Published in final edited form as:

Alcohol Clin Exp Res. 2009 May ; 33(5): 826–838. doi:10.1111/j.1530-0277.2009.00902.x.

Collateral Reports in the College Setting: A Meta-Analytic Integration

Brian Borsari and Paige Muellerleile

From the Center for Alcohol and Addictions Studies, Brown University (BB) and Providence Veterans Affairs Medical Center, Rhode Island; and Marshall University (PM), Huntington, West Virginia

Abstract

Background—The majority of research examining college drinking utilizes self-report data, and collateral reports have been used to verify participants' self-reported alcohol use

Methods—This meta-analytic integration examined the correspondence of over 970 collateral and participant dyads in the college setting.

Results and Conclusions—Results indicated that there is little bias (mean difference) between collateral estimates of participant drinking and participant's self-report. A cumulative meta-analysis revealed that this (null) effect was stable and unlikely to be altered by subsequent research or the existence of unpublished studies. Analysis of the agreement between collaterals and participant estimates (measured by intraclass correlation coefficients; ICCs) revealed moderate levels of agreement (mean ICC = 0.501). Examination of predictors of both bias and agreement in collateral and participant reports indicates a possible intentional and protective underreporting on the part of the collaterals. Ways to reduce this bias are discussed along with the value of using collaterals to verify participant self-report in the college setting.

Keywords

College; Alcohol; Collaterals; Meta-Analysis; Self-Report

It was once assumed that adults would underreport their drinking out of concern for social or legal consequences, and therefore their self-reports could not be trusted (Babor et al., 1987; Watson et al., 1984). As a result, collateral informants were employed to verify the individuals' self-report of drinking. Research participants nominate collateral informants, usually friends or family members, to provide corroboration for self-reported alcohol use and/or problems. Ideally, the collateral has frequent contact with the participant, especially in drinking situations. Statistically significant positive correlations have been found between collateral estimates of participant self-reported substance use (Babor et al., 1987; O'Farrell and Maisto, 1987), although poor collateral and participant agreement has been found in those with comorbid mental disorders (Stasiewicz et al., 2008). As more research indicates that self-reported drinking is often accurate, the use of collaterals has been discouraged as a way to obtain accurate reports of participant alcohol use (Babor and Del Boca, 1992; Babor et al., 1987, 2000; Maisto et al., 1990).

Self-report is the most common method of data collection with college students, and researchers continue to use collateral informants to confirm self-reports of drinking behavior of research participants. As in the adult literature, collaterals are most often used in the context of randomized clinical trials to assess the veracity of the self-report of participants who have received treatment. The primary downside of the use of collaterals is the significant time and effort it takes to recruit, track, and contact them in order to obtain information about the participants' drinking. Therefore it is important to determine the extent to which estimates of drinking provided by collaterals in college drinking research correspond with the self-reported alcohol use of participants. This correspondence can be expressed in a number of ways. In the present effort, we *define* bias as the mean difference between participant and collateral reports, and we define *agreement* as the reliability of the estimates that are provided by collaterals and participants (how much their estimates agree). It is important to evaluate both as a sample could exhibit no bias (no mean differences) with little or no agreement between collateral and participant reports.

College students are a unique population to collect collateral reports as there may be little motivation for college students to misrepresent their alcohol use. Heavy drinking is perceived by many students to be an integral part of the college culture for many college students (Borsari and Carey, 2001, 2006; Perkins, 2002), and is often highly salient at social functions and sporting events. Therefore it is unlikely that students view their alcohol use as socially undesirable or embarrassing. This is much different for adults and adolescents for whom heavy drinking is more likely to be seen as atypical and problematic. As a result, adults and adolescents may be more motivated to underreport their use (especially in treatment samples) than college students. Indeed, previous research with college collaterals (Laforge et al., 2005) found no evidence of systematic underreport of alcohol use. Furthermore, there was no evidence that the use of a "bogus pipeline" (where the student was told that a collateral would verify his/her self-report, but no verification actually was carried out) enhanced the accuracy of college student's self-reports. There may be situations in which collateral informants would be useful to corroborate self-report in the college setting. For example, students who have received judicial sanctions (Barnett and Read, 2005) or who are going to receive unwanted treatment as a result of their responses (as in stepped care, Borsari and O'Leary Tevyaw, 2005; Sobell and Sobell, 2000) may intentionally underreport their alcohol use. Recent research indicates that mandated students intentionally underreported their alcohol use prior to receiving an intervention due to concerns about how the information would be used (Walker and Cosden, 2007). Given the amount of research conducted with college students, we examined the research literature to further our understanding of the factors associated with collateral and participant correspondence and to determine the situations in which collateral reports can be a valuable option.

POTENTIAL PREDICTORS OF COLLATERAL/PARTICIPANT CORRESPONDENCE

Several predictors in the adult collateral literature have been hypothesized to influence the correspondence between collateral and participant reports. Therefore, we selected 7 predictors that may predict the magnitude of effect sizes of collateral and participant correspondence (with smaller bias effect sizes and larger agreement effect sizes).

Contact

The frequency of contact the collateral has with the participant is an important factor in enhancing the rates of correspondence in adults (Maisto et al., 1979; O'Farrell et al., 1984) and college students (Laforge et al., 2005). Therefore, studies with collaterals who report frequent

contact with the participants should produce smaller bias and greater agreement in collateral and participant correspondence.

Relationship

The collateral's relationship with the participant may also influence the accuracy of their estimates. Previous research with college students revealed that collaterals who were close friends of the participant demonstrated greater levels of correspondence than collaterals who were not close friends (Laforge et al., 2005). Therefore, studies that used collaterals who were close friends should exhibit smaller bias and greater agreement.

Confidence

The confidence of collaterals in their estimates of participant drinking has also been linked to collateral/participant correspondence in adults (O'Farrell et al., 1984) and college students (Laforge et al., 2005). Collaterals may have frequent contact with the participant but not in drinking situations, or the collateral may not pay attention to the participant's alcohol use. Therefore, studies wherein collaterals expressed greater confidence in their reports should exhibit smaller bias and greater agreement.

Drinking Specificity

The studies included in this meta-analysis included a wide variety of drinking behaviors, and the type of drinking behavior the collateral is asked to report may also influence his/her accuracy in corroborating the participant's drinking. Therefore, studies that used more concise questions (e.g., the highest number of drinks on one specific occasion) should report smaller bias and greater agreement than studies that used more vague questions (e.g., frequency of drinking alcohol over the past 30 days).¹

Drinking Behavior

The variety of drinking behaviors reported by the participants also fall into 3 categories that are commonly assessed: typical quantity of alcohol consumed, the frequency of alcohol consumption, and heavy drinking episodes (e.g., peak drinks in past month, binge drinking). These categories may differ in their salience for the collateral, which would influence his/her accuracy in estimating the participant's drinking. Therefore we hypothesized that heavy drinking would be most would exhibit smaller bias and greater agreement.

Mandated Status

Whether the participant is motivated to misrepresent his or her use has been recognized as affecting the validity of participant self-report (Babor and Del Boca, 1992). Two of the eight studies in this meta-analytic integration recruited mandated students who had been referred to the administration for violating campus alcohol policy (Barnett et al., 2007; Borsari and Carey, 2005). It is possible that mandated students may be more motivated to underreport their alcohol use as they had already been under the scrutiny of their school administration for their drinking (e.g., Walker and Cosden, 2007). Therefore studies that included mandated students should exhibit greater bias and less agreement.

¹We acknowledge that the length of time the collateral is asked to recall may also influence collateral and participant agreement (Babor and Del Boca, 1992). Specifically, it may be easier to recall the participant's drinking over a shorter period (e.g., over the past month) than a longer period of time (e.g., over the past year). Therefore we expected that collateral and participant agreement would be best for shorter recall periods. However, the recall period was 30 days except in 2 studies (Hagman et al., 2007; Stacy et al., 1985), precluding a robust test of this relationship.

Gender

Men and women may use alcohol in the college drinking environment differently. Specifically, alcohol use may be a more of a central role in men's social interactions than women's (Borsari and Carey, 2006). The more men interact with peers, the more drinking they report (e.g., Dorsey et al., 1999). In contrast, women may not drink as frequently or as heavily in social situations (Dowdall et al., 1998). As a result, women may have fewer opportunities to observe their female friends' alcohol use, making it more difficult for collaterals to estimate the drinking of their female friends. There is also a lesser degree of social acceptability of alcohol use (especially heavy drinking) for women (e.g., Ashmore et al., 2002; Young et al., 2005), which may lead to conscious (protective) or subconscious bias in collateral reports. Therefore, we predict that there will be greater bias and less agreement for collateral estimates of female participants' drinking.

The present effort is a meta-analytic integration of collateral– participant correspondence in their estimates of alcohol use in the college setting. Using traditional meta-analytic techniques (Mullen, 1989; Rosenthal, 1991), we examined the extent to which collateral and participant self-reports agreed. We also present a cumulative meta-analysis (Mullen et al., 2001) to determine whether the set of effect sizes gathered from research conducted to date are a sufficient and stable representation of collateral reports, which may help determine the extent to which additional resources should be devoted in collecting data from collateral informants. Finally, using an expanded dataset, we evaluate 7 factors that may moderate correspondence between participants and collaterals.

METHOD

Standard literature search techniques were utilized to conduct an exhaustive search for studies that utilized collateral informants for drinking behaviors in a college setting. Online computer searches of databases such as PubMed and PsycINFO used combinations of the words *college*, *collateral*, *alcohol*, *substance use*, and *drinking*. We also used ancestry and descendency approaches, together with the invisible college [researchers active in the domain, see Mullen (1989)]. Data available as of September 2008 were eligible for inclusion.

