that for *both* younger than 60 and 60+ respondents, the DPI performs significantly better than the CAGE with regard to two out of the four criteria (i.e., 1+ and 2+ heavy drinking episodes per month; in both cases, $p < .05$). In addition, the DPI outperformed the CAGE in identifying those consuming 60 or more drinks per month for workers ages 60+ and 88 or more drinks per month for those younger than 60 (in both cases, $p < .001$). It is also interesting that only with respect to one criterion were the AUROCs significantly different for those younger than 60 and respondents age 60+. The DPI performed significantly better in identifying more frequent (2+ episodes) periodic heavy drinking among the 60+ respondents vs. the younger than 60 respondents (AUROC = .88 vs. .78, $\chi^2 = 3.81$, $p = .05$).

Tables 4 and 5 indicate that for both younger than 60 and 60+ respondents, the DPI was generally more sensitive than the CAGE across these consumption criteria, again suggesting that the DPI was better at detecting drinking problems when respondents reported having them. As shown in Table 4, for both younger than 60 and 60+ respondents, the sensitivity of the DPI with a cutoff of 1 or 2 remained above .50 for all criteria. In contrast, a similar cutoff for the CAGE generated sensitivities as low as .12 and .15. Even with a lower than usual cutoff point (i.e., ≥ 1),

Table 4
Performance of DPI Relative to the CAGE Along Multiple Criteria

	DPI								CAGE						χ^2 (p value)			
	Under 60				60+				Under 60			60+						
Criteria	1+	2+	3+	4+	1+	2+	3+	4+	1+	2+	3+	1+	2+	3+	DPI 60+ vs. under 60	CAGE 60+ vs. under 60	60+ DPI vs. CAGE	Under 60: DPI vs. CAGE
60+ drinks monthly																		
Sensitivity	0.69	0.54	0.4	0.33	0.59	0.5	0.39	0.32	0.54	0.21	0.08	0.375	0.125	0.05				
Specificity	0.77	0.88	0.94	0.96	0.86	0.94	0.96	0.98	0.89	0.94	0.99	0.91	0.98	1.0				
PP value	0.26	0.34	0.44	0.5	0.33	0.47	0.54	0.67	0.43	0.37	0.5	0.48	0.58	0.75				
NP value	0.96	0.94	0.93	0.93	0.95	0.94	0.93	0.92	0.92	0.88	0.87	0.87	0.83	0.83				
AUROC (SE)		0.77 (0.037)				0.75 (0.036)				0.71 (0.036)			0.64 (0.035)		0.15 (ns)	1.94 (ns)	13.02 (0.0003)	3.39 (ns)
88+ drinks monthly																		
Sensitivity	0.81	0.67	0.52	0.52	0.68	0.63	0.58	0.53	0.57	0.28	0.09	0.63	0.16	0.1				
Specificity	0.75	0.86	0.92	0.95	0.83	0.91	0.94	0.97	0.85	0.93	0.98	0.89	0.97	0.99				
PP value	0.12	0.17	0.23	0.32	0.13	0.2	0.27	0.37	0.18	0.2	0.25	0.27	0.25	0.5				
NP value	0.99	0.98	0.98	0.98	0.99	0.98	0.98	0.98	0.97	0.96	0.95	0.97	0.95	0.94				
AUROC (SE)		0.83 (0.052)				0.8 (0.062)				0.71 (0.056)			0.76 (0.057)		0.14 (ns)	0.39 (ns)	0.55 (ns)	13.11 (0.0003)
1+ episodes of heavy drinking (6+ servings)/ month																		
Sensitivity	0.75	0.56	0.47	0.35	0.79	0.58	0.45	0.33	0.49	0.26	0.09	0.46	0.12	0.03				
Specificity	0.7	0.84	0.93	0.96	0.8	0.91	0.95	0.98	0.88	0.95	0.99	0.95	0.98	0.99				
PP value	0.31	0.39	0.56	0.59	0.52	0.64	0.73	0.81	0.43	0.5	0.625	0.7	0.66	0.5				
NP value	0.94	0.92	0.91	0.89	0.93	0.89	0.86	0.84	0.91	0.87	0.86	0.86	0.8	0.79				
AUROC (SE)		0.77 (0.036)				0.83 (0.03)				0.69 (0.035)			0.70 (0.032)		1.64 (ns)	0.04 (ns)	14.27 (0.0001)	5.935 (0.015)
2+ episodes of heavy drinking (6+ servings)/ month																		
Sensitivity	0.79	0.57	0.48	0.38	0.87	0.73	0.62	0.47	0.52	0.26	0.12	0.61	0.15	0.04				
Specificity	0.68	0.83	0.92	0.95	0.77	0.9	0.95	0.98	0.87	0.94	0.99	0.94	0.98	0.99				
PP value	0.24	0.29	0.42	0.47	0.39	0.56	0.68	0.77	0.34	0.37	0.625	0.64	0.58	0.5				
NP value	0.96	0.94	0.93	0.92	0.97	0.95	0.94	0.91	0.94	0.91	0.9	0.9	0.87	0.86				
AUROC (SE)		0.78 (0.04)				0.88 (0.032)				0.70 (0.041)			0.77 (0.037)		3.81 (0.05)	1.61 (ns)	8.19 (0.004)	5.11 (0.02)

