Virginia Commonwealth University
VCU Scholars Compass

Theses and Dissertations | Graduate School

2016

Identifying misconceptions associated with inaccurate survey reporting in the combined use of caffeine and alcohol

Kathryn Polak

Follow this and additional works at: https://scholarscompass.vcu.edu/etd

 Part of the Clinical Psychology Commons

© The Author

Downloaded from
https://scholarscompass.vcu.edu/etd/4287

This Thesis is brought to you for free and open access by the Graduate School at VCU Scholars Compass. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of VCU Scholars Compass. For more information, please contact libcompass@vcu.edu.

Identifying misconceptions associated with inaccurate survey reporting in the combined use of caffeine and alcohol

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University.

by

Kathryn M. Polak
B.S., Mary Baldwin College, May 2010

Director: Dace S. Svikis, Ph.D.
Professor, Department of Psychology

Virginia Commonwealth University
Richmond, Virginia
May, 2016

Acknowledgements

I would like to thank a number of people for their help on this thesis. First, I would like to acknowledge and thank my advisor, mentor, and committee chair, Dr. Dace Svikis. She has been a constant source of support and guidance throughout both this project and my time in graduate school thus far. I would also like to thank my other committee members, Drs. Heather Jones and Linda Hancock, for their advice during the completion of this thesis. In addition, I would like to thank Dr. Leroy Thacker for providing statistical consultation on this project. Finally, I wish to thank my husband, Matt, for being a continued source of love and support.

Table of Contents

Funding: This grant was supported by VCU funds available to Dr. Svikis.

List of Tables

List of Figures

Abstract

IDENTIFYING MISCONCEPTIONS ASSOCIATED WITH INACCURATE SURVEY REPORTING IN THE COMBINED USE OF CAFFEINE AND ALCOHOL

By Kathryn Polak, B.S.

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University.

Virginia Commonwealth University, 2016

Director: Dace S. Svikis, Ph.D., Professor, Department of Psychology

Research on college student use of caffeine combined with alcohol (CAC) and public health concern over such use has been hampered by the absence of psychometrically sound measures of caffeine and CAC use. The present study examined agreement between survey (CAS) and interview (TLFB) methods for collecting data on caffeine, alcohol and CAC use. Participants were *N*=50 college students randomized to complete CAS followed by TLFB or the reverse. Qualitative follow-up interviews with *N*=15 participants were used to identify factors contributing to CAS-TLFB discrepancies. Responses varied by method of administration, with largest discrepancy magnitudes found for CAC, followed by caffeine, then alcohol use. Rates of reporting use by only one method were highest for CAC (65.5%). Lack of knowledge about caffeine was common, with over half (56%) having at least one caffeine misreport. Largest discrepancies were found for CAC use, an area of public health concern, particularly among college students.

Identifying misconceptions associated with inaccurate survey reporting in the combined use of caffeine and alcohol

Alcohol consumption and binge drinking are prevalent among college students (e.g., SAMHSA, 2013b). Caffeine in the form of coffee, sodas, and tea is also widely used, and the advent of energy drinks has brought significant problems (e.g., FDA, 2012). In particular, college students often consume caffeine combined with alcohol (CAC) and such use has been associated with greater problems than alcohol use alone, at comparable quantities of alcohol (e.g., Arria et al., 2010). More research is needed to better understand CAC use and factors that contribute to risk for abuse and negative consequences. This can inform subsequent prevention and intervention efforts.

Central to this field of research is investigator ability to accurately measure quantity and frequency of caffeine, alcohol, and CAC use. While standardized measures exist for alcohol, much less is known for caffeine and even less for CAC. Further, preliminary data from the Svikis lab found many college students were misinformed about what beverages contain caffeine, which contributed to less accurate reporting of caffeine and CAC use behavior (Polak, Hancock, & Svikis, n.d.).

This study used a mixed methods approach to collect benchmark data on novel methods for collecting quantity and frequency information about caffeine, alcohol and CAC use. The study was conducted in 2 phases. In Phase 1, TLFB interview and computer-administered survey methods of data collection were compared in counterbalanced order using a sample of college students. In Phase 2, researchers interviewed participants who previously reported CAC use

about any discrepancies in their Phase 1 reporting of caffeine, alcohol and CAC use to obtain clarification and better understanding of contributing factors.

The study findings will inform and guide future development and testing of screening and assessment measures that can aid clinicians and researchers alike in the identification of college students at risk for development of caffeine and CAC problems. In the current climate of diminishing health care funding, these are the individuals that are most in need of education and intervention to prevent the development of such problems.