Studies were included if they met the following criteria: first, participants had to be college students. Second, the study had to utilize collateral informants who provided information on the participants' drinking. Third, studies and/or primary authors had to provide sufficient information from which a test of the collateral and participant discrepancy in alcohol use could be derived.² The effect sizes for bias in this analysis were based on paired sample *t*-tests, where an effect of zero indicated that there was no difference between the participants' and the collaterals' reports. A positive test indicated that the collaterals provided *lower* estimates of the participants' drinking than the participants did. A negative test indicated that the collaterals provided *higher* estimates of the participants' drinking than the participants did. The effect sizes for agreement were based on intraclass correlation coefficients (ICC) calculated and provided by the primary researchers for the purpose of this meta-analytic integration.

In the course of conducting this literature search, 11 published and unpublished articles, reports, and theses were considered for inclusion (Barnett et al., 2007; Borsari and Carey, 2005; Carey et al., 2006; Curtin et al., 2001; Greaves, 1996; Hagman et al., 2007; Laforge et al., 2005; Marlatt et al., 1998; O'Leary Tevyaw et al., 2007; Stacy et al., 1985; Tevyaw et al., 2005). One study was not included in this meta-analysis because the document did not provide enough information to derive an effect size, and the original data files had been lost (Greaves, 1996).

²The suggestions of the reviewers and action editor prompted us to request additional data and analyses from the primary study authors. As a result, in some cases the values reported in the present analysis differ from the data that were published in the original reports.

A second study (O'Leary Tevyaw et al., 2007) did not collect participant and collateral information on the same metric, making it impossible to extract an effect size. The last study that was omitted from this analysis (Tevyaw et al., 2005) recruited collaterals who were required to attend a two-session brief motivational intervention or standard education session with the participant. Because this study was qualitatively different than the others where it involved collaterals in the therapeutic context, we excluded it from the meta-analysis.

The remaining 8 studies were first analyzed as wholly independent tests of the bias hypothesis, coded separately by gender where possible. These efforts yielded 14 separate tests of collateral and participant correspondence, representing the responses of 970 collateral-participant dyads. The primary authors were contacted in order to get separate hypothesis tests of bias and agreement by participant gender for each study. All authors provided this information with 2 exceptions: one was unable to recover the data to provide these estimates (Stacy et al., 1985); and another could only recover a portion of the data from the original report (Marlatt et al., 1998). Thus, the meta-analysis of wholly independent samples was compiled by aggregating across the individual nonindependent hypothesis tests using the most conservative degrees of freedom from each nonindependent test for the aggregate statistic.

A cumulative meta-analysis (Muellerleile and Mullen, 2006) was conducted on the same data in order to determine whether additional research would be necessary to demonstrate bias; that is, whether additional research would be necessary to demonstrate that participant-collateral reports differ from one another. Following the procedures outlined in Mullen and colleagues (2001), we performed new meta-analyses each year (or wave) that new hypothesis tests were added to the database in order to demonstrate the sufficiency and stability of the database. In this case, sufficiency would indicate that no additional evidence is required to establish the existence of an effect, and stability indicates that the effect size is unlikely to change with the addition of new studies.³

The hypothesis tests for the cumulative meta-analysis on bias are described in Table 1, which represents the individual hypothesis tests of participant-collateral bias and agreement collapsed across nonindependent tests of the hypothesis. Therefore, the effect size statistics are presented as r (the weighted average of the individual nonindependent tests presented in Table 2). All hypothesis tests were coded for direction of effect (where + = the participant's self-reported use was *higher than* that of the collateral, and - = the participant's self-reported use was *lower than* that of the collateral) and using the smallest number of collateral/participant pairs (by gender) within each study.

All analyses were conducted using Mullen's Advanced BASIC meta-analytic database management system (Mullen, 1989). The significance level of an effect is provided by Z , or standard normal deviate, and its associated p -value; Z_{Fisher} is used as an indicator of effect size; and relationships between predictors and effect sizes are provided by the correlation coefficient r . BASIC provides fixed-effects analysis of the data, the underlying assumption of which is that the studies sample from a single distribution such that there is one true effect underlying the analysis. This is in contrast with a random-effects meta-analysis which assumes that the effects represent a random selection of possible effects (Schulze et al., 2003). We restricted our analysis in such a way that the studies are relatively uniform in the means by

³We decided not to conduct a cumulative meta-analysis for agreement for 2 reasons. First, because an ICC is bound by the context and question, we believed that asking researchers to collapse across these contexts and questions to provide one estimate would render a value that would be difficult to interpret. Second, because reliability is contextual, we felt that the overtime implication of cumulative analysis was secondary to the values of the ICCs within contexts. In other words, conducting a cumulative meta-analysis might establish that the average rather unreliable effect is both sufficiently established and stable. Such a result does not seem to have any practical utility if the context of the measure is what is important.

which the participant–collateral differences were collected; if we had been more inclusive in these criteria, a random-effects analysis may have been more appropriate.

The primary level data were collected at nine 4-year institutions (5 private, 4 public) with participants aged 17 to 25 years (mean age: 19.52 years). Nearly all (95%) collaterals were college students who were the same gender as the participant. Six of the studies delivered brief alcohol interventions to the participants (Barnett et al., 2007; Borsari and Carey, 2005; Carey et al., 2006; Curtin et al., 2001; Laforge et al., 2005; Marlatt et al., 1998), whereas two of the studies explicitly assessed collateral–participant bias (Hagman et al., 2007; Stacy et al., 1985).^{4,5}

In order to examine the predictors of participant–collateral bias and agreement, it was necessary to re-analyze the data from these 8 studies such that each study could contribute multiple tests of the bias hypothesis (Table 2) and the agreement hypothesis (Table 3), a procedure consistent with recommended meta-analytic methods (Lipsey and Wilson, 2001; Mullen, 1989; Schulze et al., 2003).⁶ The meta-analytic conversions of the bias effect sizes involved standard weighting procedures (i.e., $n_i - 3$) (Mullen, 1989; Rosenthal, 1991), but the conversions of ICC to Fisher's Z were somewhat different (McGraw and Wong, 1996). Although the ICC to Z_{Fisher} conversion for the particular type of ICC required to assess agreement in this case rendered an equation that was equivalent to an r to Z_{Fisher} conversion, the meta-analytic aggregation required weighting by the reciprocal of the appropriate variance (in this case, the appropriate weight was $n_i - 2$).

When possible, each hypothesis test was coded for each of the 6 predictors (with separate tests conducted separately for male and female participants). *Contact* was defined as the percentage of the collaterals in the study who reported seeing the participant “every day” or “nearly every day.” The *relationship* between the collateral and participant was the percentage of participants who reported being a “close” or “best” friend. This predictor was not significantly correlated with contact. The collaterals' *confidence* was obtained by calculating the percentage of participants in each study who reported being “very” or “mostly” confident in the estimates

⁴Examination of the effect sizes in intervention and survey studies revealed a nonsignificant ($p = 0.183$), negligible ($Z_{\text{Fisher}} = 0.041$, $r = 0.041$) effect that was obtained for the 3 hypothesis tests from survey studies (Hagman et al., 2007; Stacy et al., 1985) and a nonsignificant ($p = 0.548$), null ($Z_{\text{Fisher}} = -0.011$, $r = -0.011$) effect for the 11 hypothesis tests from intervention studies. The difference between the magnitudes of these 2 effects was not significant ($p = 0.377$), indicating that the type of study did not appear to systematically influence collateral and participant agreement.

⁵Some studies in the meta-analytic database recorded collateral estimates of alcohol-related problems (Barnett et al., 2007; Laforge et al., 2005; Marlatt et al., 1998; Stacy et al., 1985). However, we did not analyze the participant and collateral estimates on problems for 2 reasons. First, correspondence between participant and collateral reports of problems in adult samples has ranged widely, with the highest rates of agreement being found for summary scores of measures of severity and impairment of alcohol use that are more observable and concrete (O'Farrell and Maisto, 1987; Polich, 1982). Lower agreement has been found on measures of more minor consequences (e.g., minor withdrawal symptoms) that may not be observable such as having a hangover (Connors and Maisto, 2003). This tendency is especially problematic in a college sample in which a wide variety of minor problems can be endorsed (Kahler et al., 2004). Each study used different measures of problems, all of which included consequences that were difficult for collaterals to observe (e.g., felt that she/he needed more alcohol than she/he used to in order to get the same effect). Indeed, correlations between participant and collateral estimates of problems were consistently lower than those of alcohol use in these studies, ranging from 0.29 (Marlatt et al., 1998) to 0.52 (Laforge et al., 2005). Second, the studies used summary scores for their problem measures, making it impossible to determine agreement on specific items. Furthermore, many of the problem measures combined both concrete behaviors (e.g., “got sick”) and behaviors that were difficult for the collateral to observe (e.g., “had a bad time”). As a result, we questioned the utility of evaluating the agreement of summary scores as identical scores could be derived by endorsing items that were very different in terms of severity of alcohol-related problems.

⁶We recognize the assumption that hypothesis tests from the same study are independent is patently false. For example, the 20 hypothesis tests reported by Borsari and Carey (2005) were derived from the same participants at the same time. Without making this assumption of independence, 2 options are available to avoid nonindependent hypothesis tests. One option is to choose the “best” hypothesis test from each study. Given the variety in hypothesis tests within each study in the type of alcohol use assessed, confidence in the collateral report, and so on, it would have been difficult to determine which hypothesis test was the “best” one. The second option is to pool the results from all the hypothesis tests to create a single test of participant–collateral agreement. Nevertheless, both these alternatives create more problems with assumptions and arbitrariness than the present assumption of independence, and they fail to allow a nuanced examination of the predictors of participant–collateral agreement. Further, our results indicate that the degree of distortion engendered by the assumption of independence in the nonindependence hypothesis tests in the main database is (at worst) tolerable.

they provided. *Drinking specificity* addressed methodological features of the primary studies' examination of drinking behavior. Across the 8 studies, collaterals were asked to report on 7 different drinking behaviors in the study: drinks per week; frequency of drinking alcohol; heavy drinking episodes (defined as 5 or more drinks in one sitting for men, 4 or more for women; Wechsler et al., 1995); typical number of drinks per occasion; highest number of drinks on one occasion; frequency of drinking game participation; and drinking days per week. An additional 6 drinking questions were asked to collaterals in 2 studies (Hagman et al., 2007; Stacy et al., 1985): the frequency and typical quantity of beer, wine, and hard liquor consumed by the participants. Three judges were asked to rate all 13 drinking behaviors on a scale of 0 (concise: defined as "specific, precise, explicit") to 100 (vague: defined as "not precisely defined, determined, or distinguished"). Reliability of the judges' ratings was good (intraclass correlation coefficient = 0.79). The *drinking specificity* predictor was the mean of the judges' ratings. The *drinking behaviors* predictor was created by categorizing each of the drinking behaviors into those that assessed quantity (e.g., typical number of drinks per occasion), frequency (e.g., frequency of drinking alcohol), or heavy drinking (e.g., binge drinking). Finally, studies that included *mandated students* were dummy coded (mandated = 1).

