Journal of Consulting and Clinical Psychology
2002, Vol. 70, No. 4, 897–907

0022-006X/02/$5.00 DOI: 10.1037//0022-006X.70.4.897

Alcohol Use Disorders and Cognitive Abilities in Young Adulthood: A Prospective Study

Phillip K. Wood, Kenneth J. Sher, and Bruce D. Bartholow
University of Missouri—Columbia

The effect of alcohol use disorder (AUD) on cognitive and neuropsychological abilities was investigated in a prospective study of 68 freshmen who met past-year criteria for AUD on 2 or more occasions during their college years and 66 matched controls. At baseline, participants were administered a total of 14 subtests from the Wechsler Adult Intelligence Scale—Revised, Wechsler Memory Scale, and Halstead–Reitan Neuropsychological Battery. At 7-year follow-up, most measures were readministered, along with the Reflective Judgment Interview, Watson–Glaser Critical Thinking Appraisal, and Plant Test. Analyses revealed few differences between AUD and control groups. However, visuospatial deficits may be present among AUD participants with poor baseline visuospatial performance. Alcohol exposure measures yielded similar patterns to those shown with AUD.

Reviews of collegiate drinking practices consistently have noted the high prevalence of persistent alcohol use and abuse on college campuses and the many negative consequences associated with it (e.g., Berkowitz & Perkins, 1986; Prendergast, 1994; Saltz & Elandt, 1986; H. Wechsler, Isaac, Grodstein, & Sellers, 1994). Nearly half of all college students engage in "binge" drinking (five or more drinks in a row for men and four or more drinks in a row for women), and about one in five students report frequent binge drinking (H. Wechsler et al., 1994). Such heavy consumption patterns are associated with poor academic performance, reflected in lower course grades, being placed on academic probation, lower rates of college degree attainment, and spending fewer hours studying (e.g., J. L. Brown, 1989; Engs, 1977; Hughes & Dodder, 1983; Jessor, Donovan, & Costa, 1991; H. Wechsler, Dowdall, Davenport, & Castillo, 1995; M. D. Wood, Sher, & McGowan, 2000; P. K. Wood, Sher, Erickson, & DeBord, 1997), as well as health problems (H. Wechsler et al., 1995). Despite the consistency of these reports, it is unclear whether academic problems experienced by alcohol-abusing college students necessarily indicate generalized cognitive impairment, impairment in abilities particularly associated with the cognitive outcomes of higher education, or merely noncompliance with academic norms or institutional requirements.

Phillip K. Wood, Kenneth J. Sher, and Bruce D. Bartholow, Department of Psychology, University of Missouri—Columbia.

Bruce D. Bartholow is now at the Department of Psychology, University of North Carolina at Chapel Hill.

Portions of this research were presented at the 1997 meeting of the Research Society on Alcoholism in San Francisco, California. This research was supported in part by Grant R01 AA07231 08S1 from the National Institute on Alcohol Abuse and Alcoholism to Kenneth J. Sher. We thank Heather Gotham, Kristina Jackson, Matthew Meier, Susan O'Neill, Sara Jo Nixon, and Daniel Vinson for helpful comments on earlier versions of this article.

Correspondence concerning this article should be addressed to Phillip K. Wood, Department of Psychology, University of Missouri—Columbia, 210 McAlester Hall, Columbia, Missouri 65211. E-mail: woodpk@missouri.edu

Some research suggests that excessive exposure to alcohol is indeed related to cognitive test performance deficits, even for such relatively high-functioning subpopulations as college undergraduates. Parsons and Nixon (1993) noted that impaired performance has been observed across varied cognitive skills involving perceptual–motor, visuospatial, problem-solving, and learning abilities. Of these deficits, those associated with visuospatial ability have been reported most consistently (Evert & Oscar-Berman, 1995). In a study of college freshmen, Sher, Martin, Wood, and Rutledge (1997) found that participants with *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed.; *DSM–III*; American Psychiatric Association, 1980) alcohol use disorder (AUD) diagnoses demonstrated lower performance on selected neuropsychological tests than freshmen without an AUD diagnosis. In addition, individuals with alcohol dependence scored significantly lower on measures of motor speed than those who met only abuse criteria.

The literatures on the cognitive effects of social drinking (Parsons & Nixon, 1998) and on cognitive deficits in clinically ascertained adolescent substance abusers (S. A. Brown, Tapert, Granholm, & Delis, 2000; Moss, Kirisci, Gordon, & Tarter, 1994; Tarter, Mezzich, Hsieh, & Parks, 1995) have yielded mixed results. On balance, however, neurocognitive deficits appear to vary as a function of alcohol involvement, with apparently wide individual differences in susceptibility to alcohol impairment (Parsons, 1998). One potentially important individual difference is chronological age. Recent findings from studies examining alcohol effects in adolescent rats (Swartzwelder, Wilson, & Tayyeb, 1995a, 1995b) and in humans ranging from adolescence (S. A. Brown et al., 2000) to young adulthood (Acheson, Stein, & Swartzwelder, 1998) suggest that the brain may be particularly susceptible to the toxic effects of alcohol during adolescence and young adulthood (see also Fackelmann, 2000). Such findings indicate not only that young drinkers are vulnerable to neuropsychological impairment but that we need to distinguish between relatively short-term and long-term effects of drinking.

In addition to basic cognitive functions assessed by clinical neuropsychologists, Pascarella and Terenzini's (1991) review of higher education research outlined four areas of cognitive or intellectual competencies most identified as higher education out-

comes in young adulthood: general verbal skills (vocabulary ability and verbal reasoning skills), Piagetian formal reasoning (formal operations and the ability to reason abstractly and deductively), critical thinking (the ability to use logic to form correct conclusions from information, to identify central issues and assumptions in an argument, and to interpret whether conclusions are warranted on the basis of data; e.g., Furedy & Furedy, 1985), and use of reason to solve ill-structured problems (referred to as "reflective judgment"; King & Kitchener, 1994; Wood, 1997). Tests of formal operations may have particular relevance for studies of alcohol-related deficits, given that they also have been found to discriminate between clinical alcoholic and nonalcoholic populations (Nixon & Parsons, 1991).

Even in the case of cognitive deficits found to be associated with problematic alcohol use among young adults, questions exist as to how these deficits arise. Evidence of cognitive impairment resulting from problematic alcohol consumption points to a general decreased ability to learn relative to peers (e.g., Parsons & Nixon, 1993, 1998), possibly as a result of the neurotoxic effects of alcohol, which may degrade intellectual functioning or interfere with new learning. In addition, and of equal importance, problematic alcohol consumption may interfere with an individual's ability to successfully assume new sociodevelopmental roles during young adulthood (e.g., Baumrind & Moselle, 1985; Newcomb & Bentler, 1986, 1988). In particular, such consumption may interfere with the assumption of the responsibilities of a college student, preventing the individual from deriving the full educational benefits of the curricular, co-curricular, and extra-curricular experiences that constitute the crucial environments for the development of higher intellectual functioning (e.g., Astin, 1993; Stage, Watson, & Terrell, 1999). One implication of developmental theory (e.g., Baumrind & Moselle, 1985) is that such alcohol-related cognitive effects should occur in those intellectual functions that are emerging at a particular time in an individual. From this perspective, we would think that those tasks that are most closely coupled with the higher education experience would show the greatest deficits.

The current study represented an attempt to address two classes of hypotheses regarding the effects of alcohol use on cognitive abilities in a college-aged sample. First, alcohol abuse during the college years may result in specific performance deficits for those abilities associated with a college education (e.g., critical thinking, reflective judgment, and formal operations), because alcohol involvement reduces optimal engagement in the tasks and activities that promote the development of higher cognitive functions. Second, AUDs during this time may result in deficits on more traditional neuropsychological measures, particularly visuospatial ability, because of the direct, neurotoxic effects of alcohol. In addition, we hypothesized that AUDs during the college years may result in smaller gains (or even deficits) in cognitive development as a function of baseline performance. Specifically, alcohol-related cognitive deficits may be manifest only for those individuals who are at lower initial levels of performance (and thus more likely to be functioning at the limits of their competencies) but not present for those with an already high degree of functioning. Such a pattern of differential effects may help to explain discrepancies between conclusions based on college-aged samples and those based on community or clinically ascertained samples.