Statement of the Problem and Aims

Problems and Clinical Relevance

Alcohol consumption is widespread on college campuses. About 80 percent of college students report alcohol use and half of these alcohol users are binge drinkers ($\geq$4 drinks for women or $\geq$5 for men per occasion) (NIAAA, n.d.a). Such use has negative consequences; alcohol consumption has been associated with physical and sexual assault, depression, suicide attempts, and even death (NIAAA, n.d.a). These findings have prompted large scale prevention and early intervention efforts on college campuses.

In contrast to this focus on alcohol use, caffeine use has historically not received much research attention despite 95 percent of college students self-reporting daily caffeine consumption (McIlvain et al., 2011). This began to change with the advent of caffeinated energy drinks. Introduced in 1997, and heavily marketed to college students, energy drink consumption has contributed to adverse health effects, prompting increased attention and cause for concern.

In particular, the combining of alcohol with caffeinated energy drinks has garnered substantive media attention. While the FDA stopped selling prepackaged caffeinated alcohol products, mixing of the two remains prevalent (CDC, 2014). Additionally, research has shown

that CAC users experience greater problems compared to those who use alcohol by itself. CAC drinkers are more likely to report alcohol dependence and drug use compared to drinkers of alcohol alone (Arria et al., 2010; Arria et al., 2011) and they are more likely to drive while impaired and be hurt or injured (Brache et al., 2012). They are also twice as likely to report being taken advantage of sexually, taking advantage of someone else sexually, and riding with a driver who was under the influence of alcohol (O'Brien et al., 2008).

More research is needed to better understand caffeine and CAC use patterns and the factors that contribute to risk for abuse and negative consequences. Central to this field of study is the researcher/clinician ability to accurately measure quantity and frequency of caffeinated beverage consumption and in particular, CAC consumption.

While standardized measures are available to assess alcohol use, comparable measures do not exist for caffeine and CAC use. Additionally, recent data from our lab suggests student knowledge about caffeinated products may be limited. In a mixed methods survey of 111 college students, we found 51 percent had reporting errors and 15 percent failed to distinguish between caffeine containing and non-caffeinated beverages (Polak, Hancock, & Svikis, n.d.).

Aims

The purpose of the study was to: 1) compare interview and survey methods for measurement of caffeine, alcohol and CAC use in a sample of college students reporting recent regular use of both substances; and 2) where lack of agreement was found, conduct qualitative interviews to identify themes and patterns with respect to caffeine and CAC misconceptions and reporting errors.

This thesis begins with a review of relevant literature on caffeine and alcohol use and associated problems. Alternative methods for measuring quantity and frequency of such use are

summarized. In particular, advantages and disadvantages of survey as compared to interview methods for data collection are discussed. Whenever possible, the review focuses on research with college students and young adult populations. The review also includes a summary of studies of CAC use, with particular attention paid to methods used for data collection.

Next, study methods are described, summarizing the 2-phase procedures used for this project. In Phase 1, survey and interview methods for collection of caffeine, alcohol and CAC use in a sample of college students were compared. In addition, potential correlates of CAC use were examined. This was followed, in Phase 2, by a qualitative interview with those students reporting recent (past 30 days) CAC use. The purpose of this component was to determine common themes and factors that may have contributed to misperceptions and errors in self-reports of caffeine, alcohol and CAC use. This study had the following aims:

Aim 1. Assess descriptive and associative features of quantity and frequency of alcohol, caffeine, and CAC use.

Aim 2. Examine the agreement between survey and TLFB methods, identifying areas where there is lack of agreement.

Aim 3. Examine qualitative personal interview data and identify themes and patterns with respect to discrepancies and misconceptions.

Review of the Literature

Caffeine

Introduction. Caffeine is a mild stimulant that increases dopamine to produce energizing effects (Striley & Khan, 2014) such as heightened alertness and increased energy (NIC, 2015). These effects begin as early as 15 minutes after consumption and can last up to six hours (NIC, 2015). Side effects associated with caffeine use include rapid heartbeat, anxiety, difficulty sleeping, and tremors (NIH, 2015). Caffeine can be found naturally in coffee, tea, and chocolate. Categorized as a food additive, it is also present in various soft drinks (e.g., Coca Cola, Mountain Dew), foods (e.g., Wired Waffles, Energy Gummy Bears) and certain medicines (e.g., Excedrin, Anacid) (NIH, 2015).