Table 5
Comparison of the Sensitivity of Cutoffs Associated With a Common Level of Specificity Across Instruments

Criteria	Age group	Specificity level	Cutoff for DPI/CAGE at that specificity	Sensitivity of DPI/CAGE at that cutoff level	McNemar's test statistic
60+ drinks monthly	Under 60	0.94	3+/2+	0.4/0.21	7.36**
	60+	0.98	4+/2+	0.32/0.125	9.78**
88+ drinks monthly	Under 60	0.85	2+/1+	0.67/0.57	4.74*
	60+	0.97	4+/2+	0.53/0.16	9.78**
1+ episodes of heavy drinking (6+ servings)/month	Under 60	0.95	4+/2+	0.35/0.26	0.42
	60+	0.98	4+/2+	0.33/0.12	9.78**
2+ episodes of heavy drinking (6+ servings)/month	Under 60	0.94	4+/2+	0.38/0.26	0.42
	Under 60	0.98	4+/2+	0.47/0.15	9.78**

* $p < 0.05$. ** $p < 0.01$.

the CAGE still missed 62% of respondents who drink 60 or more drinks per month on average, whereas the DPI (with a similar cutoff of one or more problems) missed only 41% of such individuals. As shown in Table 5, the DPI consistently demonstrated significantly higher sensitivities than the CAGE for those ages 60+.[3] For those younger than 60, DPI sensitivities were significantly higher than those of the CAGE with regard to total consumption but not with regard to episodic heavy drinking.

Interestingly, as can be seen in Table 4, no similar differences are apparent with respect to specificity. In no case did specificity fall below .68 for either younger than 60 or 60+ respondents. This was true whether a cutoff of one or two was used with the DPI and CAGE, indicating that neither instrument was worse than the other in generating false positives.

Although, as noted previously, age-related differences in the criterion validity of the DPI were apparent only with regard to one consumption-related criteria (i.e., 2+ episodes of heavy drinking), the results presented in Table 6 suggest that the criterion validity of the DPI (at least with regard to alcohol-related problem criteria) may be somewhat more age-contingent than that of the CAGE. Specifically, as shown in Table 6, relative to respondents younger than 60 and reporting to be DPI positive (i.e., DPI > 0), a higher proportion of respondents aged 60+ and reporting to be DPI positive tended to answer affirmatively to questions regarding a variety of alcohol-related problems. Among respondents younger than 60 with DPI scores of 1+ or 2+, at least 12% mentioned the existence of problems at work, 22% mentioned the existence of two or more recent health problems, and approximately half mentioned having at least one such problem (i.e., family, work, or health) in their lives. In contrast, among respondents ages 60+ with DPI scores of 1+ or 2+, at least 16% reported problems at work (this despite the fact that many of those in this category were already fully retired), 24% mentioned the existence of two or more recent health problems, and more than three quarters mentioned having at least one such problem in their lives. Although the differences in percentages between the two age groups were significant only with respect to hospitalization ($\chi^2 = 10.01$, $p < .05$), with the exception of family problems, the DPI (using either a 1+ or 2+ point cutoff) consistently performed better in capturing the extent of problems often associated with drinking for blue-collar workers age 60 and older than for those between the ages of 43 and 59. In contrast, as shown at the bottom of Table 6, the CAGE identified a similarly increasing proportion of alcohol-related problems (e.g., hospitalization, no problems, 1+ problems) at higher CAGE levels for both those younger than 60 and those 60+, with no criterion manifesting a statistically significant difference in the percentages of those identified across groups.