The present study was designed to assess the relationship of AUDs during the college years with cognitive performance and to test whether any observed effects could be due to other plausible third-variable explanations. As such, this study differed in four ways from previous research: (a) Cognitive performance and intellectual performance were directly assessed rather than inferred from self-reports or educational achievement; (b) participants were classified on the basis of an extended longitudinal observation period using structured diagnostic measures; (c) baseline performance was available for several measures; and (d) the effects of recent alcohol or other drug consumption were systematically investigated with a detailed, structured assessment of substance use during the previous month. In addition, we considered a number of potential confounds (e.g., baseline intellectual ability and previous educational achievement) by including a control group matched to the AUD group on a baseline composite measure of verbal ability (consisting of Wechsler Adult Intelligence Scale—Revised [WAIS–R] Vocabulary and Similarities subtest scores, ACT composite score, and high school class rank) and by assessing the influence of a number of relevant covariates such as other concurrent drug use and recent (i.e., past day or month) alcohol consumption.

Method

Participant Selection and Screening

Baseline Sample

Participants were drawn from an ongoing, prospective study of offspring of alcoholics that began in the 1987–1988 academic year, when the participants were freshmen at a large midwestern university. An initial sample ($N = 3{,}156$) was screened for the presence of family history of alcoholism with an extensive battery of interview and questionnaire measures. However, because family history status was not of central interest in the present study, we do not provide details of the screening procedure here. A detailed description of sampling, ascertainment procedures, and participant characteristics has been presented elsewhere (Sher, Walitzer, Wood, & Brent, 1991). The baseline sample retained for this study ($N = 489$) consisted of roughly equal numbers of individuals with and without a family history of paternal alcoholism and roughly equal numbers of men and women. The mean age of the sample was 18.2 years at screening, and most of the participants (94%) were White.

In addition to the baseline assessment, participants were assessed on four subsequent occasions (Years 2, 3, 4, and 7 of the ongoing study). Participants who were the focus of the current study were assessed at an additional wave of data collection, when the sample was 7 years postmatriculation (Year 8, cognitive follow-up). The Diagnostic Interview Schedule, Version III–A (DIS–III–A; Robins, Helzer, Croughan, Williams, & Spitzer, 1985), was used for assessment at baseline and Year 2, and the revised DIS–III (DIS–III–R; Robins, Helzer, Cottler, & Goldring, 1989) was used at Years 3, 4, and 7. On the basis of DIS score at each year, participants were classified as having (a) no (past-year) AUD, (b) (past-year) alcohol abuse without dependence, or (c) (past-year) alcohol dependence. As a means of maintaining consistency across all waves of data collection, *DSM–III* (American Psychiatric Association, 1980) diagnostic criteria were used throughout.

Cognitive Follow-Up Sample

For inclusion in the cognitive follow-up, participants from the baseline subsample were required to meet eligibility criteria on several variables, including age at baseline data collection (younger than 20 years), completion of all assessments during the first four waves of data collection, and complete academic and assessment data (used to create the participant matching variable). Participants were eligible for inclusion in the AUD

group if they met criteria for a *DSM–III* (past-year) AUD (i.e., alcohol abuse or dependence diagnosis) at two or more of the four initial waves of data collection. Participants were eligible for inclusion in the (non-AUD) control group if they were not diagnosed with (past-year) alcohol abuse or dependence at any time during the initial four waves of assessment. Note that although prevalence rates of AUDs are roughly similar across the *DSM–III*, the revised *DSM–III* (*DSM–III–R*), and the fourth edition of the *DSM* (*DSM–IV*; American Psychiatric Association, 1987, 1994), *DSM–III* alcohol dependence (which requires physical dependence for diagnosis) is a narrower, and therefore rarer, condition than outlined in the *DSM–III–R* and *DSM–IV*, which adopt a broader notion of dependence. To assess the effect of participant attrition on the available prospective data, we calculated general linear models comparing baseline performance on all cognitive measures among participants who did and did not participate in the prospective arms of the study. No statistically significant differences between participants and nonparticipants were found (all ps $> .10$). Although this finding is somewhat reassuring, given the fact that several measures were compared, it must be noted that the statistical power to detect systematic biases was small in that only 33 participants dropped out of the study.

Of the original 489 participants assessed throughout the first four waves of data collection, 102 AUD participants (68 men and 34 women) and 113 controls (77 men and 36 women) were deemed eligible for inclusion in the cognitive follow-up sample according to the criteria just described. Attempts were made to contact, schedule, and assess each of these individuals who were still residing within Missouri or the surrounding states.[1] In total, 153 participants completed the assessments for the cognitive follow-up (71% of those eligible). In an effort to ascertain whether those individuals who elected to participate in the cognitive follow-up differed from those who did not, general linear models analyses including sex, AUD status, and participation status were conducted. These analyses revealed no statistically significant differences in baseline cognitive performance between participants and nonparticipants with one minor exception: A statistically significant interaction between AUD status and participation was found, $F(1, 210) = 4.29$, $p = .04$. Data from several participants were excluded from all analyses for the following reasons: 6 participants reported consumption of alcohol less than 24 hr before the assessment or were suspected of being drug or alcohol impaired during the assessment, or both; 5 participants had suspected neurological impairment, including serious head injury ($n = 1$) and neurologic disease (e.g., AIDS-related dementia, multiple sclerosis, or pituitary tumor; $n = 4$), since the baseline assessment; and 8 participants could not be included because no eligible match could be found for them.[2]

The final sample on which all analyses were based consisted of 68 AUD participants and 66 controls ($N = 134$). Although differences in degree attainment have been found in the overall sample from which this study's participants were drawn (M. D. Wood et al., 2000), no differences were found in this study between AUD participants and controls in terms of level of self-reported academic degree attainment, $\chi^2(2, N = 134) = 3.38$, $p = .18$ (see Table 1). The probable reason for this negative finding is that participants were matched on baseline general academic ability and performance. Forty-four (64.7%) of the AUD participants had a positive family history of alcoholism, as compared with 28 (42.4%) of the non-AUD participants.

Table 1
Frequency of AUD and Control Participants by Level of Educational Attainment

Group	Level of education		
	Less than bachelor's	Bachelor's	Graduate work or degree
AUD			
n	24	29	13
%	36.36	43.94	19.70
Control			
n	26	36	6
%	38.24	52.94	8.82

Note. $N = 134$. AUD = alcohol use disorder.

Participant Matching

As a means of ensuring some degree of comparability between the two groups in terms of baseline general academic ability, AUD participants were matched to same-sex controls through a composite variable of high school class rank, ACT English and Math subtests, and baseline scores on the WAIS–R Vocabulary and Similarities subtests. Percentile ranks were computed on the basis of each participant's relative standing in the full baseline sample on this composite measure (to within 5 percentile points).

Measures

Measures used in the present study generated three types of data. First, measures administered at both baseline and the cognitive follow-up permitted analysis of mean change across time (e.g., WAIS–R Vocabulary). Second, in some of the measures given at baseline and follow-up, the scoring changed across administrations (e.g., new forms of the Visual Reproduction and Logical Memory subtests from the Wechsler Memory Scale—Revised [WMS–R]). Finally, measures administered only at follow-up allowed inclusion of presumed cognitive outcomes of higher education (e.g., the Watson–Glaser Critical Thinking Appraisal [WGCTA], the Reflective Judgment Interview [RJI], and the Plant Test) and a traditional neurocognitive test not administered at baseline, the Visual Memory Span subtest of the WMS–R.