Prevalence. Caffeine is the most commonly consumed psychoactive substance in the world with 80 percent of individuals reporting daily caffeine use (James, 1997). In the US, 80 to 90 percent of adults report regular use of caffeine (Fulgoni, Keast, & Lieberman, 2015). Average consumption of caffeine is about 200 mg per day, which is equivalent to five soft drinks or two cups of coffee (Frary et al., 2005). A recent study found that about 95 percent of college students have consumed caffeine in the past two weeks (McIlvain et al., 2011). Thirty-five to 50 percent of young adults report regular use of caffeine (Fulgoni, Keast, & Lieberman, 2015).

Heavy Caffeine Use. Low to moderate levels of caffeine use are generally considered safe and may even have beneficial effects (e.g., increased energy and alertness) (NIC, 2015). For healthy individuals, the American Medical Association Council on Scientific Affairs recommends consuming non-harmful/moderate amounts of caffeine, limited to approximately to 200 to 300 mg (equivalent to two to three 8 oz. cups of brewed or dip coffee, five servings of caffeinated soft drinks, or five servings of tea) (NIH, 2015).

However, heavy use can lead to problems. Heavy caffeine consumption is defined as 500 to 600 mg per day (equal to 5 to 6 cups of coffee or 13 to 15 soft drinks) (Mayo Clinic, 2014), and is associated with such adverse consequences as headaches, insomnia, tachycardia, and muscle tremors (e.g., Reissig et al., 2009; Clauson et al., 2008). The American Medical Association has designated 800 mg of caffeine (i.e., ten 8 oz cups of coffee) to be excessive (NIH, 2015).

Quantifying Caffeine Intake. Assessing the amount of caffeine consumed has proven difficult because caffeine content varies by product type (e.g., coffee vs tea); serving size (e.g., 6 oz cup vs 12 oz can); and method of preparation (e.g., brewed vs instant coffee). This difficulty is exacerbated by the fact that the variation in caffeine content has increased exponentially within and across beverage types.

Coffee, tea and soft drinks. Variation in caffeine levels for coffee, tea, and soft drinks are summarized in Tables 1, 2, and 3, respectively. Each table lists popular types for each beverage category as well as the associated serving sizes and caffeine content.

Table 1. Caffeine Levels in Popular Types of Coffee

Type of Coffee	**Serving Size**	**Caffeine (mg)**
Starbucks Coffee	venti, 20 fl. oz	415
Panera Coffee	regular, 16.8 fl. oz.	189
Dunkin' Donuts Coffee	medium, 14 fl. oz.	178
Maxwell House Ground Coffee—100% Colombian, Dark Roast, Master Blend, or Original Roast	2 Tbs., makes 12 fl. oz.	100-160
Keurig Coffee K-Cup, all varieties	1 cup, makes 8 fl. oz.	75-150
Folgers Classic Roast Instant Coffee	2 tsp., makes 12 fl. oz.	148
McDonald's Coffee	large, 16 fl. oz.	133
Dunkin' Donuts, Panera, or Starbucks Decaf Coffee	16 fl. oz.	15-25
Maxwell House Decaf Ground Coffee	2 Tbs., makes 12 fl. oz.	2-10

Table 2. Caffeine Levels in Popular Types of Tea

Type of Tea	Serving Size	Caffeine (mg)
Starbucks Tazo Awake—Brewed Tea or Tea Latte	grande, 16 fl. oz.	135
Black tea, brewed for 3 minutes	8 fl. oz.	30-80
Snapple Lemon Tea	16 fl. oz.	62
Lipton Pure Leaf Iced Tea	18.5 fl. oz.	60
Green tea, brewed for 3 minutes	8 fl. oz.	35-60
Nestea Unsweetened Iced Tea Mix	2 tsp., makes 8 fl. oz.	20-30
Arizona Iced Tea, green, all varieties	16 fl. oz.	15
Lipton Decaffeinated Tea—black or green, brewed	8 fl. oz.	5
Herbal Tea, brewed	8 fl. oz.	0

Table 3. Caffeine Levels in Popular Soft Drinks

Type of Soft Drink	Serving Size (fl. oz.)	Caffeine (mg)
Pepsi MAX	12	69
Mountain Dew, regular or diet	12	54
Diet Coke	12	47
Dr Pepper or Sunkist, regular or diet	12	41
Pepsi	12	38
Coca-Cola, Coke Zero, or Diet Pepsi	12	35
Barq's Root Beer, regular	12	23
7-Up, Fanta, Fresca, ginger ale, or Sprite	12	0
Root beer, most brands, or Barq's Diet Root Beer	12	0

* The FDA limits the amount of caffeine in soft drinks to 71 mg per 12 oz.