Discussion

These results suggest that when using problem drinking (assessed in terms of a positive CAGE

[3] Given that it is impossible to directly compare the sensitivity between two instruments, McNemar's test was conducted to examine the degree to which both instruments, given the same specificity, have similar patterns regarding true positives and false negatives.

Table 6
Percentage of Respondents Reporting Alcohol-Related Problems by DPI Cutoff and Age Group

Problem	DPI cutoff level: Under 60 1+	Under 60 2+	Under 60 3+	Under 60 4+	60+ 1+	60+ 2+	60+ 3+	60+ 4+	χ^2 difference between under 60 and 60+[1]
Family problems	11.34	13.33	17.14	21.74	10.84	10	11.11	13.04	1.55
Problems at work	11.86	13.24	15.38	17.86	16.28	18.18	22.22	27.27	4.12
2 or more health problems	23.57	21.95	25	20.59	27	23.73	17.07	11.11	7.24
Hospitalization	12.14	10.98	8.33	11.76	13	13.56	17.07	14.81	10.01*
No problems	47.57	49.18	42.42	45.45	23.73	24.24	29.17	21.43	4.25
1+ problems	52.43	50.82	57.58	54.55	76.27	75.76	70.83	78.57	4.25

Problem	CAGE cutoff level: Under 60 1+	Under 60 2+	Under 60 3+	60+ 1+	60+ 2+	60+ 3+	χ^2 difference between under 60 and 60+[1]
Family problems	17.38	26.32	0	19.51	18.18	33.33	6.00
Problems at work	12.73	10.71	0	15.79	28.57	0	0.52
2 or more health problems	23.08	30.00	25.00	25.00	16.67	0	5.16
Hospitalization	12.31	16.67	25.00	15.91	25.00	25.00	0.62
No problems	43.48	28.57	25.00	24.14	22.22	0	5.21
1+ problems	56.52	71.43	75.00	75.86	77.78	100.00	5.21

Note. χ^2 difference is calculated using percentages based on nondependent DPI and CAGE cutoff levels (i.e., DPI cutoffs being 0, 1, 2, 3, 4, and greater than 4, and CAGE cutoffs being 0, 1, 2, 3, and 4) whereas table reflects percentages assuming intercolumn dependence (i.e., 1+, 2+, 3+, etc.).
* $p < 0.05$.

score) as the criterion, the DPI offers higher sensitivities for workers ages 60+ than for workers ages 40–59 without any noticeable difference across age groups in specificity. Although it may be argued that this enhanced sensitivity is an artifact of a two-point CAGE cutoff, our findings remained essentially unchanged even when we applied the one-point CAGE cutoff criteria commonly applied by clinicians screening older individuals for alcohol problems.[4]

Moreover, our findings suggest that the DPI may provide enhanced accuracy relative to the more widely used CAGE screening instrument in correctly distinguishing between those with and without risky patterns of alcohol consumption. Specifically, we found that, across all four risky consumption criteria (i.e., number of drinks in past month and number of heavy drinking episodes in past month), the DPI offered significantly higher AUROC curve scores for both younger than 60 and 60+ respondents. We also found that across these same consumption criteria, sensitivity was significantly higher for the DPI than for the CAGE. Regardless of cutoff level, the CAGE failed to present simultaneously good sensitivity and specificity with respect to the four consumption criteria. In contrast, despite somewhat lower positive predictive values, the DPI was able to do so. Overall, our findings suggest that, consistent with the predictions of Finney et al. (1991), the DPI may offer considerable utility over the CAGE in discriminating between heavy drinkers and nonheavy drinkers in older populations.