Traditional Tests of Neurocognitive Performance

Measures of neurocognitive performance consisted of scales that previous research has indicated may be sensitive in regard to detecting impairments associated with alcohol abuse (e.g., Nixon, 1995). Fourteen tests or subtests, drawn from three larger neurocognitive instruments, were administered at baseline: the Vocabulary, Similarities, Digit Span, Block Design, and Digit Symbol subtests of the WAIS–R (D. Wechsler, 1981); the Trail Making Test (Parts A and B) and the booklet version of McCampbell and DeFilippis's (1979) Category Test, taken from the Halstead–Reitan Neuropsychological Battery (HRNB; Reitan, 1969); and the Personal and Current Information, Mental Control, Orientation, Visual Reproduction (immediate and delayed), Paired Associate Learning, and Logical Memory

[1] However, several individuals did not participate, for the following reasons: 12 individuals could not be contacted after repeated attempts had been made; 29 individuals were contacted but were unable to schedule a time for the required face-to-face assessment; 5 individuals were scheduled but did not keep their appointments; 8 individuals in the control group were excluded from eligibility because their matching AUD-diagnosed participant could not be contacted; 4 individuals refused to participate in the cognitive follow-up; and 4 individuals withdrew from the ongoing study subsequent to the Year 4 data collection.

[2] To allow a sufficient number of matched pairs, two ALC–non-ALC pairs differed by as many as 10 percentile points on the composite measure. To keep as many eligible AUD participants in the sample as possible, we included data from 2 AUD participants for whom no match could be found in the analyses, resulting in an imbalance in the numbers of participants in the two groups. Seven participants misunderstood the directions for the WMS Paired Associates subtest, resulting in a sample size of 127 for this subtest.

(immediate and delayed) subtests of the WMS (D. Wechsler, 1944; see Russell, 1975).

Many of the baseline measures were readministered at follow-up; however, the Logical Memory, Paired Associate Learning and Visual Reproduction subtests of the WMS were substituted with those subtests from the WMS–R, and the WMS–R Visual Memory Span subtest (D. Wechsler, 1981) was added. Note, however, for the Visual Reproduction and Logical Memory subtests that only the immediate (and not delayed) assessments were administered. In addition, the WMS Personal and Current Information, Mental Control, and Orientation subtests were not readministered at follow-up.

Tests of Higher Intellectual Functioning

At follow-up, three tests thought to measure the cognitive outcomes of higher education were administered: the RJI (Kitchener & King, 1985), the WGCTA (Watson & Glaser, 1980), and the Plant Test (a measure of formal operations; Kuhn & Brannock, 1977). These tests were selected as representing types of intellectual abilities that improve as a result of attending college (see Pascarella & Terenzini, 1991, for a comprehensive review).

The RJI is a semistructured instrument designed to assess how individuals justify their answers to "ill-structured" problems about which even qualified experts could disagree. The interview consists of four dilemmas (taken from science, current events, history, and religion) presented in a random sequence and accompanied by a set of standardized probe questions.[3] Averages of the four dilemma ratings are used to form a composite RJI score. The intraclass correlation coefficient based on two raters (and assuming that these raters are randomly drawn from a larger pool) is .88. The coefficient alpha based on the dilemmas was .69.

The WGCTA is an 80-item, multiple-choice pencil-and-paper measure composed of five subsections: (a) inference, (b) recognition of assumptions, (c) deduction, (d) interpretation, and (e) evaluation of arguments (Watson & Glaser, 1980). In the current study, only composite scores were used. The coefficient alpha across the 80 items was .83.

In the Plant Test, participants are shown four plants, two appearing healthy and two appearing unhealthy. Plants used in the present study were silk and plastic; the unhealthy plants were made to look so by exposure to heat. Stipulated treatment regimens for each plant, consisting of information on amount of water, plant food, and leaf lotion, are also presented to the participant. The participant is then asked to predict the outcome of a fifth, unseen plant based on a description of the treatment it receives. Next, the participant is asked to isolate the variable in the treatment strategy responsible for the plant outcome. Finally, the participant is asked to indicate whether the effects of an experimentally defined irrelevant variable were responsible for the plant's outcome and how she or he knows the effects of that variable. Scores typically range from 0 (*concrete*; no concept of variable isolation) to 4 (*formal*; isolation of operative variable and logical exclusion of inoperative variables). In the current study, a fifth scoring level was added (similar to formal, but the participant also states that the answer is based on the information given and could change in light of receipt of additional information). In the present study, each interview was audiotaped. Initially, responses to the test were scored by two raters, but this procedure yielded unacceptably low agreement rates (weighted $\kappa = .53$). As a result, it was decided to assign scores by consensus of the first two authors. In the case of both the RJI and the Plant Test, raters were unaware of all participant characteristics.

Substance Use and Substance Use Disorders

Recent substance use. We administered the Timeline Follow-Back Interview, a retrospective diary method that provides data regarding daily drinking and other drug use over a period of up to 1 year (see B. B. Cohen & Vinson, 1995, and Sobell, Sobell, Leo, & Cancilla, 1988, for more detailed descriptions of the technique). In the present study, participants indicated their substance involvement during the past 28 days, separately for alcohol and illicit drugs. Reliability coefficients between .70 and 1.00 for various composite outcome measures have been reported for the procedure (Miller, Heather, & Hall, 1991). On the basis of these data, we chose to define two average daily consumption measures for alcohol and other substances, defined respectively as average number of drinks consumed daily over the past 28 days and number of days on which other drugs were used. Other measures were explored in additional analyses (such as maximum number of days of abstinence from alcohol and other drugs and maximum number of days of consecutive use), but these analyses revealed much the same pattern as that described for the average consumption measures and so are not reported here.

Substance use disorders. As previously noted, alcohol and drug use disorders were assessed with the DIS–III-A (Robins et al., 1985) and the DIS–III-R (Robins et al., 1989). Details concerning interviewer training and quality control have been provided by Sher et al. (1991).

Past alcohol use. At Years 1, 2, 3, 4, and 7, we assessed alcohol use and heavy alcohol use. We used past-year estimates because estimates over an extended time frame are less likely to be affected by periodic fluctuations in drinking patterns. Each item was assessed with ordinal response scales. Frequency variables were scaled to reflect drinking occasions per week, and quantity variables were scaled to reflect standard drink equivalents. Total consumption was estimated through a quantity–frequency composite. Frequency of heavy drinking was estimated from a composite variable composed of three indicators (coefficient alphas ranged from .78 to .82 over the 7 years of the study) representing mean number of heavy drinking occasions per week (based on the frequency of past-month reports of being "drunk," being "high on alcohol," or consuming five or more drinks in a single sitting).

Health-Related Variables

We assessed by interview each participant's history of neurologic diseases or other medical problems associated with neurologic complications (including loss of consciousness as a result of injury) to identify, for exclusion, participants who may have suffered cognitive impairment owing to injury or illness after the baseline assessment.

Procedure

Baseline Assessment

At baseline, participants were contacted by telephone and asked to take part in a project assessing the development of health behaviors during college. Those who consented were scheduled for a series of assessment appointments. During the initial sessions, participants were administered selected sections of the DIS–III-A by trained interviewers. An independent, trained editor later reviewed completed interview forms. Editing questions arising as a result of the initial interview were resolved by referral to the session audiotapes, which also were reviewed at random. During a later session, participants were administered the neurocognitive test battery in the following order: WAIS–R Vocabulary and Similarities; WMS Personal and Current Information, Orientation, Mental Control, and Logical Memory; WAIS–R Digit Span; WMS Visual Reproduction and Paired Associate Learning; WAIS–R Block Design and Digit Symbol; HRNB Trail Making Test (Parts A and B); WMS Logical Memory (delayed) and Visual Reproduction (Russell, 1975); and HRNB Booklet Category Test. Extensive

[3] A more detailed description of the rating procedures for the RJI can be found in King and Kitchener (1994) or P. K. Wood (1997). Briefly, each dilemma topic of the RJI is transcribed separately. Identifying information is removed from the transcript, and the order of rating of the individual transcripts is randomized. Dilemma transcript ratings that are discrepant by more than one level are resubmitted to raters, along with transcripts that originally met this criterion. If, after rerating, raters continue to assign discrepant scores to a transcript, a consensus score is assigned to the transcript.

analyses of these baseline neuropsychological data can be found in Sher et al. (1997).