RESULTS

General Effects

Table 1 displays the results of the 14 wholly independent hypothesis tests of bias obtained by averaging the individual effect sizes presented in Table 2. With each hypothesis test weighted by the sample size, there was a nonsignificant, practically null effect ($\bar{r} = 0.007$; $Z_{\text{Fisher}} = 0.007 \pm 0.064$, $p = 0.290$) of difference between collateral and participant reports. Therefore, there appeared to be no practical discrepancy between participant and collateral estimates of alcohol use. A diffuse comparison of these effect sizes [or homogeneity test, $\chi^2_{(13)} = 17.833$, $p = 0.164$] indicates that the effect sizes reported in Table 1 are not significantly different from one another. Table 1 also displays the results of the 14 wholly independent hypothesis tests of agreement obtained by averaging the individual effect sizes presented in Table 3. With each hypothesis test weighted by the reciprocal variance for the ICC, there was a significant, moderate effect (average ICC = 0.537, $Z_{\text{Fisher}} = 0.601 \pm 0.065$, $p < 0.001$) of agreement between collateral and participant reports. A diffuse comparison of these effect sizes [$\chi^2_{(13)} = 8.204$, $p = 0.830$] indicates that the effect sizes reported in Table 1 are not significantly different from one another. For both bias and agreement, a funnel plot of the effect sizes against the sample sizes, though sparse, revealed no evidence of any publication bias that would produce results different from those reported above.

The combined results of the 102 tests of the magnitude of collateral/participant bias with each hypothesis test weighted by its reciprocal variance estimate ($n_i - 3$) revealed a significant,⁷ practically null ($\bar{r} = 0.029$; $Z_{\text{Fisher}} = 0.029 \pm 0.025$, $p = 0.035$) effect. Therefore, there appeared to be little bias between participant and collateral estimates of alcohol use. A diffuse comparison of these effect sizes [$\chi^2_{(101)} = 217.70$, $p < 0.001$] indicates that the effect sizes reported in Table 2 are significantly different from one another. The independent and nonindependent databases yielded the same conclusion: there was little bias between participant and collateral reports of drinking behavior.

Finally, Table 3 presents 102 tests of the magnitude of collateral/participant agreement, which revealed a significant, moderate (average ICC = 0.501; 0.551 ± 0.026 , $p < 0.001$) effect. This

⁷All confidence intervals and significance levels are computed based on nonindependent data. The nonindependent nature of these tests is likely to inflate the Type I error rate, so these tests should be interpreted with caution.

result does not reach the accepted standard reliability estimate for standardized measures by trained raters (Nunnally and Bernstein, 1994). However, in a context where participants and collaterals are not trained to produce reliable estimates on a variety of drinking behaviors during a recall period of at least 30 days, this level of agreement may be considered good (Wong, personal communication, 12/1/2008). A diffuse comparison of these effect sizes [$\chi^2_{(101)} = 273.393, p < 0.001$] indicates that the effect sizes reported in Table 3 are significantly different from one another. The results of this analysis indicated considerable variability among the effect sizes of the ICCs.

Cumulative Meta-Analytic Integration of Collateral and Participant Bias

We performed a cumulative meta-analysis (Muellerleile and Mullen, 2006; Mullen et al., 2001) on the wholly independent bias dataset to examine the evidence that participant self-reports of drinking behavior were significantly different than collateral reports of the participant's drinking behavior. Figure 1 displays the results of the cumulative meta-analysis, combining the average effect, the fail-safe ratio, and the cumulative slope. The average effect size or mean effect across the 7 waves of the database hovered under and then settled just above zero.

Because the fail-safe ratio did not rise above 1 over the course of the database, there is insufficient evidence for bias; in other words, there appears to be no evidence for a discrepancy between participants' and collaterals' reports. The fail-safe ratio developed by Mullen and colleagues (2001) was intended to address the "file drawer" problem, meaning that there are likely to be some number of studies with null results that were never published, languishing in researchers' file drawers. Rosenthal (1979) developed a means of calculating a "fail-safe" number of unavailable studies with null results that would be required to bring the significance level down to just significant at $p = 0.05$. Rosenthal also suggested (Rosenthal, 1979, 1984; Rosenthal and Hall, 1981) that a reasonable benchmark against which the fail-safe number could be compared was $5k + 10$, where k refers to hypothesis tests. That is, there were unlikely to be 5 times the number of retrieved studies remaining unpublished in file drawers, and the value of 10 was added to ensure that the smallest database must have a fail-safe number of at least 15. The fail-safe ratio is the ratio of Rosenthal's fail-safe number to the benchmark value, and when it exceeds 1, it has reached a level that will tolerate unpublished null results. In this case, the indicator did not rise above 1, so there was no evidence that this database was sufficient to show an effect different from zero across the 6 waves in the database. Thus, there was no evidence that a bias exists between participant and collateral reports of drinking.

The cumulative slope indicates that the database is stable. The cumulative slope is the absolute value of the slope of the line derived from regressing effect size upon the number of hypothesis tests in the database. As the cumulative slope approaches 0, the effect becomes stable or less likely to change with the addition of more tests of the hypothesis. In the case of collateral reports of participants' drinking behavior, the cumulative slope indicates stability in the database by wave four, when the eighth hypothesis test was conducted in 2005. At this point, there was enough evidence to suggest that this practically null discrepancy between collateral and participant reports of participant drinking behavior would not change over time as additional tests of the hypothesis were added. This was the case in our cumulative analysis, and there is no reason to believe that additional tests of the hypothesis will suddenly reveal discrepancies between collateral and participant reports.

Predictors of Participant and Collateral Correspondence

To better understand the observed variation in the correspondence between collateral and participant responses, we examined the influence of the 7 predictors on bias and agreement.

Contact—A significant, positive relationship ($r = 0.315, p < 0.001$) emerged between the frequency of the collateral's contact with the participant and the bias effect size. Therefore the more contact between the collateral and the participant, the greater the positive discrepancy in their estimates, such that the collaterals underestimated the participants' alcohol use. A nonsignificant, negative relationship ($r = -0.067, p = 0.092$) emerged between the frequency of the collateral's contact with the participant and the collateral/participant agreement. Therefore contact between the collateral and participant did not influence the reliability of their estimates.

Relationship—A significant, positive relationship ($r = 0.449, p < 0.001$) emerged between the relationship the collateral had with the participant and collateral/participant bias effect size. Therefore the closer the relationship between the collaterals and participants, the greater the positive discrepancies in their estimates, indicating that the collaterals underestimated the participants' alcohol use. A nonsignificant negative relationship ($r = -0.072, p = 0.081$) emerged between the relationship the collateral had with the participant and collateral/participant agreement.

Confidence—A significant, positive relationship ($r = 0.384, p < 0.001$) was observed between the collateral's confidence in their ratings and participant/collateral bias. Therefore more confident the collateral was in his or her estimate of the participant's drinking, the greater the positive discrepancy in their estimates, such that the collaterals underestimated the participants' alcohol use. A positive, nonsignificant relationship ($r = 0.012, p = 0.491$) emerged between the collaterals' confidence in their ratings and the collateral/participant agreement.

Drinking Specificity—A significant, positive relationship ($r = 0.145, p = 0.007$) emerged between drinking specificity and the participant and participant/collateral bias. Therefore, the more vague the question assessing the alcohol behaviors, the more likely the collateral underestimated the participants' alcohol use. A negative, nonsignificant relationship ($r = -0.034, p = 0.213$) emerged between the specificity of the drinking behavior and collateral/participant agreement.

Drinking Behavior—A nonsignificant, practically null ($\bar{r} = -0.013; Z_{\text{Fisher}} = -0.013 \pm 0.041, p = 0.578$) effect was obtained for the 33 bias hypothesis tests that assessed quantity, indicating that the collateral reports of drinking quantity did not differ from the participants' reports. A null ($\bar{r} = -0.011; Z_{\text{Fisher}} = -0.011 \pm 0.041, p = 0.704$) effect was obtained for the 41 bias hypothesis tests that assessed frequency, indicating that the collateral reports of drinking frequency did not differ from the participants' reports. Finally, a nonsignificant, small, positive ($\bar{r} = 0.163; Z_{\text{Fisher}} = 0.164 \pm 0.053, p = 0.220$) effect was obtained for the 28 bias hypothesis tests that assessed heavy drinking, indicating that collaterals in these studies underreported participant's heavy drinking. Although there was no difference in the magnitude of the effect size between quantity and frequency, the effect size for heavy drinking was significantly different than both quantity and frequency. These results suggested that collaterals were more accurate in their estimates of drinking quantities or frequencies, but less accurate in their estimates regarding participants' heavy drinking episodes.

Regarding collateral and participant agreement on different types of drinking behaviors, significant and moderate effects were found for quantity (average ICC = 0.500; $Z_{\text{Fisher}} = 0.549 \pm 0.041, p \leq 0.001$), frequency (average ICC = 0.491; $Z_{\text{Fisher}} = 0.538 \pm 0.041, p < 0.001$), and heavy drinking (average ICC = 0.521; $Z_{\text{Fisher}} = 0.578 \pm 0.053, p < 0.001$). The agreement effect sizes for each drinking behavior were not significantly different.