This conclusion is reinforced when we compare the performance results for the DPI generated in the current study with similar performance data generated for other screening questionnaires in other studies. For example, even with a cutoff point of greater than or equal to two, the DPI sensitivity/specificity levels of .63/.91 and .73/.90 (for 88 + drinks per month and 2+ episodes of heavy drinking per month) with respect to older respondents compare favorably to the sensitivity and specificity results for similar criteria (e.g., .69/.84 and .48/.99 for the CAGE and .33/.91 for the AUDIT) reported in the largely clinical studies of individuals age 60 or older included in the review by Bradley, Bush, McDonell, Malone, and Fihn (1998). Although these comparative results

[4] Detailed results available upon request from the first author.

should be taken with considerable caution given that most studies examining the sensitivity and specificity of the CAGE ask about lifetime alcohol abuse and dependence and utilize clinical samples of individuals who are younger than 50 years old, overall, they suggest that the DPI may offer a higher degree of accuracy in discriminating between older individuals with and without drinking problems than several of the most widely used screening instruments.

In addition to comparing the ability of the DPI and CAGE to distinguish between older workers with and without drinking problems, a second objective of the current study was to identify the DPI cutoff score providing maximum potential for actual problem detection. Based on the sensitivity and specificity of the DPI in identifying various adverse conditions often associated with problematic drinking, a cutoff of either one or two, depending on the outcome of interest and the age (under 60 vs. 60+), appears to maximize discriminatory accuracy. For both younger than 60 and 60+ respondents, a higher DPI cutoff (i.e., of two) results in a substantial increase in specificity and positive predictive value with only a minor to moderate decline in sensitivity in identifying a drinking problem (assessed on the basis of the CAGE with a cutoff of two). However, for the identification of heavy drinking, although similar enhancements in specificity and positive predictive value may be generated with only minor to moderate decline in sensitivity for older respondents, for those in the younger than 60 group, such specificity and positive predictive value enhancements come only at the expense of a more severe decline in sensitivity.

Turning to the criterion validity of the DPI with respect to problematic consequences criterion (see Table 6), as noted earlier, regardless of cutoff level, the DPI was found to perform better with regard to the family problems criteria for those younger than 60 than for those 60+. In contrast, with regard to the problems at work criteria, the DPI consistently performed better for respondents age 60 and older than for those younger than 60 regardless of the cutoff level. We believe that these apparent age-related DPI sensitivities have much to do with when such problems generally occur over the course of one's life. Trice's (1965) early research suggests that alcohol tends to take its toll on the family before it does so in the workplace. Recent research by Frone (2005) also suggests that many alcohol-dependent employees successfully manage to keep alcohol-related problems from interfering with their work, with only those having the most severe problems eventually using alcohol at work or coming to work impaired. In this context, whereas the most salient consequences of problem drinking for those younger than 60 may be those relating to the family, family problems may be relatively less salient (and therefore less likely to be reported) than more recent or emergent work-related problems among those ages 60+.

With regard to an appropriate DPI cutoff, ultimately, the choice between a DPI cutoff of one or two will depend upon the purpose of the screening. If the screening is being conducted for clinical purposes (e.g., by an employee-assistance professional), it may be beneficial to select a cutoff of one to ensure that no true positives are excluded from the opportunity to seek help. However, if the purpose of the screening is more epidemiological in nature, a higher cutoff may be preferable to reduce the risk of including false positives. The inclusion of such false positives results in upwardly biased prevalence estimates which may, in turn, result in less than optimal public policy and the misallocation of valuable prevention and treatment resources.

Limitations

The possibility of response consistency bias serves as one possible limitation of the current study. In theory, when a single interviewer administers both screening and criterion instruments at the same point in time, there is the possibility that response consistency bias will upwardly inflate sensitivity estimates. However, we deem this risk to be limited in the current study because screening and criterion instruments were also administered to the same respondents by other interviewers at two earlier points in time, with a high degree of stability in each of the measurements across time. Indeed, as noted previously, the cross-temporal correlations for the DPI were all at or above 0.54, and the percent of respondents reporting a DPI score of two or more at T1 (13%) was not significantly different from the percent reporting a score of two or more at time T3 (13.4%).