Cognitive Follow-Up Assessment

Potential participants for the cognitive follow-up were contacted by telephone and asked to take part in the next phase of the research project. Those who were successfully recruited were scheduled for a one-session assessment in Columbia, St. Louis, or Kansas City, Missouri, or Chicago, depending on participants' proximity to Columbia. The order of the assessments was as follows: health interview; Timeline Follow-Back Interview; WAIS–R Vocabulary and Similarities; WMS–R Logical Memory; WAIS–R Digit Span; WMS–R Logical Memory Span, Visual Reproduction, and Paired Associate Learning; WAIS–R Block Design and Digit Symbol; HRNB Trail Making Test (Parts A and B); Plant Test; RJI; HRNB Booklet Category Test; WGCTA; Brief Symptom Inventory (Derogatis, 1993; Derogatis & Spencer, 1982); and drug, alcohol, and demographic questionnaire items. A 10-min break was inserted after administration of the Plant Test and before the RJI, and a 5-min break was inserted after the RJI. In addition, breaks were given at the request of the participant or at the discretion of the tester if any participant appeared fatigued. Participants received a payment of $75 for completion of all follow-up assessments and additional stipends for travel to the testing location. Table 2 provides a summary of the measures administered and indicates which measures involved revised scoring procedures at follow-up.

Results

Before we describe the analyses related to our primary goals, it is helpful to first characterize the general context of intellectual growth associated with this period of human development. Consistent with national norms, we saw improvements in performance across a broad range of tests, ranging from 0.41 standard deviations on WAIS–R Similarities to 0.65 standard deviations on the Booklet Category Test. Thus, to the extent that AUDs could interfere with cognitive growth beyond the level assessed at baseline performance, we would have expected to detect measurable impairment. These gains in performance were also similar to the 0.56 standard deviation gain in general verbal ability during the college years (Pascarella & Terenzini, 1991).

Although repeated measures analyses including these variables were not possible owing to changes in the scale, covariance analyses were performed to test for between-groups differences at follow-up while controlling for baseline scores on earlier versions of the tests. Although none of the participants in the control group were diagnosed with AUD at any of the four assessments, they did report alcohol consumption during these 4 years; however, levels were obviously much lower than those exhibited by the AUD participants. In addition, these participants were followed again 7 years after initial testing. Non-AUD participants averaged 3.44, 3.78, 3.27, 4.16, and 6.78 drink equivalents per week across the five measurement occasions; corresponding values for the AUD participants were 17.22, 15.57, 11.38, 11.07, and 6.97.[4]

In the case of measures on which baseline and follow-up measures were identical, we conducted paired *t* tests to examine mean changes in the control participants' performance over the follow-up interval. These tests were significant for WAIS–R Vocabulary and Similarities, Digit Span, and Block Design; WMS Paired Associates; and the HRNB Booklet Category Test, indicating significant growth in cognitive abilities in this period of young adulthood that could not easily be explained as a practice phenomenon given the 6–7-year interval between measurement occasions (all $ps < .001$ except for WAIS–R Similarities [$p = .0257$]). Matched *t*-test statistics were not significant (all $ps > .59$) for Digit Symbol or HRNB Trail Making Tests Parts A and B, indicating no significant improvement on these measures. This pattern is consonant with the view that such measures represent a combination of motor and processing speed. Such measures have been implicated as proximal mediators of age-related poor performance (e.g., Salthouse, 1992). Against this general background of growth, we examined effects of AUDs during the college years on cognitive functioning.

We had previously documented baseline differences between participants with and without AUD during their freshman year (Sher et al., 1997). Because the direction of effects in such cross-sectional findings is ambiguous, we designed the study to be sensitive to further changes in cognitive functioning associated with prospective drinking history. That is, we matched AUD participants with controls based on initial general intellectual ability and performance to assess the effect of AUDs during the ensuing 6-year follow-up period.

Given the study goals and design, two different types of hypotheses were investigated. In the case of those outcome measures for which corresponding baseline measures existed, we examined whether performance differed as a function of AUD status and sex using analyses of covariance (ANCOVAs) with baseline performance as a covariate. We also estimated the interaction of baseline performance (i.e., the covariate) with our AUD variable so as to examine whether there were differential AUD effects for those high versus low on the underlying ability measured at baseline as well as to test the homogeneity of slope assumption underlying ANCOVA (different AUD effects for those high and low in underlying ability might be detected if the size of an AUD effect were dependent on the reserve capacity of the function being tested; Satz, 1993). Baseline scores were first centered to a mean of zero to eliminate nonessential collinearity (Aiken & West, 1991). When there was no significant interaction, interaction terms were dropped to conform to a traditional ANCOVA model, and the model was reestimated. Finally, when no baseline covariates were available for a measure, general linear models involving AUD group and sex were specified. This was the case for the Visual Memory Span subtest of the WMS–R and for the tests of cognitive abilities thought to be related to college outcomes (i.e., critical thinking skills, reflective judgment, and levels of formal operations). Additional analyses, discussed subsequently, were conducted to determine whether any obtained deficits could be attributed to either long-term or short-term effects of alcohol consumption patterns.[5]

[4] Although it would initially appear that these two groups converged in their alcohol consumption after the college years (in Year 7 of the study), more detailed data regarding daily alcohol consumption suggest that the AUD group continued to consume at levels comparable to Year 4 ($M = 12.70$ drinks per week), whereas the control group increased slightly ($M = 7.23$ drinks per week).

[5] In addition, because participants were drawn from a larger study that oversampled participants with a family history of alcoholism, we also conducted analyses that included family history status as a covariate. These

Table 2
Measures and Assessment Design

Scoring procedures at baseline and follow-up		
Same	Different	Administered only at follow-up
Vocabulary (WAIS–R)	Logical Memory	Reflective Judgment Interview
Similarities (WAIS–R)	Visual Reproduction	Watson–Glaser Critical Thinking Appraisal
Digit Span (WAIS–R)	Paired Verbal Associates	Plant Test
Block Design (WAIS–R)		Timeline Follow-Back Interview
Digit Symbol (WAIS–R)		Visual Memory Span (WMS–R)
Trails Part A (HRNB)		
Trails Part B (HRNB)		
Booklet Category (HRNB)		
General Severity Index (BSI)		

Note. WAIS–R = Wechsler Adult Intelligence Scale—Revised; WMS–R = Wechsler Memory Scale—Revised; HRNB = Halstead–Reitan Neuropsychological Battery; BSI = Brief Symptom Inventory.