Mandated Status—A significant, small, positive ($\bar{r} = 0.220; Z_{\text{Fisher}} = 0.224 \pm 0.063, p < 0.001$) effect was obtained for the 52 hypothesis tests that assessed collateral/participant

agreement in mandated students, indicating that collaterals in these studies underreported participant drinking. A nonsignificant, negative ($\bar{r} = -0.013$; $Z_{\text{Fisher}} = -0.013 \pm 0.028$, $p = 0.320$) effect was obtained for the 50 hypothesis tests that included nonmandated participants. The difference between the magnitudes of the 2 effects was highly significant ($p < 0.001$), indicating that the collateral and participant discrepancies in mandated students were significantly greater than those in nonmandated samples.

A significant effect (average ICC = 0.419; $Z_{\text{Fisher}} = 0.446 \pm 0.065$, $p < 0.001$) was obtained for the 52 hypothesis tests that assessed collateral/participant agreement in mandated students, indicating reasonable participant/collateral agreement. Similarly, a significant, moderate (average ICC = 0.516; $Z_{\text{Fisher}} = 0.571 \pm 0.028$, $p < 0.001$) effect was obtained for the 50 hypothesis tests that included nonmandated participants. The difference between the magnitudes of the 2 effects was significant ($p = 0.044$), indicating that agreement in nonmandated samples was somewhat better than agreement in mandated samples.

Gender—Because the primary authors provided us with separate tests of bias by gender, we were able to examine this variable in the independent database (Table 1). A nonsignificant, positive ($\bar{r} = 0.041$; $Z_{\text{Fisher}} = 0.041 \pm 0.120$, $p = 0.243$) effect was obtained for the 6 independent hypothesis tests that assessed participant/collateral bias in male students. A nonsignificant, negative ($\bar{r} = -0.037$; $Z_{\text{Fisher}} = -0.037 \pm 0.096$, $p = 0.670$) effect was obtained for the 7 independent hypothesis tests that included female participants. The difference between the magnitudes of the 2 effects was not significant ($p = 0.356$), indicating that the collateral and participant discrepancies in male and female participants were not significantly different. Similar results (both magnitude and direction) were obtained in the larger, nonindependent database (see Table 2).

To test gender differences in agreement, we examined the independent database (Table 1). A significant effect (average ICC = 0.505, $Z_{\text{Fisher}} = 0.556 \pm 0.119$, $p < 0.001$) was obtained for the 6 independent hypothesis tests that assessed participant/collateral agreement in male students. A diffuse comparison of these 6 independent hypothesis tests of agreement indicated that they were not significantly different from one another [$\chi^2_{(5)} = 0.832$, $p = 0.975$]. A significant effect (average ICC = 0.552, $Z_{\text{Fisher}} = 0.621 \pm 0.092$, $p < 0.001$) also was obtained for the 7 independent hypothesis tests that included female participants, and a diffuse comparison of these effect sizes revealed that they were not significantly different [$\chi^2_{(6)} = 6.608$, $p = 0.359$]. The difference between the magnitudes of the 2 effects was not significant ($p = 0.390$), indicating that the collateral and participant agreement in male and female participants were not significantly different. As was the case with bias, similar results (both magnitude and direction) were obtained in the larger, nonindependent database (see Table 3).

DISCUSSION

Findings from this meta-analytic integration did not provide evidence that collaterals exhibited significant bias, but good agreement, when estimating the alcohol use of participants. The bias that was detected was consistent with previous research (Laforge et al., 2005), in which the majority of collaterals (73%) underreported participant alcohol use. The cumulative meta-analysis indicated that the bias between collaterals and participants was sufficient and stable, unlikely to be altered by subsequent research using collaterals in the college setting. Although the agreement between collateral and participant reports was quite variable (ICC's ranging from 0.0 to 1.0), the average ICC of 0.51 indicated that the reliability of collateral estimates was good, especially given the circumstances under which the reliability estimates were produced. To better understand these findings, we examined 7 predictors of bias and agreement.

Regarding bias, 6 of the 7 predictors we examined were significantly related to collateral and participant agreement. First, inconsistent with previous research (Laforge et al., 2005; Maisto et al., 1979; O'Farrell et al., 1984), frequency of contact was not associated with a less bias between collateral and participant reports. Instead, collaterals who saw the participant more frequently reported less alcohol use than the participant. Second, the closer the relationship between the collaterals and the participants, the larger the positive discrepancies in the estimates. This is an unexpected finding as closer relationships had been associated with more agreement in previous research with college students (Laforge et al., 2005). However, there are several reasons why having a close relationship with the participant may not lead to less bias in the college setting: the two may not consistently drink together or may drink together so frequently that the different drinking experiences may come together in memory as some kind of weighted average that it is accurate overall but inaccurate for any one drinking occasion. In addition, collaterals who are close friends of the participant may be less comfortable reporting on the participants' alcohol use, leading to lower, more conservative estimates. Third, collaterals who reported greater confidence in their estimates were more likely to provide estimates that were lower than the participants' self reported use. This discrepancy between confidence and bias is consistent with the literature on eyewitness testimony, which has found little connection between the confidence and accuracy (Brewer et al., 2002; Smith et al., 1989). Fourth, consistent with the adult literature on the validity of self-report (Babor and Del Boca, 1992; Babor et al., 1987), the more specific the drinking behavior, the greater the agreement between the collateral and participant estimates. The more specific the drinking behavior reported, it is more likely that the collateral estimates matched the participants' self-reported alcohol use. Fifth, collaterals and participants exhibited less discrepancy on general drinking behaviors, frequency, and quantity as opposed to heavy drinking. This was unexpected as we thought that heavy drinking episodes would be more salient and memorable to the collateral and the participant, resulting in greater agreement (e.g., Perkins, 2002). However, it is possible that the participant and collateral reported on *different* heavy drinking episodes, resulting in the discrepancy. Finally, mandated participants and their collaterals also demonstrated less agreement than participants who had not received an alcohol infraction. Specifically, collaterals reported that the participants were drinking *less* than the participants reported, whereas collaterals for nonmandated students demonstrated practically no discrepancy.

Taken together, this pattern of findings suggests a possible protective bias on the part of the collateral. Specifically, there appears to be a small yet significant tendency for collaterals who are close friends to underestimate the participants' use, a trend also evident for collaterals who are reporting on the use of participants who have received alcohol violations. This bias may be the result of demand characteristics created by the method of collecting collateral data. Participants in this research underwent a detailed consenting process, where it would have been made clear that their responses would be kept confidential and not be used against them in any way. If participants wanted to underreport their use, then a negative effect would have been found—the participants' self-reported use would have been lower than that of the collateral. Instead, the opposite occurred. Collaterals were most often contacted by phone, and given a verbal description of the study, and then asked to provide information. As a result, the collaterals may have been concerned that their responses would be used against the participants, often a close friend. Thus, collaterals may have underestimated the drinking of the participants in order to protect them from consequences of their alcohol use. The observed trend to underestimate the use of the participants may also have resulted from an intentional bias on the part of the participant. That is, the participant may have underreported his or her use to the collateral, who in turn reported this inaccurate perception of drinking to the researchers. There are a number of reasons why the participant would want to underreport his or her use to the collateral, including being embarrassed about continued drinking after a referral incident (for mandated students), concern over how the information will be used (e.g., Walker and Cosden,

2007), or if the collateral had expressed concern about the participant's drinking (for a close friend). These social pressures may have led to an underrepresentation of alcohol use to peers that was not replicated in the context of a confidential survey.

In general, the results of the agreement analysis indicate that the estimates of reliability do not vary systematically with most of the predictors we examined. In other words, the average ICC obtained in this analysis did not change depending on the amount of contact or type of relationship between the participants and collaterals, the degree of confidence that collaterals had in their estimates, the specificity of the drinking assessments, the type of drinking behavior, or the participant gender. The only significant predictor of agreement was mandated status: collaterals and mandated students demonstrated significantly poorer agreement (average ICC = 0.419) than collaterals and nonmandated participants (average ICC = 0.516). It is likely that this discrepancy may have been the result of intentional misrepresentation of alcohol use, especially the protective behavior of the collateral for the mandated student. It is also possible that the observed differences may be partially due to unintentional error resulting from being asked to report the drinking behavior of a friend over an extended period of time.

These findings have direct implications for researchers in the college setting. If the purpose of the use of collaterals is to verify participant self-report, it appears that they will do so with little bias and relatively good agreement. However, to limit bias, specific questions should be asked regarding the drinking behavior of the participant. Furthermore, clear, concise instructions should be provided to the collateral, including how the information will be used, especially if the participants are close friends or mandated students. It may be preferable to email the collateral detailed instructions and assurances or refer them to a project webpage that contain this information. Therefore it is important to implement proper procedures for accurate self-report. These include consideration of the sensitivity of the information sought, the specificity of the information the participant is asked to report, the personal characteristics of the participant, the time period covered by the measures, and the demand characteristics of the situation (Babor and Del Boca, 1992; Babor et al., 1987; Midanik, 1988). In addition, care should be taken when using collaterals that were actively involved in the treatment of the participant (e.g., Tevyaw et al., 2005).

The findings of the study should be considered in the context of some limitations. First, this meta-analysis examined the agreement between collateral estimates of the alcohol use of participants and the participants' self-report. Thus, it is impossible to determine from these data what the actual alcohol use of the participants may be. There is no "gold standard" of alcohol consumption, and many college students are unaware of standard serving sizes and tend to underestimate the amount that they consume (White et al., 2005). Therefore, it is possible that both participants and collaterals' estimates of actual drinking may be inaccurate. Second, we were not able to examine dyad-specific characteristics of the collaterals and participants. Instead, given the nature of the data we reconstructed from the available studies, we could only examine sample-specific predictors. The examination of personal factors that may contribute to correspondence would also be valuable (e.g., hostility, degree of protectiveness of the participant). Third, the homogeneity of college students may have contributed to the lack of large effects for the predictors. It may be that these predictors play more of a role in the accuracy of collaterals in more heterogeneous samples (e.g., adult alcoholics), in which there is a wide variety of drinking contexts and consequences. Finally, although these predictors were significantly related to the collateral and participant bias, it should be noted that their overall effect was rather small. Therefore it is possible that the significance achieved by some predictors may have resulted from Type I error (inflated significance due to re-sampling from the same study) rather than from an effect with real-world implications.