A second potential limitation of the study stems from the temporary or permanent disengagement of a third of our sample from the labor force at the time screening and criterion data were collected. Although many individuals receiving retirement benefits eventually return to the labor force on either a part- or full-time basis (Doeringer, 1990), it is possible that with the inclusion of such a large proportion of retirees, our sample can be generalized not so much to older, blue-collar workers as to the population of older, blue-collar workers *and retirees*. Still, retire-

ment studies suggest that the probability of postretirement "bridge employment" is greater than 50% and is greatest among younger and healthier employees (Kim & Feldman, 2000). Given that retirees in the current sample are precisely such individuals (i.e., 78% were age 60 or younger and 90% were reported to be in generally good health), we viewed these individuals as temporarily unemployed but nevertheless remaining within the older, blue-collar labor force. However, to rule out the possibility of any such bias, we reran our analyses, including only those individuals actually employed at T3 (i.e., either never having taken retirement or having taken retirement but continuing to work on a full- or part-time basis for the same or alternative employer).[5] The results of these analyses were similar to the results presented previously. Specifically, using the CAGE as a criterion, the AUROC remained essentially the same (i.e., a statistically insignificant decline to .79), with sensitivities remaining consistently higher among 60+ than among employees younger than 60. Furthermore, for those currently employed, the DPI continued to offer greater detection potential (i.e., higher AUROC scores) than the CAGE with respect to all four consumption criteria (albeit *significantly* higher AUROCs were generated only with regard to the two episodic drinking criteria). These results suggest that the inclusion or exclusion of these retired workers has little meaningful effect on the validity or relative performance of the DPI as an alcohol-screening device with respect to older, blue-collar workers.

A third limitation has to do with the ethnic and gender composition of our sample (i.e., 82% white and 68% male). As noted earlier, researchers have found significant variation in the criterion validity of screening instruments by ethnicity and gender (Cherpitel, 1997). Although our findings suggest that the DPI may offer a greater degree of accuracy relative to the CAGE when screening older adults for alcohol problems, given the nature of our sample, we cannot be certain whether this holds true for older members of ethnic minorities or women.

To explore this issue a bit further, we ran a post hoc analysis of the DPI's sensitivity and specificity for those ages 60 and younger and those older than 60 (similar to that presented in Table 3) separately for men and women. The results of this post hoc analysis indicated that although DPI sensitivities (regardless of age) were generally lower for females (e.g., for a DPI cutoff of 2+ and a criterion of a CAGE score of 2+, the sensitivity was .83 for men and .71 for women), our findings of an overall insignificant effect of age on the validity of the DPI was found to hold true for both men and women. For both men and women, regardless of a criterion of one or two affirmative CAGE responses, there was no significant difference in the AUROCs for those ages 60 and younger and those older than 60. Although these findings suggest that our results may hold for both older men and women, this issue is deserving of a far more thorough analysis that goes beyond the realm of the current analysis.

A final limitation may stem from the use of the CAGE as our "gold standard" and of a six-point cutoff in the assessment of heavy episodic drinking. Although it is one of the most widely used screening instruments, as noted earlier, the CAGE has been found to perform differentially for men and women. Consequently, its use as a "gold standard" in the current study may have implications with regard to the ability to generalize our findings to women. Still, the results of the post hoc analysis suggest that regardless of gender, the DPI performed relatively the same as the CAGE for both age groups examined in this study. Moreover, studies comparing the CAGE with other screening instruments (e.g., MAST) have found the CAGE to offer superior performance with respect to older adults (Maisto, Connors, & Allen, 1995). In addition, our evaluation of the criterion validity of the DPI was not based strictly on its performance relative to the CAGE. Rather, we examined the sensitivity and specificity of the DPI in relation to a variety of criteria other than the CAGE.

Regarding the six-point cutoff, such a cutoff is used both in the WHO AUDIT (Babor et al., 2001) and the Enhanced CAGE (Bradley, Kivlahan, Bush, McDonell and Fihn, 2001). This relatively high threshold may have artificially inflated sensitivities for both the DPI and the traditional CAGE measure used in the current study. Nevertheless, because the same high threshold criterion was applied across both screening mechanisms, the conclusions drawn regarding each instrument's performance relative to the other should apply regardless of the cutoff. Moreover, given the fact that in several of the high-risk occupations included in the current analysis, the consumption of four or five drinks per drinking episode may be more the rule than the exception (Parker & Harford, 1992; Sonnenstuhl, 1996), based on Wechsler and Austin (1998) we believe that a higher than

[5] As unionized workers, a handful or individuals formally retired but were then hired back by the same employer as external, contract employees.

usual cutoff is not simply justified but actually advantageous.