Traditional Neurocognitive Performance Tests

Results of analyses of variables with corresponding baselines are presented in the top section of Table 3.[6] Note that, in the case of each of these models, there was a strong effect for baseline covariate, $29.18 < F(1, 126) < 149.10$. Correlations between baseline and retesting ranged from .39 to .77. Only one test, WMS–R Visual Reproduction, demonstrated a significant main effect due to AUD status when scores were adjusted for baseline performance. Control participants performed significantly better on this measure at follow-up (adjusted $M = 38.05$, $M = 37.98$, $SD = 2.14$) than AUD participants (adjusted $M = 36.95$, $M = 37.14$, $SD = 2.67$). A significant interaction between the covariate (baseline score) and AUD group was also found. This interaction indicated that the difference between the AUD and control groups was present only for individuals scoring below a given value at baseline. Using the Johnson–Neyman procedure (Johnson & Neyman, 1936), we estimated that the performance of the AUD participants was significantly poorer than that of the control participants only among those who scored less than 12, roughly the 43rd percentile of the control group used in this study (Johnson–Neyman lower bound critical score: 12.41) at baseline.[7]

In addition, a significant AUD × Sex interaction was found for WMS–R Logical Memory subtest score, $F(1, 126) = 5.06$, $p = .026$. AUD women ($M = 29.74$, $SD = 6.14$) scored lower than non-AUD women ($M = 31.96$, $SD = 5.81$), AUD men ($M = 32.41$, $SD = 5.66$), or non-AUD men ($M = 32.41$, $SD = 5.66$).[8]

Tests Associated With College Outcomes

Next, we examined three constructs—critical thinking, reflective judgment, and formal operations—that have been particularly associated with the cognitive outcomes of college. No statistically significant main effects or interactions involving AUD were found.[9]

Taken together, the analyses presented so far yielded no statistically significant effects for collegiate AUD, with the exception of WMS–R Visual Reproduction. In evaluating such null findings, however, it is important to consider whether the observed differences between the AUD and non-AUD groups could be thought clinically important even though they were not statistically significant. One way to evaluate this is to take into account the average performance of the AUD group on the follow-up measures expressed as a standardized value based on control group performance. Taken across all follow-up measures except Visual Reproduction (for which, as mentioned, a significant main effect was

analyses yielded no statistically significant main effects involving family history, and patterns of statistical significance for AUD were identical to those presented here. We elected to present analyses that did not include family history as a covariate to demonstrate that the failure to find effects for AUD does not appear to be due to the collinearity of AUD with family history.

[6] Note that four measures revealed significant sex differences (WAIS–R Digit Span, WAIS–R Block Design, WMS–R Visual Memory Span, and WGCTA). In all cases, performance of men was slightly better than that of women. Given that these effects did not involve major research questions, we do not discuss them further here.

[7] Recall that participants for the present study were drawn from a population that was oversampled for family history of alcoholism, and note that family history of alcoholism has been associated with deficits in visuospatial information processing in both adults (Schandler, Cohen, & Antick, 1992; Schandler, Cohen, McArthur, Antick, & Brannock, 1991) and elementary school children (Schandler, Brannock, Cohen, Antick, & Caine, 1988). For this reason, follow-up analyses including family history status as well as an interaction of family history status with baseline Visual Reproduction subtest score were conducted. In these analyses, both the main effect and the Baseline Score × AUD Diagnosis interaction remained significant. To investigate whether this observed main effect was an artifact of both family history and general psychological distress, we considered family history of alcoholism as well as baseline and retest scores on the Brief Symptom Inventory General Severity Index as possible third-variable explanations for this effect. The statistically significant effect for AUD diagnosis and the AUD Diagnosis × Baseline Score interaction remained when these variables were added to the model simultaneously, individually, and in any pairwise combination.

[8] Alternatively, if adolescent alcohol involvement is particularly pernicious for women, one would expect that the matched design of the study would yield significant main effects favoring men. As such, the significant effects favoring men on WAIS–R Digit Span, Digit Symbol, and Block Design and the HRNB Booklet Category Test may have resulted from such differential overmatching by sex.

[9] Chi-square analyses similar to those reported by Nixon and Parsons (1991) also indicated no differences in the relative frequencies of AUD and control participants at each level of operational thought, $\chi^2(2, N = 130) = 2.62$, $p > .30$. For comparability with data reported by Nixon and Parsons, scoring levels 0 and 1 were combined as concrete operational, and scoring levels 2 and 3 were combined as transitional.

Table 3
F Values for General Linear Models Controlling, When Possible, for Baseline Performance (N = 134)

Measure	Source						
	Baseline (B)	Sex (S)	B × S	AUD (A)	B × A	S × A	B × S × A
Vocabulary (WAIS–R)	149.10**	0.03	0.50	0.00	0.03	0.95	0.65
Similarities (WAIS–R)	44.96**	0.02	2.00	2.64	0.07	0.34	0.07
Digit Span (WAIS–R)	100.07**	4.39*	2.28	0.49	0.01	0.86	0.24
Block Design (WAIS–R)	87.61**	6.27*	0.42	0.60	0.37	0.34	0.21
Digit Symbol (WAIS–R)	97.65**	0.03	0.49	0.04	0.31	0.02	0.31
Trails Part A (HRNB)	22.74**	0.00	0.87	0.00	0.20	0.13	3.25†
Trails Part B (HRNB)	32.33**	2.46	3.25†	0.15	2.35	0.04	0.08
Booklet Categories (HRNB)	46.26**	0.33	0.86	0.45	0.10	0.32	0.38
Logical Memory (WMS–R)	47.31**	0.02	2.20	0.27	0.27	5.06*	0.07
Visual Reproduction (WMS–R)	50.77**	0.35	0.52	8.32**	7.43**	0.06	0.98
Paired Associates[a] (WMS–R)	29.18**	0.66	0.12	0.36	1.84	1.07	0.30
Reflective Judgment Interview[b,c]		3.12†		0.43		0.11	
Plant Test[b,c]		0.02		0.33		0.52	
Visual Memory Span[b] (WMS–R)		3.92*		0.03		0.65	

Note. Numerator $df = 1$ in all tests reported here. Denominator $df = 126$ for all measures with baseline scores and 127 otherwise. AUD = alcohol use disorder; WAIS–R = Wechsler Adult Intelligence Scale—Revised; HRNB = Halstead–Reitan Neuropsychological Battery; WMS–R = Wechsler Memory Scale—Revised.
[a] $n = 127$ (7 individuals misunderstood directions). [b] Administered only at follow-up. [c] $n = 133$.
† $p < .09$. * $p < .05$. ** $p < .01$.

found), average standardized performance ranged from 0.01 to 0.29 of a standard deviation, with a median of 0.06. Thus, if we consider the observed magnitude of differences between the AUD and control groups, there do not appear to be large but statistically nonsignificant differences between the groups.

Such comparisons of means, however, do not provide insight into the statistical power of the general linear models reported here. Although the study was initially designed to detect differences of a clinically significant magnitude, the observed stability of performance over time and the distribution of performance on the measures permit a closer examination of statistical power. Specifically, we calculated the statistical power associated with a single degree of freedom increment in R^2 for models separately under conditions of the highest baseline covariate values, the lowest values, and no covariate. These power calculations were somewhat conservative, because we assumed that no additional variation in the dependent variable was explained by other effects in the model. According to J. Cohen's (1988) general guideline that an increment in R^2 of .02 represents a small effect size, only moderate power (.74) was observed for the high baseline condition, and power values of .45 and .38 were observed for the low and no baseline covariate conditions, respectively. The power, however, associated with a moderate increment ($r = .13$, as suggested by Cohen) was greater than .90 under all three models. Some caution is appropriate in using these guidelines when evaluating the strength of interactions. As pointed out by Aiken and West (1991) and Judd, McClelland, and Smith (1996), tests of interactions, by virtue of their collinearity with tests of main effects, may result in relatively modest increments in overall R^2.

Recent Alcohol Use

Given that there is some question as to whether the observed effects of impairment in these populations may be due to the short-term effects of alcohol consumption, we conducted a separate set of analyses examining information on past-30-day alcohol consumption and other drug use. Given the low base rates of other drug use, we elected to aggregate all other drug use into a single variable. These two consumption variables (alcohol and other drug consumption) were used as covariates in a series of general linear models exploring whether recent consumption affected performance, either singly or in combination with base rate or AUD group. These covariates were centered to zero to reduce collinearity in tests of interaction terms.

In the case of Visual Reproduction—the one measure for which an AUD main effect and an interaction of base rate with AUD were noted in our main analyses—no effects for recent alcohol and other drug consumption were found. This was true both when the consumption measures were entered as main effects and when possible interactions between recent consumption and AUD status or baseline Visual Reproduction score were considered. The main effect for AUD and the interaction of baseline with AUD remained statistically significant ($p < .01$). As such, it does not appear that recent alcohol and other drug consumption accounted for the observed main effects of WMS–R Visual Reproduction in this sample.