Perhaps the best way to summarize the findings of this meta-analytic integration is that, in the general college context, collaterals may slightly underreport the drinking of participants. The agreement between participants and collaterals is reasonably good, although less so with mandated students. Although this pattern of findings may be the result of intentional underreporting on the part of the collaterals and/or participants, it is also possible that they reflect the difficulty of estimating a variety of drinking behaviors over an extended period of time across different contexts. As a result, there does not appear to be a significant return on the considerable investment of time and money required to collect collateral data with college student drinkers. Therefore, researchers should consider carefully the purpose and design of their studies before implementing self-report verification procedures via collaterals with college student samples.

ACKNOWLEDGMENTS

This work was supported by National Institute on Alcohol Abuse and Alcoholism Grant R01-AA015518 to B. Borsari. The authors thank all of the researchers who graciously and promptly contributed their data, and also recognize the importance of the guidance of Alice Eagly and Seok P. Wong in converting ICCs to effect sizes for meta-analytic integration. Finally, the authors would also like to acknowledge the late Brian Mullen for his thoughtful comments and guidance in the conceptualization of this project.

REFERENCES

- Ashmore RD, Del Boca FK, Beebe M. "Alkie", "Frat Brother", and "Jock": perceived types of college students and stereotypes about drinking. *J Appl Soc Psychol* 2002;32:885–907.
- Babor, TF.; Del Boca, FK. Just the facts: enhancing measurement of alcohol consumption using self-report methods. In: Litten, R.; Allen, J., editors. *Measuring Alcohol Consumption*. New York: Humana; 1992. p. 3-19.
- Babor TF, Steinberg K, Anton R, Del Boca FK. Talk is cheap: measuring drinking outcomes in clinical trials. *J Stud Alcohol* 2000;61:55–63. [PubMed: 10627097]
- Babor TF, Stephens RS, Marlatt GA. Verbal report methods in clinical research on alcoholism: response bias and its minimization. *J Stud Alcohol* 1987;48:410. [PubMed: 3312821]
- Barnett NP, Murphy JG, Colby SM, Monti PM. Brief motivational intervention vs. computerized alcohol education with college students mandated to intervention. *Addict Behav* 2007;32:2529–2548. [PubMed: 17707594]
- Barnett NP, Read JP. Mandatory alcohol intervention for alcohol-abusing college students: a systematic review. *J Subst Abuse Treat* 2005;29:147. [PubMed: 16135343]
- Borsari B, Carey KB. Peer influences on college drinking: a review of the research. *J Subst Abuse* 2001;13:391–424. [PubMed: 11775073]
- Borsari B, Carey KB. Two brief alcohol interventions for mandated college students. *Psychol Addict Behav* 2005;19:296–302. [PubMed: 16187809]
- Borsari B, Carey KB. How the quality of peer relationships influences college alcohol use. *Drug Alc Rev* 2006;25:361–370.
- Borsari B, O'Leary Tevyaw T. Stepped care: a promising treatment strategy for mandated college students. *NASPA J* 2005;42:381–397.
- Brewer N, Keast A, Rishworth A. The confidence-accuracy relationship in eyewitness identification: the effects of reflection and disconfirmation on correlation and calibration. *J Exp Psychol Appl* 2002;8:44–56. [PubMed: 12009176]
- Carey KB, Carey MP, Maisto SA, Henson JM. Brief motivational interventions for heavy college drinkers: a randomized controlled trial. *J Consult Clin Psychol* 2006;74:943–954. [PubMed: 17032098]
- Connors GJ, Maisto SA. Drinking reports from collateral individuals. *Addiction* 2003;98:21–29. [PubMed: 14984239]
- Curtin L, Stephens RS, Bonenberger JL. Goal setting and feedback in the reduction of heavy drinking in female college students. *J Coll Stud Psychotherapy* 2001;15:17.

- Dorsey AM, Scherer CW, Real K. The college tradition of 'drink 'til you drop': the relation between students' social networks and engaging in risky behaviors. *Health Commun* 1999;11:313.
- Dowdall GW, Crawford M, Wechsler H. Binge drinking among American college women. *Psychol Women Q* 1998;22:705–715.
- Greaves, CK. PhD Dissertation. Blacksburg, VA: Department of Psychology, Virginia Polytechnic Institute; 1996. An evaluation of two approaches to drinking risk reduction with college students: cognitive-behavioral skills training and motivational feedback.
- Hagman BT, Clifford PR, Noel NE, Davis CM, Cramond AJ. The utility of collateral informants in substance use research involving college students. *Addict Behav* 2007;32:2317–2323. [PubMed: 17329032]
- Kahler CW, Strong DR, Read JP, Palfai TP, Wood MD. Mapping the continuum of alcohol problems in college students: a Rasch Model analysis. *Psychol Addict Behav* 2004;18:322. [PubMed: 15631604]
- Laforge RG, Borsari B, Baer JS. The utility of collateral informant assessment in college alcohol research: results from a longitudinal prevention trial. *J Stud Alcohol* 2005;66:479. [PubMed: 16240555]
- Lipsey, MW.; Wilson, DB. *Practical Meta-Analysis*. Thousand Oaks, CA: Sage; 2001.
- Maisto SA, McKay JM, Connors GJ. Self-report issues in substance abuse: state of the art and future directions. *Behav Assess* 1990;12:117–134.
- Maisto SA, Sobell LC, Sobell MB. Comparison of alcoholics' self reports of drinking behavior with reports of collateral informants. *J Consult Clin Psychol* 1979;47:106–112. [PubMed: 429642]
- Marlatt GA, Baer JS, Kivlahan DR, Dimeff LA, Larimer ME, Quigley LA, Somers JM, Williams E. Screening and brief intervention for high-risk college student drinkers: results from a 2-year follow-up assessment. *J Consult Clin Psychol* 1998;66:604. [PubMed: 9735576]
- McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychol Methods* 1996;1:30–46.
- Midanik LT. Validity of self-reported alcohol use: a literature review and assessment. *Br J Addict* 1988;83:1019–1030. [PubMed: 3066418]
- Muellerleile P, Mullen B. Sufficiency and stability of evidence for public health interventions using cumulative meta-analysis. *Am J Public Health* 2006;96:515–522. [PubMed: 16449603]
- Mullen, B. *Advanced BASIC Meta-Analysis*. Hillsdale, New Jersey: Lawrence Erlbaum Associates; 1989.
- Mullen B, Muellerleile P, Bryant B. Cumulative meta-analysis: a consideration of indicators of sufficiency and stability. *Pers Soc Psychol Bull* 2001;27:1450–1462.
- Nunnally, JC.; Bernstein, IH. *Psychometric Theory*. Vol. 3rd edn.. New York: McGraw-Hill; 1994.
- O'Farrell TJ, O'arrell TJ, Cutter HSG, Batog RD, Dentch G, Fortgang J. Correspondence between one-year retrospective reports of pre-treatment drinking by alcoholics and their wives. *Behav Assess* 1984;6:263–274.
- O'arrell TJ, Maisto SA. The utility of self-report and biological measures of alcohol consumption in alcoholism treatment outcome studies. *Adv Behav Res Ther* 1987;9:91–125.
- O'Leary Tevyaw T, Borsari B, Colby SM, Monti PM. Peer enhancement of a brief motivational intervention with mandated college students. *Psychol Addict Behav* 2007;21:114–119. [PubMed: 17385961]
- Perkins HW. Social norms and the prevention of alcohol misuse in collegiate contexts Suppl. *J Stud Alcohol* 2002;14:164–172.
- Polich JM. The validity of self-reports in alcoholism research. *Addict Behav* 1982;7:123–132. [PubMed: 7102443]
- Rosenthal R. The "file drawer problem" and tolerance for null results. *Psychol Bull* 1979;86:638–641.
- Rosenthal, R. *Meta-Analytic Procedures for Social Research*. Beverly Hills, CA: Sage; 1984.
- Rosenthal, R. *Meta-Analytic Procedures for Social Research*. Thousand Oaks, CA: Sage; 1991.
- Rosenthal R, Hall J. Critical values of Z for combining independent probabilities. *Replications Soc Psychol* 1981;1:1–6.
- Schulze, R.; Holling, H.; Bohning, D. *Meta-Analysis*. Seattle: Hogrefe & Huber; 2003.
- Smith VL, Kassin SM, Ellsworth PC. Eyewitness accuracy and confidence: within-versus between-subjects correlations. *J Appl Psychol* 1989;74:356–359. [PubMed: 2708267]

- Sobell MB, Sobell LC. Stepped care as a heuristic approach to the treatment of alcohol problems. *J Consult Clin Psychol* 2000;68:573–579. [PubMed: 10965632]
- Stacy AW, Widaman KF, Hays R, DiMatteo MR. Validity of self reports of alcohol and other drug use: a multitrait-multimethod assessment. *J Pers Soc Psychol* 1985;49:219–232. [PubMed: 4020615]
- Stasiewicz PR, Vincent PC, Bradizza CM, Connors GJ, Maisto SA, Mercer ND. Factors affecting agreement between severely mentally ill alcohol abusers' and collaterals' reports of alcohol and other substance abuse. *Psychol Addict Behav* 2008;22:78–87. [PubMed: 18298233]
- Tevyaw, TO.; Colby, SM.; Borsari, B.; Luboyeski, EJ.; Monti, PM. Including peers in BMI for mandated college students. A Look Towards the Future: Outcome Data on Innovative Brief Interventions (Chan K chair); Symposium presented at the annual conference of the Research Society on Alcoholism; Santa Barbara, CA. 2005.
- Walker S, Cosden M. Reliability of college student self-reported drinking behavior. *J Subst Abuse Treat* 2007;33:405–409. [PubMed: 17499957]
- Watson CG, Tilleskjor C, Hoedeccheck-Schow EA, Pucel J, Jacobs L. Do alcoholics give valid self-reports? *J Stud Alcohol* 1984;45:344–348. [PubMed: 6482438]
- Wechsler H, Dowdall GW, Davenport A, Rimm EB. A gender-specific measure of binge drinking among college students. *Am J Public Health* 1995;85:982–985. [PubMed: 7604925]
- White AM, Kraus CL, Flom JD, Kestenbaum LA, Mitchell JR, Shah K, Swartzwelder HS. College students lack knowledge of standard drink volumes: implications for definitions of risky drinking based on survey data. *Alcohol Clin Exp Res* 2005;29:631–638. [PubMed: 15834229]
- Young AM, Morales M, McCabe SE, Boyd CJ, d'Arcy H. Drinking like a guy: frequent binge drinking among undergraduate women. *Subst Use Misuse* 2005;40:241–267. [PubMed: 15770887]

Fig. 1.
Results of cumulative meta-analysis.