Conclusion

Despite these limitations, we believe that these findings support the original argument made by Finney et al. (1991) justifying the development of the DPI as an alternative screening tool for use with older adults. Certainly, taken in the context of the validity and reliability data presented in their original study, the data presented suggest that the DPI may indeed offer significant advantages over the CAGE when the target population includes large numbers of older adults employed or soon to retire from high-risk occupations. Although the DPI is longer than the CAGE and somewhat more complex to administer because of its use of a Likert-type scoring mechanism, the instrument's inclusion of a broader range of adverse consequences specifically relevant to older workers appears to generate validities that may in many cases justify the additional questionnaire space, as well as the extra time and cognitive burdens placed on respondents. In this context, we encourage future researchers to consider using the DPI when conducting population studies of problem drinking among older high-risk workers.

References

Adams, W. L., Barry, K. L., & Fleming, M. F. (1996). Screening for problem drinking in older primary care patients. *Journal of the American Medical Association, 276,* 1964–1967.

Alvarez, F. J., & Del Rio, M. C. (1994). Screening for problem drinkers in a general population survey in Spain by use of the CAGE scale. *Journal of Studies on Alcohol, 55,* 471–474.

Babor, T. F., Higgins-Biddle, J. C., Saunders, J. B., & Monteiro, M. G. (2001). *AUDIT: The Alcohol Use Disorders Identification Test: Guidelines for use in primary care*. Geneva, Switzerland: Department of Mental Health and Substance Dependence, World Health Organization.

Bacharach, S. B., Bamberger, P. A., & Sonnenstuhl, W. J. (2002). Driven to drink: Managerial control, work-related risk factors, and employee problem drinking. *Academy of Management Journal, 45,* 637–658.

Bacharach, S. B., Bamberger, P. A., Sonnenstuhl, W. J., & Vashdi, D. (2004). Retirement, risky alcohol consumption and drinking problems among blue-collar workers. *Journal of Studies on Alcohol, 65,* 537–545.

Barnea, Z., & Teichman, M. (1994). Substance misuse and abuse among the elderly: Implications for social work intervention. *Journal of Gerontological Social Work, 21,* 133–148.

Bisson, J., Nadeau, L., & Demers, A. (1999). The validity of the CAGE scale to screen for heavy drinking and drinking problems in a general population survey. *Addiction, 94,* 715–722.

Blow, F. C. (1998). *Substance abuse among older adults* (DHHS Publication No. SMA 98–3179, Treatment Improvement Protocol (TIP) series, No. 26). Rockville, MD: Center for Substance Abuse Treatment.

BLS. (2005). *Labor force statistics from the current population survey.* Retrieved February 25, 2004, from http://data.bls.gov/PDQ/outside.jsp?survey=ln

Bradley, K. A., Boyd-Wickizer, J., Powell, S. H., & Burman, M. L. (1998). Alcohol screening questionnaires in women: A critical review. *Journal of the American Medical Association, 280,* 166–171.

Bradley, K. A., Bush, K. R., McDonell, M. B., Malone, T., & Fihn, S. D. (1998). Screening for problem drinking: Comparison of CAGE and AUDIT. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol Use Disorders Identification Test. *Journal of General Internal Medicine, 13,* 379–388.

Bradley, K. A., Kivlahan, D. R., Bush, K. R., McDonnell, M. B., & Fihn, S. D. (2001). Variations on the CAGE alcohol screening questionnaire: Strengths and limitations in VA general medical patients. *Alcohol Clin Exper Res, 25(10),* 1472–1478.

Buchsbaum, D. G., Buchanan, R. G., Welsh, J., Centor, R. M., & Schnoll, S. H. (1992). Screening for drinking disorders in the elderly using the CAGE questionnaire. *Journal of the American Geriatrics Society, 40,* 662–665.

Bush, B., Shaw, S., Cleary, P., Delbanco, T. L., & Aronson, M. D. (1987). Screening for alcohol abuse using the CAGE questionnaire. *American Journal of Medicine, 82,* 231–235.

Cherpitel, C. J. (1997). Brief screening instruments for alcoholism. *Alcohol Health and Research World, 21,* 348–351.

Clements, R. (1998). A critical evaluation of several alcohol screening instruments using the CIDI-SAM as a criterion measure. *Alcoholism: Clinical & Experimental Research, 22,* 985–993.

Conigrave, K. M., Hall, W. D., & Saunders, J. B. (1995). The AUDIT questionnaire: Choosing a cut-off score. Alcohol Use Disorder Identification Test. *Addiction, 90,* 1349–1356.