With minor exceptions, no main effects for recent alcohol consumption or other drug use were found, nor were there interaction terms involving these variables and AUD status or baseline performance. A statistically significant interaction between baseline and past-30-day alcohol consumption was found for the WAIS–R Vocabulary subtest, $F(1, 124) = 4.95$, $p = .03$. The standardized regression weight associated with this interaction ($\beta = .33$) was difficult to interpret, given that the direction of the interaction was positive, meaning that higher baseline scores and higher past-30-day alcohol consumption were associated with higher levels of performance on retesting or, conversely, that those individuals

who scored low at baseline and who had low patterns of past-30-day alcohol consumption scored lower than would be expected on the basis of the main effects found in the model. In addition to these difficulties, it should be noted that no significant main effects for past-30-day alcohol consumption were found, suggesting that the interaction might represent a crossover pattern. In the case of WAIS–R Block Design, a marginal main effect was found for past-30-day alcohol consumption ($\beta = -.12$, $p = .08$).

WMS–R Visual Memory Span showed a significant but small main effect for past-30-day drug use, $F(1, 124) = 4.83$, $p = .03$; $\beta = -.19$. In light of the number of analyses of covariance conducted on these data, the analyses involving recent alcohol and other drug consumption are perhaps at best suggestive of possible effects of recent consumption that merit replication in other research.

Cumulative Alcohol Exposure During the College Years

Presumably, any deficits related to AUD status would be a direct function of ethanol exposure if the critical mediating mechanism were alcohol-related neurotoxicity. As noted earlier, participants with AUDs consumed three to four times as much ethanol during their college years as controls. To investigate this issue further, we computed measures of aggregate ethanol consumption based on measures of alcohol quantity–frequency and frequency of intoxication averaged over Years 1–4 or, alternatively, Years 1–4 and Year 7. We conducted a series of regression analyses, treating these aggregate measures as continuous variables analogous to those just reported for AUD. Not surprisingly, these analyses yielded findings comparable to those obtained when AUD status was used, although differences were somewhat smaller in magnitude. Specifically, in the case of WMS Visual Reproduction, a significant main effect was found for alcohol quantity–frequency as well as for frequency of intoxication (although this effect was not statistically significant when the analysis was based on Waves 1–4 only, $p = .10$). The interaction between quantity–frequency and baseline performance was also recovered in both of the frequency of intoxication measures ($p < .002$), except when based on alcohol quantity–frequency from Waves 1–4 ($p = .09$).

A few new findings also emerged, although none were of great magnitude or consistency. Briefly, in the case of both WAIS–R Vocabulary and WMS Logical Memory, significant interactions between frequency of intoxication and sex were found based on both 4-year and 7-year intervals ($p < .05$). Men, in addition to scoring higher than women, appeared to show little impairment in scores as a function of frequency of intoxication. Women, however, seemed to show more impairment at higher levels of quantity–frequency of alcohol consumption. The size of this difference, however, was somewhat small: When women's scores were grouped into quartiles based on quantity–frequency, the highest frequency of intoxication represented approximately one sixth (for WMS Logical Memory) to one third of a standard deviation. Additional interactions of baseline and quantity–frequency from Waves 1–4 were found for WAIS–R Block Design ($\beta = .21$) and HRNB Trails Part B ($\beta = .16$; both ps $= .02$). As with the preceding discussion of 30-day alcohol use, we are cautious in interpreting these interactions given that the sign of the regression weights was positive, meaning that higher levels of quantity–frequency of alcohol use together with elevated levels of baseline performance resulted in higher performance at retesting (or, conversely, that individuals with low levels of frequency of intoxication and poor baseline performance performed lower than the main effects of the model would suggest). It should also be noted that no main effects were detected in these models for frequency of intoxication.

Discussion

The primary goal of this study was to determine whether AUD during the college years (defined as multiple AUD diagnoses) would have a deleterious effect on neuropsychological test performance and the development of cognitive skills associated with a college education. The use of a prospective design, the inclusion of a carefully screened sample of AUD and control participants matched on general intellectual functioning at baseline, and the use of both traditional neuropsychological tests and measures designed to assess higher intellectual functioning were among the strengths of this study.

The single most striking result was the general lack of differences between the AUD and control groups on our neuropsychological and cognitive outcome measures (with baseline intellectual functioning controlled). The range of measures included was broad, involving traditional tests of intellectual functioning, specialized tests designed to detect neuropsychological impairment, and measures of higher intellectual functioning related to adult cognitive development. The statistical power associated with the design of the study was only moderate for detection of small effect sizes, but it was satisfactory for detection of moderate or large effect sizes. At least at face value, these findings suggest that if cognitive effects of AUDs during the college years exist, they are relatively minor or not persistent. In defense of the design of the study, however, it must be noted that much of the previous research has not adjusted for initial academic performance and has made claims of dramatic cognitive impairment due to collegiate AUD.

Note, however, that impairment related to AUD status was found with the WMS–R Visual Reproduction subtest conditional on baseline performance. (Recall that our baseline measure of visual reproduction was the original WMS, whereas our follow-up measure was the WMS–R.) More specifically, only participants who performed poorly at baseline (i.e., below 12) were found to show prospective effects of alcohol abuse at follow-up. These effects remained after control for the effects of recent alcohol and other drug use, either alone or in combination with baseline score or AUD status. Thus, it appears that some individuals characterized by visuospatial deficits at baseline are particularly susceptible to selective alcohol-related deficits. These effects are consistent with the finding that visuospatial impairment appears to be among the most consistent and sensitive measures of alcohol-related cognitive impairment (Evert & Oscar-Berman, 1995). Also note that, in examining the larger baseline sample from which this study sample was drawn, Sher et al. (1997) found that individuals diagnosed with AUDs were significantly poorer in terms of their visuospatial performance than those without AUDs. The fact that the effect involving AUD found here was conditional based on lower baseline levels implies that individuals with low capacities on this measure (Satz, 1993) might be particularly sensitive to alcohol-related impairment.

From the standpoint of public health, the overall pattern of findings is somewhat reassuring in that this population (which is

characterized by very heavy alcohol use) does not appear to be incurring obvious damage. Although our study would appear to differ from other research showing alcohol-related deficits at all ages (Hochla & Parsons, 1982; Noonberg, Goldstein, & Page, 1985; Ryan, 1982; Ryan & Butters, 1980), even the "youngest" groups in previous studies were typically older (average ages of 30–40 years) than participants in the current study. (In addition, the research designs of these studies did not control for initial cognitive ability.) As such, the failure to find differences between our AUD and control groups is largely consonant with the "age sensitivity hypothesis" for alcohol, according to which the brains of older adults (i.e., older than 50 years) may be more vulnerable to the effect of chronic alcohol use than those of young adults (Ellis & Oscar-Berman, 1989; Noonberg et al., 1985). In addition, younger individuals (i.e., adolescents) may also be particularly vulnerable to alcohol-related impairment (Acheson et al., 1998; S. A. Brown et al., 2000; Swartzwelder et al., 1995a), suggesting that age sensitivity might be nonlinear, with the nadir of sensitivity occurring in young adulthood.

Before we conclude, a few comments on study limitations are in order. Most critical was the fact that participants were not assessed until matriculation in college, when they had already begun their exposure to alcohol but had achieved a sufficiently high level of intellectual performance to gain entrance into a university. Ideally, both cognitive functioning and alcohol consumption would be assessed over the course of students' drinking careers, beginning before initial exposure to alcohol. Longitudinal assessments would resolve cumulative lifetime exposure, trajectories of cognitive growth, and the effects of educational experiences (both before and after college matriculation) on consumption and cognitive ability.