Table 1
Independent Hypothesis Tests Included in the Meta-Analytic Database

Wave	Study	Bias hypothesis tests				Agreement hypothesis tests				Gender
		Statistic	DOE	n	Z _{Fisher}	Statistic, ICC	DOE	n	Z _{Fisher}	
1	Stacy et al. (1985)	$r_{(192)} = 0.068$	+	193	0.068	0.457	+	193	0.494	B
2	Marlatt et al. (1998)	$r_{(48)} = 0.247$	−	49	−0.252	0.505	+	49	0.555	M
2	Marlatt et al. (1998)	$r_{(49)} = 0.311$	−	50	−0.321	0.567	+	50	0.643	F
3	Curtin et al. (2001)	$r_{(50)} = 0.296$	−	51	−0.305	0.332	+	28	0.345	F
4	Borsari and Carey (2005)	$r_{(15)} = 0.308$	+	16	0.319	0.415	+	16	0.441	M
4	Borsari and Carey (2005)	$r_{(1)} = 0.384$	+	2	0.405	0.076	+	2	0.076	F
4	Laforge et al. (2005)	$r_{(87)} = 0.051$	+	88	0.051	0.542	+	88	0.607	M
4	Laforge et al. (2005)	$r_{(121)} = 0.021$	+	122	0.021	0.634	+	122	0.749	F
5	Carey et al. (2006)	$r_{(70)} = 0.141$	+	71	0.142	0.474	+	71	0.515	M
5	Carey et al. (2006)	$r_{(110)} = 0.034$	+	111	0.034	0.491	+	111	0.537	F
6	Barnett et al. (2007)	$r_{(38)} = 0.181$	+	39	0.183	0.561	+	31	0.635	M
6	Barnett et al. (2007)	$r_{(38)} = 0.079$	+	39	0.080	0.382	+	39	0.402	F
6	Hagman et al. (2007)	$r_{(21)} = 0.122$	−	22	−0.123	0.445	+	29	0.479	M
6	Hagman et al. (2007)	$r_{(116)} = 0.028$	+	117	0.028	0.599	+	117	0.691	F

M, male; F, female; B, both; DOE, direction of effect.

Table 2
Bias Hypothesis Tests Included in the Expanded Meta-Analytic Database

Study	Statistic	DOE	n	Z _{Fisher}	Contact ^a	Relationship ^b	Confidence ^c	Drinking specificity ^d	Drinking behavior	Mandated	Gender
Stacy et al. (1985)	$t_{(192)} = 0.000$	+	193	0.000	68	61	.	13	F	0	B
Stacy et al. (1985)	$t_{(192)} = 1.216$	+	193	0.088	68	61	.	12	Q	0	B
Stacy et al. (1985)	$t_{(193)} = 0.681$	+	194	0.049	68	61	.	20	Q	0	B
Stacy et al. (1985)	$t_{(193)} = 0.960$	+	194	0.069	68	61	.	15	F	0	B
Stacy et al. (1985)	$t_{(192)} = 0.684$	+	194	0.049	68	61	.	19	F	0	B
Stacy et al. (1985)	$t_{(193)} = 2.102$	+	194	0.151	68	61	.	17	Q	0	B
Marlatt et al. (1998)	$t_{(48)} = 1.472$	-	49	-0.211	.	.	.	58.5	F	0	M
Marlatt et al. (1998)	$t_{(49)} = 2.030$	-	50	-0.286	.	.	.	58.5	F	0	F
Marlatt et al. (1998)	$t_{(75)} = 2.519$	-	76	-0.287	75	.	.	58.5	F	0	M
Marlatt et al. (1998)	$t_{(95)} = 4.823$	-	96	-0.477	75	.	.	58.5	F	0	F
Marlatt et al. (1998)	$t_{(48)} = 0.000$	+	49	0.000	75	.	.	51.2	Q	0	M
Marlatt et al. (1998)	$t_{(50)} = 1.184$	-	51	-0.167	75	.	.	51.2	Q	0	F
Marlatt et al. (1998)	$t_{(73)} = 3.606$	-	74	-0.410	75	.	.	51.2	Q	0	M
Marlatt et al. (1998)	$t_{(94)} = 2.611$	-	95	-0.266	75	.	.	51.2	Q	0	F
Curtin et al. (2001)	$t_{(50)} = 3.179$	-	51	-0.436	94	40	50	58	Q	0	F
Curtin et al. (2001)	$t_{(50)} = 2.463$	-	51	-0.342	94	35	50	51.2	Q	0	F
Curtin et al. (2001)	$t_{(50)} = 1.206$	-	51	-0.170	94	35	50	13	F	0	F
Curtin et al. (2001)	$t_{(55)} = 1.891$	-	56	-0.252	92	40	50	58	Q	0	F
Curtin et al. (2001)	$t_{(56)} = 2.301$	-	57	-0.303	92	40	50	51.2	Q	0	F
Curtin et al. (2001)	$t_{(58)} = 2.570$	-	59	-0.331	92	40	50	13	F	0	F
Borsari and Carey (2005)	$t_{(27)} = 2.517$	+	28	0.467	85	98	56	61	H	1	M
Borsari and Carey (2005)	$t_{(2)} = 4.715$	+	3	1.919	100	98	67	61	H	1	F
Borsari and Carey (2005)	$t_{(18)} = 1.444$	+	19	0.334	95	96	63	61	H	1	M
Borsari and Carey (2005)	$t_{(7)} = 0.487$	+	8	0.183	88	96	63	61	H	1	F
Borsari and Carey (2005)	$t_{(27)} = 3.343$	+	28	0.606	85	98	56	41	H	1	M

Study	Statistic	DOE	<i>n</i>	<i>Z</i> _{Fisher}	Contact ^a	Relationship ^b	Confidence ^c	Drinking specificity ^d	Drinking behavior	Mandated	Gender
Borsari and Carey (2005)	$t_{(2)} = 2.646$	+	3	1.384	100	98	67	41	H	1	F
Borsari and Carey (2005)	$t_{(18)} = 2.931$	+	19	0.645	95	96	63	41	F	1	M
Borsari and Carey (2005)	$t_{(7)} = 1.468$	+	8	0.530	88	96	63	41	F	1	F
Borsari and Carey (2005)	$t_{(27)} = 3.591$	+	28	0.645	85	98	56	63.3	F	1	M
Borsari and Carey (2005)	$t_{(1)} = 1.127$	+	2	0.730	100	98	67	63.3	F	1	F
Borsari and Carey (2005)	$t_{(18)} = 3.004$	+	19	0.659	95	96	63	63.3	Q	1	M
Borsari and Carey (2005)	$t_{(7)} = 1.661$	+	8	0.593	88	96	63	63.3	Q	1	F
Borsari and Carey (2005)	$t_{(27)} = 2.094$	+	28	0.393	85	98	56	58	Q	1	M
Borsari and Carey (2005)	$t_{(2)} = 1.352$	+	3	0.850	100	98	67	58	Q	1	F
Borsari and Carey (2005)	$t_{(18)} = .904$	+	19	0.212	95	96	63	58	F	1	M
Borsari and Carey (2005)	$t_{(7)} = 0.587$	−	8	−0.220	88	96	63	58	F	1	F
Borsari and Carey (2005)	$t_{(27)} = 1.890$	−	28	−0.356	85	98	56	58.5	F	1	M
Borsari and Carey (2005)	$t_{(2)} = 0.000$	+	3	0.000	100	98	67	58.5	F	1	F
Borsari and Carey (2005)	$t_{(18)} = 1.930$	+	19	0.441	95	96	63	58.5	H	1	M
Borsari and Carey (2005)	$t_{(7)} = 1.158$	+	8	0.425	88	96	63	58.5	H	1	F
Borsari and Carey (2005)	$t_{(21)} = 1.729$	+	22	0.361	86	98	46	61	H	1	M
Borsari and Carey (2005)	$t_{(1)} = 1.398$	+	2	1.137	100	98	100	61	H	1	F
Borsari and Carey (2005)	$t_{(15)} = 0.737$	+	16	0.189	93	96	47	61	H	1	M
Borsari and Carey (2005)	$t_{(6)} = 0.169$	−	7	−0.069	85	96	71	61	H	1	F