Connors, G. J., & Volk, R. J. (2004). *Self-report screening for alcohol problems among adults.* Retrieved November 20, 2004, from http://www.niaaa.nih.gov/publications/arh26–4/308–315.htm

Dawson, D. A. (2003). Methodological issues in measuring alcohol use. *Alcohol Research and Health World, 27,* 18–29.

DeLong, E. R., DeLong, D. M., & Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics, 44,* 837–845.

Doeringer, P. B. (1990). Economic security, labor market flexibility and bridges to retirement. In P. B. Doeringer (Ed.), *Bridges to Retirement* (pp. 3–22). Ithaca, NY: Cornell University Press.

Ewing, J. A. (1984). Detecting alcoholism. The CAGE questionnaire. *Journal of the American Medical Association, 252,* 1905–1907.

Finney, J. W., Moos, R. H., & Brennan, P. L. (1991). The Drinking Problems Index: A measure to assess alcohol-

related problems among older adults. *Journal of Substance Abuse, 3,* 395–404.

Fitzgerald, J. L., & Mulford, H. A. (1992). Elderly vs. younger problem drinker "treatment" and recovery experiences. *British Journal of Addiction, 87,* 1281–1291.

Frone, M. R. (2005). *Alcohol and drug use in the workplace: A national study.* Paper presented at the April meeting of the Society for Industrial and Organizational Psychology, Los Angeles.

Gambert, S. R., & Katsoyannis, K. K. (1995). Alcohol-related medical disorders of older heavy drinkers. In T. P. Beresford & E. Gomberg (Eds.), *Alcohol and aging.* New York: Oxford University Press.

Gfroerer, J. C., Penne, M., Pemberton, M., & Folsom, R. (2003). Substance abuse treatment need among older adults in 2020: The impact of the aging baby-boom cohort. *Drug & Alcohol Dependence, 69,* 127–135.

Gfroerer, J. C., Penne, M. A., Pemberton, M. R., & Folsom, R. E., Jr. (2002). The aging baby boom cohort and future prevalence of substance abuse. In S. P. Korper & L. C. Council (Eds.), *Substance use by older adults: Estimates of future impact on the treatment system* (DHHS Publication No. SMA 03–3763, Analytic Series A-21). Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies.

Heck, E. J., & Williams, M. D. (1995). Using the CAGE to screen for drinking-related problems in college students. *Journal of Studies on Alcohol, 56,* 282–286.

Hwalek, M., & Richter, F. (1983). *Measuring alcoholism among the elderly.* Paper presented at the November scientific meeting of the American Gerontological Society of America, San Francisco.

Kim, S., & Feldman, D. C. (2000). Working in retirement: The antecedents of bridge employment and its consequences for quality of life in retirement. *Academy of Management Journal, 43,* 1195.

La Greca, A. J., Akers, R. L., & Dwyer, J. W. (1988). Life events and alcohol behavior among older adults. *Gerontologist, 28,* 552–558.

Liskow, B., Campbell, J., Nickel, E. J., & Powell, B. J. (1995). Validity of the CAGE questionnaire in screening for alcohol dependence in a walk-in (triage) clinic. *Journal of Studies on Alcohol, 56,* 277–281.

Maisto, S. A., Connors, G. J., & Allen, J. P. (1995). Contrasting self-report screens for alcohol problems: A review. *Alcoholism: Clinical & Experimental Research, 19,* 1510–1516.

Martin, J. C., Blum, T. C., & Roman, P. M. (1992). Drinking to cope and self-medication: Characteristics of jobs in relation to workers' drinking behavior. *Journal of Organizational Behavior, 13,* 55.

Martin, J. C., & Roman, P. M. (1996). Job satisfaction, job reward characteristics, and employees' problem drinking behaviors. *Work and Occupations, 23,* 4.

Mayfield, D., McLeod, G., & Hall, P. (1974). The CAGE questionnaire: Validation of a new alcoholism screening instrument. *American Journal of Psychiatry, 131,* 1121–1123.

Metz, C. E. (1978). Basic principles of ROC analysis. *Seminars in Nuclear Medicine, 8,* 283–298.