Given the design parameters of the current study, some specific limitations should be noted. As with any individual research study, the degree to which these data can be generalized to other populations depends on the particular hypothesis and target population of interest. All of the participants were college students and thus normatively high on most neurocognitive measures; as a result, they may represent a somewhat resilient group. Other segments of the youth and young adult population might be more vulnerable to alcohol-related cognitive impairment.[10] In addition, matching those with and without AUDs on baseline general cognitive ability implies that these individuals are not representative of all young adults with or without an AUD. These data may also be overmatched (Breslow & Day, 1980; Chapman & Chapman, 1973) for inferences regarding alcohol use and cognitive performance, given that the effects of alcohol consumption may be especially pernicious in adolescence. Finally, because individuals were selected on the basis of AUD status rather than lifetime alcohol consumption, inferences about the neurotoxicity of alcohol based on these data may be limited. (However, note that ancillary analyses of time-sampled, aggregated alcohol exposure during the college years did explore dose–response relations and converged with analyses based on AUD status.)

Because the AUD participants in our study began their involvement with alcohol during adolescence, research demonstrating the particularly pernicious effects of alcohol on the developing adolescent is germane. Recall that Swartzwelder et al. (1995a, 1995b) found the effects of alcohol to be much more pronounced in the immature hippocampus of adolescent rats. Thus, there may be a critical period of brain development during adolescence characterized by susceptibility to alcohol-related insult. Recent work on clinical samples of adolescents is certainly consistent with this view (S. A. Brown et al., 2000). There is substantial reason to believe that if this research generalizes to adolescent humans, the heavily alcohol-involved participants in our study may have been particularly affected before college matriculation. Data from the larger study from which this sample was taken do not allow us to sufficiently resolve the extent and intensity of such adolescent alcohol involvement, however. If such a time-dependent pattern of vulnerability exists, it would explain some of the seeming discrepancies between the present study and the baseline study of initial AUD (Sher et al., 1997), which, in addition to the visuospatial deficits reported here, revealed differences on measures of attention and motor speed. This suggests that in a generally high cognitive functioning group such as college students, deficits in the area of visuospatial functioning are more likely to occur than deficits in other cognitive domains. Note, however, that the lack of AUD effects on Block Design scores suggests a highly specific type of visuospatial deficit that does not generalize across all visuospatial tasks.

Given the emerging mixed patterns of findings regarding the cognitive effects of AUDs in youthful populations, we need to undertake a more systematic approach to examining the problem. Specifically, population-based, longitudinal studies that begin assessment early in schooling and before heavy alcohol exposure can help resolve the seemingly diverse set of findings that characterize this underresearched area at present. Not only should syndromal diagnoses and measures of consumption be examined as predictors, but we should also consider looking at specific indicators of neuroadaptation that might be particularly relevant for understanding brain damage. For example, alcohol withdrawal has been associated with neuronal cell death (Nutt, 1999), and recent findings indicate that withdrawal appears to produce deficits in spatial memory in rodent models and that these deficits are related to cell loss in the hippocampus (Lukoyanov, Madiera, & Paula-Barbosa, 1999). Perhaps frequency of withdrawal, although rare in this population, may mark a more morbid pattern of alcohol involvement. Thus, we should move beyond gross consumption and diagnostic measures to examine individual parameters that may be hypothesized to directly affect brain integrity. We should also be attentive to individual differences in baseline functioning, in that this factor might index the capacity of the resource being studied and, consequently, moderate the relation between alcohol-related variables and their cognitive consequences.

[10] We are indebted to Sara Jo Nixon for this point.

References

Acheson, S. K., Stein, R. M., & Swartzwelder, H. S. (1998). Impairment of semantic and figural memory by acute ethanol: Age-dependent effects. *Alcoholism: Clinical and Experimental Research, 22,* 1437–1442.

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions.* Newbury Park, CA: Sage.

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev.). Washington, DC: Author.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Astin, A. W. (1993). *What matters in college? Four critical years revisited.* San Francisco: Jossey-Bass.

Baumrind, D., & Moselle, K. A. (1985). A developmental perspective on adolescent drug use. *Advances in Alcohol and Substance Abuse, 5,* 41–67.

Berkowitz, A. D., & Perkins, H. W. (1986). Problem drinking among college students: A review of recent research. *Journal of American College Health, 35,* 21–28.

Breslow, N. E., & Day, N. E. (1980). *Statistical methods in cancer research: Vol. 1. The analysis of case-control studies.* Lyon, France: International Agency for Research on Cancer.

Brown, J. L. (1989). Alcohol consumption among Kansas State University freshmen by probation and non-probation status. *Journal of Alcohol and Drug Education, 34,* 14–21.

Brown, S. A., Tapert, S. F., Granholm, E., & Delis, D. C. (2000). Neurocognitive functioning of adolescents: Effects of protracted alcohol use. *Alcoholism: Clinical and Experimental Research, 24,* 164–171.

Chapman, L. J., & Chapman, J. P. (1973). *Disordered thought in schizophrenia.* New York: Appleton-Century-Crofts.

Cohen, B. B., & Vinson, D. C. (1995). Retrospective self-report of alcohol consumption: Test–retest reliability by telephone. *Alcoholism: Clinical and Experimental Research, 19,* 1156–1161.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Derogatis, L. R. (1993). *Brief Symptom Inventory (BSI): Administration, scoring, and procedures manual* (3rd ed.). Minneapolis, MN: National Computer Systems.

Derogatis, L. R., & Spencer, P. (1982). *Brief Symptom Inventory (BSI).* Baltimore: Clinical Psychometric Research.

Ellis, R. J., & Oscar-Berman, M. (1989). Alcoholism, aging, and functional cerebral asymmetries. *Psychological Bulletin, 106,* 128–147.

Endicott, J., Andreason, N., & Spitzer, R. L. (1978). *Family history—Research diagnostic criteria (FH-RDC).* Washington, DC: National Institute of Mental Health.

Engs, R. C. (1977). Drinking patterns and drinking problems of college students. *Journal of Studies on Alcohol, 38,* 2144–2156.

Evert, D. L., & Oscar-Berman, M. (1995). Alcohol-related cognitive impairments: An overview of how alcoholism may affect the workings of the brain. *Alcohol Health and Research World, 19,* 89–96.

Fackelmann, K. (2000, October 18). Teen drinking, thinking don't mix. *USA Today,* pp. 1D, 2D.

Furedy, C., & Furedy, J. (1985). Critical thinking: Toward research and dialogue. In J. Donald & A. Sullivan (Eds.), *Using research to improve teaching* (New Directions for Teaching and Learning No. 23, pp. 51–70). San Francisco: Jossey-Bass.

Hochla, N. A. N., & Parsons, O. A. (1982). Premature aging in female alcoholics: A neuropsychological study. *Journal of Nervous and Mental Disease, 170,* 241–245.

Hughes, S. P., & Dodder, R. A. (1983). Alcohol consumption patterns among college populations. *Journal of College Student Personnel, 24,* 257–264.

Jessor, R., Donovan, J. E., & Costa, F. M. (1991). *Beyond adolescence: Problem behavior and young adult development.* New York: Cambridge University Press.

Johnson, P. O., & Neyman, J. (1936). Tests of certain linear hypotheses and their applications to some educational problems. *Statistical Research Memoirs, 1,* 57–93.

Judd, C. M., McClelland, G. H., & Smith, E. R. (1996). Testing treatment by covariate interactions when treatment varies within subjects. *Psychological Methods, 1,* 366–378.

King, P. M., & Kitchener, K. S. (1994). *Developing reflective judgment: Understanding and promoting intellectual growth and critical thinking in adolescents and adults.* San Francisco: Jossey-Bass.

Kitchener, K. S., & King, P. M. (1985). *Reflective judgment scoring manual.* (Available from K. S. Kitchener, School of Education, University of Denver, Denver, CO 80208, or P. M. King, Department of Higher Education and Student Affairs, Bowling Green State University, Bowling Green, OH 43403)

Kuhn, D., & Brannock, J. (1977). Development of the isolation of variables scheme in experimental and "nature experiment" contexts. *Developmental Psychology, 13,* 9–14.

Lukoyanov, N. V., Madiera, M. D., & Paula-Barbosa, M. M. (1999). Behavioral and neuroanatomical consequences of chronic ethanol intake and withdrawal. *Physiology and Behavior, 66,* 337–346.