Study	Statistic	DOE	<i>n</i>	<i>Z</i> _{Fisher}	Contact ^a	Relationship ^b	Confidence ^c	Drinking specificity ^d	Drinking behavior	Mandated	Gender
Borsari and Carey (2005)	$t_{(22)} = 2.750$	+	23	0.557	86	98	46	41	H	1	M
Borsari and Carey (2005)	$t_{(1)} = 7.000$	+	2	2.644	100	98	100	41	H	1	F
Borsari and Carey (2005)	$t_{(15)} = 1.741$	+	16	0.436	93	96	47	41	F	1	M
Borsari and Carey (2005)	$t_{(6)} = 1.000$	+	7	0.398	85	96	71	41	F	1	F
Borsari and Carey (2005)	$t_{(22)} = 1.429$	+	23	0.300	86	98	46	63.3	F	1	M
Borsari and Carey (2005)	$t_{(1)} = 2.000$	+	2	1.444	100	98	100	63.3	F	1	F
Borsari and Carey (2005)	$t_{(15)} = 2.855$	+	16	0.683	93	96	47	63.3	Q	1	M
Borsari and Carey (2005)	$t_{(6)} = 0.179$	+	7	0.073	85	96	71	63.3	Q	1	F
Borsari and Carey (2005)	$t_{(22)} = 1.383$	+	23	0.291	86	98	46	58	Q	1	M
Borsari and Carey (2005)	$t_{(1)} = 0.700$	+	2	0.653	100	98	100	58	Q	1	F
Borsari and Carey (2005)	$t_{(15)} = 0.502$	−	16	−0.129	93	96	47	58	F	1	M
Borsari and Carey (2005)	$t_{(6)} = 0.385$	+	7	0.157	85	96	71	58	F	1	F
Borsari and Carey (2005)	$t_{(22)} = 2.170$	−	23	−0.447	86	98	46	58.5	F	1	M
Borsari and Carey (2005)	$t_{(1)} = 0.000$	+	2	0.000	100	98	100	58.5	F	1	F
Borsari and Carey (2005)	$t_{(15)} = 0.287$	+	16	0.074	93	96	47	58.5	Q	1	M
Borsari and Carey (2005)	$t_{(6)} = 0.548$	−	7	−0.222	85	96	71	58.5	Q	1	F
Laforge et al. (2005)	$t_{(89)} = 0.201$	+	90	0.021	81	93	90	38.2	F	0	M
Laforge et al. (2005)	$t_{(89)} = 0.118$	+	90	0.012	81	93	90	51.2	Q	0	M
Laforge et al. (2005)	$t_{(89)} = 1.037$	+	90	0.110	81	93	90	41	H	0	M
Laforge et al. (2005)	$t_{(87)} = 0.580$	+	88	0.062	81	93	90	61	H	0	M
Laforge et al. (2005)	$t_{(127)} = 0.101$	−	128	−0.009	84	95	92	38.2	F	0	F

Study	Statistic	DOE	n	Z _{Fisher}	Contact ^a	Relationship ^b	Confidence ^c	Drinking specificity ^d	Drinking behavior	Mandated	Gender
Laforge et al. (2005)	$t_{(128)} = 1.634$	–	129	–0.144	84	95	92	51.2	Q	0	F
Laforge et al. (2005)	$t_{(125)} = 1.784$	+	126	0.159	84	95	92	41	H	0	F
Laforge et al. (2005)	$t_{(121)} = 0.943$	+	122	0.086	84	95	92	61	H	0	F
Carey et al. (2006)	$t_{(105)} = 1.945$	+	106	0.189	.	.	44	61	H	0	M
Carey et al. (2006)	$t_{(199)} = 0.840$	+	200	0.060	.	.	44	61	H	0	F
Carey et al. (2006)	$t_{(105)} = 2.539$	+	106	0.245	.	.	44	41	H	0	M
Carey et al. (2006)	$t_{(199)} = 1.870$	+	200	0.132	.	.	44	41	H	0	F
Carey et al. (2006)	$t_{(107)} = 0.409$	+	108	0.040	.	.	44	58	Q	0	M
Carey et al. (2006)	$t_{(201)} = 0.502$	–	202	–0.035	.	.	44	58	Q	0	F
Carey et al. (2006)	$t_{(70)} = 0.626$	+	71	0.075	.	.	44	63.3	F	0	M
Carey et al. (2006)	$t_{(110)} = 0.626$	–	111	–0.060	.	.	44	63.3	F	0	F
Barnett et al. (2007)	$t_{(38)} = 1.583$	+	39	0.254	87	82	86	51.2	Q	1	M
Barnett et al. (2007)	$t_{(38)} = 1.098$	+	39	0.177	87	82	86	58.5	F	1	M
Barnett et al. (2007)	$t_{(38)} = 0.598$	+	39	0.097	87	82	86	61	H	1	M
Barnett et al. (2007)	$t_{(41)} = 2.943$	+	42	0.445	90	86	93	51.2	Q	1	M
Barnett et al. (2007)	$t_{(41)} = 0.731$	+	42	0.114	90	83	93	58.5	F	1	M
Barnett et al. (2007)	$t_{(41)} = 0.055$	+	42	0.009	90	86	93	61	H	1	M
Barnett et al. (2007)	$t_{(50)} = 3.145$	+	51	0.431	90	93	98	51.2	Q	1	F
Barnett et al. (2007)	$t_{(50)} = 0.063$	+	51	0.009	92	93	98	58.5	F	1	F
Barnett et al. (2007)	$t_{(50)} = 0.778$	+	51	0.110	92	93	98	61	H	1	F
Barnett et al. (2007)	$t_{(62)} = 0.890$	+	63	0.113	79	83	95	51.2	Q	1	F
Barnett et al. (2007)	$t_{(62)} = 0.070$	–	63	–0.009	79	83	95	58.5	F	1	F
Barnett et al. (2007)	$t_{(62)} = 0.931$	–	63	–0.118	79	86	95	61	H	1	F
Hagman et al. (2007)	$t_{(116)} = 0.024$	–	117	–0.002	.	.	.	58.5	F	0	F
Hagman et al. (2007)	$t_{(126)} = 0.832$	+	127	0.024	.	.	.	18.7	F	0	F
Hagman et al. (2007)	$t_{(127)} = 1.252$	+	128	0.111	.	.	.	17	Q	0	F

Study	Statistic	DOE	n	Z _{Fisher}	Contact ^a	Relationship ^b	Confidence ^c	Drinking specificity ^d	Drinking behavior	Mandated	Gender
Hagman et al. (2007)	$t_{(120)} = 1.564$	+	121	0.142	.	.	.	13	F	0	F
Hagman et al. (2007)	$t_{(123)} = 0.403$	-	124	-0.036	.	.	.	12.7	Q	0	F
Hagman et al. (2007)	$t_{(126)} = 0.172$	+	127	0.015	.	.	.	14.5	F	0	F
Hagman et al. (2007)	$t_{(125)} = 0.616$	-	126	-0.055	.	.	.	20	Q	0	F
Hagman et al. (2007)	$t_{(21)} = 0.902$	-	22	-0.074	.	.	.	58.5	F	0	M
Hagman et al. (2007)	$t_{(29)} = 1.934$	-	30	-0.352	.	.	.	18.7	F	0	M
Hagman et al. (2007)	$t_{(29)} = 0.915$	-	30	-0.169	.	.	.	17	Q	0	M
Hagman et al. (2007)	$t_{(29)} = 0.769$	+	30	0.142	.	.	.	13	F	0	M
Hagman et al. (2007)	$t_{(29)} = 0.171$	-	30	-0.032	.	.	.	12.7	Q	0	M
Hagman et al. (2007)	$t_{(28)} = 0.587$	-	30	-0.111	.	.	.	14.5	F	0	M
Hagman et al. (2007)	$t_{(28)} = 1.364$	-	29	-0.255	.	.	.	20	Q	0	M

n; pairs of collaterals and participants; Q, quantity; F, frequency; H, heavy drinking; M, males; F, females; B, both; DOE, direction of effect 0 = nonmandated students, 1 = mandated students.

^aPercent spent with student “every day” or “nearly every day”

^bpercent of students who reported being a “close” or “best” friend

^cpercent of students who reported that they were very or mostly confident in their self-report;

^drating of drinking behavior assessed (0 = concise, 100 = vague).

Pluses and minuses in brackets indicate direction of effect entry.

Table 3
Agreement Hypothesis Tests Included in the Expanded Meta-Analytic Database

Study	ICC(1)	DOE	n	Z _{Fisher}	Contact ^a	Relationship ^b	Confidence ^c	Drinking specificity ^d	Drinking behavior	Mandated	Gender
Stacy et al. (1985)	0.532	+	193	0.593	68	61	—	13	F	0	B
Stacy et al. (1985)	0.440	+	193	0.472	68	61	.	12	Q	0	B
Stacy et al. (1985)	0.411	+	194	0.437	68	61	.	20	Q	0	B
Stacy et al. (1985)	0.296	+	194	0.305	68	61	.	15	F	0	B
Stacy et al. (1985)	0.609	+	194	0.707	68	61	.	19	F	0	B
Stacy et al. (1985)	0.420	+	194	0.448	68	61	.	17	Q	0	B
Marlatt et al. (1998)	0.529	+	49	0.589	.	.	.	58.5	F	0	M
Marlatt et al. (1998)	0.501	+	50	0.551	.	.	.	58.5	F	0	F
Marlatt et al. (1998)	0.437	+	76	0.469	75	.	.	58.5	F	0	M
Marlatt et al. (1998)	0.566	+	96	0.642	75	.	.	58.5	F	0	F
Marlatt et al. (1998)	0.566	+	49	0.642	75	.	.	51.2	Q	0	M
Marlatt et al. (1998)	0.689	+	51	0.846	75	.	.	51.2	Q	0	F
Marlatt et al. (1998)	0.513	+	74	0.567	75	.	.	51.2	Q	0	M