Moos, R. H., Schutte, K., Brennan, P., & Moos, B. S. (2004). Ten-year patterns of alcohol consumption and drinking problems among older women and men. *Addiction, 99,* 829–838.

Parker, D., & Harford, T. C. (1992). The epidemiology of alcohol consumption and dependence across occupations in the United States. *Alcohol Health and Research World, 16,* 97–105.

Peterson, J. S., & Zwerling, C. (1998). Comparison of health outcomes among older construction and blue-collar employees in the United States. *American Journal of Industrial Medicine, 34,* 280–287.

Rice, D. (1999). Economic costs of substance abuse, 1995. *Proceedings of the Association of American Physicians, 111,* 119–125.

Rix, S. E. (2001). *Health and safety issues in an aging workforce.* Washington, DC: AARP Public Policy Institute.

Royal College of Physicians & Surgeons of Canada, Health and Public Policy Committee. (1991). *Statement on alcohol and alcohol related problems: Alcohol, a major drug of concern for Canada.* Kingston, ON: Royal College of Physicans and Sureons of Canada.

Schonfeld, L., & Dupree, L. W. (1995). Treatment approaches for older problem drinkers. *International Journal of the Addictions, 30,* 1819–1842.

Smart, R. G., Adlaf, E. M., & Knoke, D. (1991). Use of the CAGE scale in a population survey of drinking. *Journal of Studies on Alcohol, 52,* 593–596.

Sonnenstuhl, W. J. (1996). *Working sober: The transformation of an occupational drinking culture.* Ithaca, NY: ILR Press.

Trice, H. M. (1965). Reaction of supervisors to emotionally disturbed employees. *Journal of Occupational Medicine, 7,* 177–188.

Trice, H. M., & Roman, P. M. (1972). *Spirits and demons at work.* Ithaca, NY: ILR Press.

Wechsler, H., & Austin, S. B. (1998). Binge drinking: The five/four measure. *Journal of Studies on Alcohol, 59,* 122–124.

(*Appendix follows*)

Appendix: The Drinking Problems Index and the CAGE Questions

Items for the Drinking Problems Index

People often report that a variety of things can happen as a result of their drinking.
*In the **LAST 12 MONTHS***, how often have you:
(*Please circle one response for each statement):*

1. Had a family member worry or complain about your drinking
2. Become intoxicated or drunk after drinking
3. Felt you were spending too much money on drinking
4. Felt confused after drinking
5. Had problems between you and a member of your family because of your drinking
6. Had a friend worry or complain about your drinking
7. Skipped meals because of drinking
8. Had a craving for a drink the first thing after you woke up
9. Neglected your appearance because of drinking
10. Neglected the appearance of your living quarters because of drinking
11. Had a drink to help you forget your worries
12. Got a "buzz" or "high" after drinking
13. Lost friends because of your drinking
14. Had a fall or accident as a result of drinking
15. Felt isolated from people because of your drinking
16. Gone to anyone for help about your drinking
17. Neglected your work and/or daily tasks because of drinking

Response format: *1 = Never, 2 = Once or twice, 3 = Occasionally, 4 = Fairly often, 5 = Often*

Items for the CAGE

*In the **LAST 12 MONTHS***, have you:
(*Please circle one response for each statement*):

1. You felt that you ought to cut down on your drinking
2. People annoyed you by criticizing your drinking
3. You felt bad or guilty about your drinking
4. You had a drink first thing in the morning to steady your nerves or get rid of a hangover

Response format: 0 = *No*, 1 = *Yes*

Received March 8, 2005
Revision received August 18, 2005
Accepted August 21, 2005 ■

Correction to Adler et al. (2005)

As stated in "The Impact of Deployment Length and Experience on the Well-Being of Male and Female Soldiers," by Amy B. Adler, Ann H. Huffman, Paul D. Bliese, and Carl Andrew Castro (*Journal of Occupational Health Psychology,* 2005, Vol. 10, No. 2, pp. 121–137), all service members re-deploying from the Bosnia Area of Operations were required by Department of Defense policy (tasker P 231639Z FEB 96) to complete the psychological screening survey. The screening data were collected as part of routine clinical care under the Privacy Act Regulation, and secondary analysis of these data was conducted under a protocol approved by the Walter Reed Army Institute of Research, Human Use Review Committee. It should be noted, however, that subjects were not asked to consent to the secondary analyses of the screening data for research purposes.