McCampbell, E., & DeFilippis, N. A. (1979). The development of a booklet form of the Category Test: A preliminary report. *Clinical Neuropsychology, 1,* 33–35.

Miller, W. R., Heather, N., & Hall, W. (1991). Calculating standard drink units: International comparisons. *British Journal of Addictions, 86,* 43–47.

Moss, H. B., Kirisci, L., Gordon, H. W., & Tarter, R. E. (1994). A neuropsychological profile of adolescent alcoholics. *Alcoholism: Clinical and Experimental Research, 18,* 159–163.

Newcomb, M. D., & Bentler, P. M. (1986). Drug use, educational aspirations, and workforce involvement: The transition from adolescence to young adulthood. *American Journal of Community Psychology, 14,* 303–321.

Newcomb, M. D., & Bentler, P. M. (1988). *Consequences of adolescent drug use: Impact on the lives of young adults.* Newbury Park, CA: Sage.

Nixon, S. J. (1995). Assessing cognitive impairment. *Alcohol Health and Research World, 19,* 97–103.

Nixon, S. J., & Parsons, O. A. (1991). Alcohol-related efficiency deficits using an ecologically valid test. *Alcoholism: Clinical and Experimental Research, 15,* 601–606.

Noonberg, A., Goldstein, G., & Page, H. A. (1985). Premature aging in male alcoholics: "Accelerated aging" or "increased vulnerability"? *Alcoholism: Clinical and Experimental Research, 9,* 334–338.

Nutt, D. (1999). Alcohol and the brain: Pharmacological insights for psychiatrists. *British Journal of Psychiatry, 175,* 114–119.

Parsons, O. A. (1998). Neurocognitive deficits in alcoholics and social drinkers: A continuum? *Alcoholism: Clinical and Experimental Research, 22,* 954–961.

Parsons, O. A., & Nixon, S. J. (1993). Neurobehavioral sequelae of alcoholism. *Neurologic Clinics, 11,* 205–218.

Parsons, O. A., & Nixon, S. J. (1998). Cognitive functioning in sober social drinkers: A review of the research since 1986. *Journal of Studies on Alcohol, 59,* 180–190.

Pascarella, E. T., & Terenzini, P. T. (1991). *How college affects students.* San Francisco: Jossey-Bass.

Prendergast, M. L. (1994). Substance use and abuse among college students: A review of recent literature. *Journal of American College Health, 43,* 99–113.

Reitan, R. M. (1969). *Manual for administration of neuropsychological test batteries for adults and children.* Indianapolis, IN: Author.

Robins, L. N., Helzer, J. E., Cottler, L., & Goldring, E. (1989). *NIMH Diagnostic Interview Schedule: Version III, Revised (DIS–III-R).* St. Louis, MO: Washington University.

Robins, L. N., Helzer, J. E., Croughan, J., Williams, J. B. W., & Spitzer, R. L. (1985). *NIMH Diagnostic Interview Schedule, Version III–A.* Washington, DC: Public Health Service.

Russell, E. W. (1975). A multiple scoring method for the assessment of complex memory functions. *Journal of Consulting and Clinical Psychology, 43,* 800–809.

Ryan, C. (1982). Alcoholism and premature aging: A neuropsychological perspective. *Alcoholism: Clinical and Experimental Research, 6,* 22–30.

Ryan, C., & Butters, N. (1980). Learning and memory impairments in young and old alcoholics: Evidence for the premature-aging hypothesis. *Alcoholism: Clinical and Experimental Research, 4,* 288–293.

Salthouse, T. A. (1992). *Mechanisms of age–cognition relations in adulthood.* Hillsdale, NJ: Erlbaum.

Saltz, R., & Elandt, D. (1986). College student drinking studies: 1976–1985. *Contemporary Drug Problems, 13,* 117–159.

Satz, P. (1993). Brain reserve capacity on symptom onset after brain injury: A formulation and review of evidence for threshold theory. *Neuropsychology, 7,* 273–295.

Schandler, S. L., Brannock, J. C., Cohen, M. J., Antick, J., & Caine, K. (1988). Visuospatial learning in elementary school children with and without a family history of alcoholism. *Journal of Studies on Alcohol, 49,* 538–545.

Schandler, S. L., Cohen, M. J., & Antick, J. R. (1992). Activation, attention and visuospatial learning in adults with and without a family history of alcoholism. *Alcoholism: Clinical and Experimental Research, 16,* 566–571.

Schandler, S. L., Cohen, M. J., McArthur, D. L., Antick, J. R., & Brannock, J. C. (1991). Spatial learning deficits in adult children of alcoholic parents. *Journal of Consulting and Clinical Psychology, 59,* 312–317.

Sher, K. J., Martin, E. D., Wood, P. K., & Rutledge, P. C. (1997). Alcohol use disorders and neuropsychological functioning in first-year undergraduates. *Experimental and Clinical Psychopharmacology, 5,* 1–12.

Sher, K. J., Walitzer, K. S., Wood, P. K., & Brent, E. E. (1991). Characteristics of children of alcoholics: Putative risk factors, substance use and abuse, and psychopathology. *Journal of Abnormal Psychology, 100,* 427–448.

Sobell, L. C., Sobell, M. B., Leo, G. I., & Cancilla, A. (1988). Reliability of a timeline method: Assessing normal drinkers' reports of recent drinking and a comparative evaluation across several populations. *British Journal of Addictions, 83,* 393–402.

Stage, F., Watson, L. W., & Terrell, M. (1999). *Enhancing student learning: Setting the campus context.* Washington, DC: American College Personnel Association.

Swartzwelder, H. S., Wilson, W. A., & Tayyeb, M. I. (1995a). Age-dependent inhibition of long-term potentiation by ethanol in immature versus mature hippocampus. *Alcoholism: Clinical and Experimental Research, 19,* 1480–1485.

Swartzwelder, H. S., Wilson, W. A., & Tayyeb, M. I. (1995b). Differential sensitivity of NMDA receptor-mediated synaptic potentials to ethanol in immature versus mature hippocampus. *Alcoholism: Clinical and Experimental Research, 19,* 320–323.

Tarter, R. E., Mezzich, A. C., Hsieh, Y.-C., & Parks, S. M. (1995). Cognitive capacity in female adolescent substance abusers. *Drug and Alcohol Dependence, 39,* 15–21.

Watson, G., & Glaser, E. (1980). *Watson–Glaser Critical Thinking Appraisal.* New York: Harcourt, Brace & World.

Wechsler, D. (1944). *Wechsler Memory Scale.* New York: Psychological Corporation.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale—Revised.* New York: Psychological Corporation.

Wechsler, H., Dowdall, G. W., Davenport, A., & Castillo, S. (1995). Correlates of college student binge drinking. *American Journal of Public Health, 85,* 921–926.

Wechsler, H., Isaac, N. E., Grodstein, F., & Sellers, D. E. (1994). Continuation and initiation of alcohol use from the first to the second year of college. *Journal of Studies on Alcohol, 55,* 41–45.

Wood, M. D., Sher, K. J., & McGowan, A. D. (2000). Collegiate alcohol involvement and role attainment in early adulthood: Findings from a prospective high-risk study. *Journal of Studies on Alcohol, 61,* 278–289.

Wood, P. K. (1997). A secondary analysis of claims regarding the Reflective Judgment Interview: Internal consistency, sequentiality and intra-individual differences in ill-structured problem solving. In J. Smart (Ed.), *Higher education: Handbook of theory and research* (pp. 245–314). Edison, NJ: Agathon.

Wood, P. K., Sher, K. J., Erickson, D. J., & DeBord, K. A. (1997). Predicting academic problems in college from freshman alcohol involvement. *Journal of Studies on Alcohol, 58,* 200–210.

Received November 6, 2000
Revision received October 7, 2001
Accepted November 14, 2001 ■