Study	ICC(1)	DOE	n	Z _{Fisher}	Contact ^a	Relationship ^b	Confidence ^c	Drinking specificity ^d	Drinking behavior	Mandated	Gender
Marlatt et al. (1998)	0.527	+	95	0.586	75	.	.	51.2	Q	0	F
Curtin et al. (2001)	0.387	+	51	0.408	94	40	50	58	Q	0	F
Curtin et al. (2001)	0.330	+	51	0.343	94	35	50	51.2	Q	0	F
Curtin et al. (2001)	0.471	+	51	0.511	94	35	50	13	F	0	F
Curtin et al. (2001)	0.175	+	56	0.177	92	40	50	58	Q	0	F
Curtin et al. (2001)	0.316	+	57	0.327	92	40	50	51.2	Q	0	F
Curtin et al. (2001)	0.313	+	59	0.324	92	40	50	13	F	0	F
Borsari and Carey (2005)	0.202	+	28	0.205	85	98	56	61	H	1	M
Borsari and Carey (2005)	0.250	–	3	–0.255	100	98	67	61	H	1	F
Borsari and Carey (2005)	0.716	+	19	0.899	95	96	63	61	H	1	M
Borsari and Carey (2005)	0.237	–	8	–0.242	88	96	63	61	H	1	F
Borsari and Carey (2005)	0.147	+	28	0.148	85	98	56	41	H	1	M
Borsari and Carey (2005)	0.795	–	3	–1.085	100	98	67	41	H	1	F

Study	ICC(1)	DOE	<i>n</i>	Z_{Fisher}	Contact ^a	Relationship ^b	Confidence ^c	Drinking specificity ^d	Drinking behavior	Mandated	Gender
Borsari and Carey (2005)	0.331	+	19	0.344	95	96	63	41	F	1	M
Borsari and Carey (2005)	0.065	+	8	0.065	88	96	63	41	F	1	F
Borsari and Carey (2005)	0.182	+	28	0.184	85	98	56	63.3	F	1	M
Borsari and Carey (2005)	0.056	+	3	0.056	100	98	67	63.3	F	1	F
Borsari and Carey (2005)	0.053	+	19	0.053	95	96	63	63.3	Q	1	M
Borsari and Carey (2005)	0.540	+	8	0.604	88	96	63	63.3	Q	1	F
Borsari and Carey (2005)	0.190	+	28	0.192	85	98	56	58	Q	1	M
Borsari and Carey (2005)	0.402	−	3	−0.426	100	98	67	58	Q	1	F
Borsari and Carey (2005)	0.675	+	19	0.820	95	96	63	58	F	1	M
Borsari and Carey (2005)	0.236	+	8	0.241	88	96	63	58	F	1	F
Borsari and Carey (2005)	0.639	+	28	0.756	85	98	56	58.5	F	1	M
Borsari and	0.000	+	3	0.000	100	98	67	58.5	F	1	F

Study	ICC(1)	DOE	<i>n</i>	Z_{Fisher}	Contact ^a	Relationship ^b	Confidence ^c	Drinking specificity ^d	Drinking behavior	Mandated	Gender
Carey (2005)											
Borsari and Carey (2005)	0.885	–	19	–1.398	95	96	63	58.5	H	1	M
Borsari and Carey (2005)	0.955	–	8	–1.886	88	96	63	58.5	H	1	F
Borsari and Carey (2005)	0.492	+	23	0.539	86	98	46	61	H	1	M
Borsari and Carey (2005)	0.718	+	2	0.904	100	98	100	61	H	1	F
Borsari and Carey (2005)	0.875	+	16	1.354	93	96	47	61	H	1	M
Borsari and Carey (2005)	0.644	+	7	0.765	85	96	71	61	H	1	F
Borsari and Carey (2005)	0.599	+	23	0.692	86	98	46	41	H	1	M
Borsari and Carey (2005)	0.000	+	2	0.000	100	98	100	41	H	1	F
Borsari and Carey (2005)	0.807	+	16	1.118	93	96	47	41	F	1	M
Borsari and Carey (2005)	0.145	–	7	–0.146	85	96	71	41	F	1	F
Borsari and Carey (2005)	0.289	+	23	0.297	86	98	46	63.3	F	1	M

Study	ICC(1)	DOE	n	Z _{Fisher}	Contact ^a	Relationship ^b	Confidence ^c	Drinking specificity ^d	Drinking behavior	Mandated	Gender
Borsari and Carey (2005)	0.565	+	2	0.640	100	98	100	63.3	F	1	F
Borsari and Carey (2005)	0.079	+	16	0.079	93	96	47	63.3	Q	1	M
Borsari and Carey (2005)	0.829	+	7	1.185	85	96	71	63.3	Q	1	F
Borsari and Carey (2005)	0.578	+	23	0.659	86	98	46	58	Q	1	M
Borsari and Carey (2005)	0.318	+	2	0.329	100	98	100	58	Q	1	F
Borsari and Carey (2005)	0.854	+	16	1.271	93	96	47	58	F	1	M
Borsari and Carey (2005)	0.360	+	7	0.377	85	96	71	58	F	1	F
Borsari and Carey (2005)	0.547	+	23	0.614	86	98	46	58.5	F	1	M
Borsari and Carey (2005)	1.000	+	2	4.952	100	98	100	58.5	F	1	F
Borsari and Carey (2005)	0.560	+	16	0.633	93	96	47	58.5	Q	1	M
Borsari and Carey (2005)	0.486	+	7	0.531	85	96	71	58.5	Q	1	F

Study	ICC(1)	DOE	n	Z _{Fisher}	Contact ^a	Relationship ^b	Confidence ^c	Drinking specificity ^d	Drinking behavior	Mandated	Gender
Laforge et al. (2005)	0.644	+	90	0.765	81	93	90	38.2	F	0	M
Laforge et al. (2005)	0.570	+	90	0.648	81	93	90	51.2	Q	0	M
Laforge et al. (2005)	0.594	+	90	0.684	81	93	90	41	H	0	M
Laforge et al. (2005)	0.316	+	88	0.327	81	93	90	61	H	0	M
Laforge et al. (2005)	0.59	+	128	0.678	84	95	92	38.2	F	0	F
Laforge et al. (2005)	0.656	+	129	0.786	84	95	92	51.2	Q	0	F
Laforge et al. (2005)	0.622	+	126	0.728	84	95	92	41	H	0	F
Laforge et al. (2005)	0.667	+	122	0.805	84	95	92	61	H	0	F
Carey et al. (2006)	0.575	+	106	0.655	.	.	44	61	H	0	M
Carey et al. (2006)	0.56	+	201	0.633	.	.	44	61	H	0	F
Carey et al. (2006)	0.396	+	108	0.419	.	.	44	41	H	0	M
Carey et al. (2006)	0.484	+	198	0.528	.	.	44	41	H	0	F
Carey et al. (2006)	0.593	+	109	0.682	.	.	44	58	Q	0	M
Carey et al. (2006)	0.463	+	200	0.501	.	.	44	58	Q	0	F

Study	ICC(1)	DOE	n	Z _{Fisher}	Contact ^a	Relationship ^b	Confidence ^c	Drinking specificity ^d	Drinking behavior	Mandated	Gender
Carey et al. (2006)	0.191	+	71	0.193	.	.	44	63.3	F	0	M
Carey et al. (2006)	0.417	+	111	0.444	.	.	44	63.3	F	0	F
Barnett et al. (2007)	0.477	+	31	0.519	87	82	86	51.2	Q	1	M
Barnett et al. (2007)	0.355	+	31	0.371	87	82	86	58.5	F	1	M
Barnett et al. (2007)	0.427	+	31	0.456	87	82	86	61	H	1	M
Barnett et al. (2007)	0.775	+	38	1.033	90	86	93	51.2	Q	1	M
Barnett et al. (2007)	0.476	+	38	0.518	90	83	93	58.5	F	1	M
Barnett et al. (2007)	0.672	+	36	0.814	90	86	93	61	H	1	M
Barnett et al. (2007)	0.578	+	41	0.659	90	93	98	51.2	Q	1	F
Barnett et al. (2007)	0.308	+	39	0.318	92	93	98	58.5	F	1	F
Barnett et al. (2007)	0.315	+	41	0.326	92	93	98	61	H	1	F
Barnett et al. (2007)	0.388	+	56	0.409	79	83	95	51.2	Q	1	F
Barnett et al. (2007)	0.293	+	52	0.302	79	83	95	58.5	F	1	F
Barnett et al. (2007)	0.392	+	55	0.414	79	86	95	61	H	1	F

Study	ICC(1)	DOE	n	Z _{Fisher}	Contact ^a	Relationship ^b	Confidence ^c	Drinking specificity ^d	Drinking behavior	Mandated	Gender
Hagman et al. (2007)	0.591	+	117	0.679	.	.	.	58.5	F	0	F
Hagman et al. (2007)	0.629	+	127	0.740	.	.	.	18.7	F	0	F
Hagman et al. (2007)	0.552	+	128	0.621	.	.	.	17	Q	0	F
Hagman et al. (2007)	0.559	+	121	0.631	.	.	.	13	F	0	F
Hagman et al. (2007)	0.498	+	124	0.547	.	.	.	12.7	Q	0	F
Hagman et al. (2007)	0.723	+	127	0.914	.	.	.	14.5	F	0	F
Hagman et al. (2007)	0.605	+	126	0.701	.	.	.	20	Q	0	F
Hagman et al. (2007)	0.198	+	30	0.201	.	.	.	58.5	F	0	M
Hagman et al. (2007)	0.466	+	30	0.505	.	.	.	18.7	F	0	M
Hagman et al. (2007)	0.472	+	30	0.513	.	.	.	17	Q	0	M
Hagman et al. (2007)	0.689	+	30	0.846	.	.	.	13	F	0	M
Hagman et al. (2007)	0.341	+	30	0.355	.	.	.	12.7	Q	0	M
Hagman et al. (2007)	0.384	+	29	0.405	.	.	.	14.5	F	0	M
Hagman et al. (2007)	0.484	+	29	0.528	.	.	.	20	Q	0	M

n, pairs of collaterals and participants; *Q*, quantity, *F*, frequency; *H*, heavy drinking; *M*, males; *F*, females; *B*, both; *DOE*, direction of effect. 0 = nonmandated students, 1 = mandated students.

^aPercent spent with student “every day” or “nearly every day”

^bpercent of students who reported being a “close” or “best” friend;

^cpercent of students who reported that they were very or mostly confident in their self-report;

^drating of drinking behavior assessed (0 = concise, 100 = vague)

Plus and minuses in brackets indicate direction of effect entry.