Caffeinated Energy Drinks. Energy drinks (EDs) usually contain high levels of caffeine combined with other alleged energy-boosting ingredients (e.g., taurine, guarana, and B vitamins) (McLellan & Lieberman, 2012). In addition to full size beverages, these drinks are consumed as shots sold in smaller containers (e.g., 5-hour Energy, Amp), containing comparable amounts of caffeine and other ingredients (Kurtx et al., 2013).

EDs were introduced to the US in 1997 through the marketing of Austrian-based, Red Bull (UVA, n.d.). Currently, over seven million EDs are sold worldwide every day, with over 500 ED products from which to choose (Reissig, Strain, & Griffiths, 2009). These beverages are particularly popular among college students and marketing efforts often target this group (Reissig, Strain, & Griffiths, 2009). EDs are promoted for enhancing alertness and improving cognitive/athletic-based performances. Typical marketing strategies are aggressively target youth and, in particular, young males (Reissig, Strain, & Griffiths, 2009).

Unlike other caffeine-containing beverages, EDs are classified as dietary supplements and are thus not subject to FDA food regulations on caffeine quantities and manufacturers are not required to report the ingredients of such products. As a result, EDs often contain high amounts of caffeine compared to other caffeinated beverages (e.g., Arria & O'Brien, 2011), ranging from 50 mg to over 500 mg per can (Reissig, Strain, & Griffiths, 2009) Additionally, they often go unlabeled and are not standardized across individual servings. As shown in Table 4, the caffeine content of many popular energy drinks and their serving sizes vary from 80 mg (Red Bull) to 280 mg (Jolt Energy Drink).

Table 4. Caffeine Levels in Popular Energy Drinks

Type of Energy Drink	Serving Size (fl. oz.)	Caffeine (mg)
Jolt Energy Drink	23.5	280
Rockstar Citrus Punched	16	240
NOS Active Sports Drink (Coca-Cola)	22	221
5-hour Energy	1.9	208
Full Throttle	16	200
Monster Energy	16	160
Rockstar	16	160
Venom Energy Drink	16	160
NOS Energy Drink (Coca-Cola)	16	160
AMP Energy Boost Original (PepsiCo)	16	142
Mountain Dew Kick Start	16	92
Red Bull	8.4	80

There has been a rapid rise in ED consumption (Heckman et al., 2010), causing concern among health providers because of the often unsafe amounts of caffeine being consumed. The number of emergency department visits related to ED use doubled from 2007 to 2011 (10,068 and 20,783, respectively) (SAMHSA, 2013a). One specific energy shot, 5-Hour Energy, has

been mentioned in about 90 FDA filings and has been involved in 13 deaths since 2009 (FDA, 2012). The FDA recently issued a warning about powdered pure caffeine, which is now being marketed; one teaspoon is equivalent to 25 cups of coffee and thus four teaspoons can be a fatal dose for an adult (FDA, 2014). Table 5 summarizes the FDA filings of energy drinks for the five year period from 2008 to 2012 (FDA, 2012).

Table 5. Energy Drink-Related FDA Filings from 2008 to 2012.

Type of Energy Drink	**Number of Adverse Events**
Red Bull	21
5-Hour Energy	92
Monster	40
Rockstar	13
TOTAL	166

*16 deaths are included in these adverse events.

Measurement of Caffeine Consumption. In studies of caffeine, researchers have generally relied upon self-report measures that often do not distinguish between types of caffeine (e.g., brewed and instant coffee) and variations in serving size (e.g., 6 oz vs 8 oz cup). Approximations are often the norm, with a cup of coffee said to have 100 mg of caffeine, tea 40 mg, and soft drinks 40 mg (CSPI, 2014; BPRU, 2003). As the number and types of caffeine sources have increased, accurate measurement has grown even more challenging. As shown in Table 1, the amount of caffeine in a serving of coffee can vary from 133 mg (McDonald's

Coffee) to 415 mg (Starbucks Coffee) with similar variations found for tea (range from 0 mg [herbal tea] to 135 mg [Starbucks Tazo Awake]; Table 2) and soft drinks (vary from 0 mg [root beer] to 69 mg [Pepsi MAX] in each 12 oz can; Table 3). In addition, decaffeinated options are available for all these substance types. However, even those can contain modest amounts of caffeine; for example, decaffeinated coffee can contain anywhere from 2 to 25 mg of caffeine (CSPI, 2014).

Alternatives to Self-Report. Levels of caffeine consumption can also be estimated through collection of saliva and blood samples that are assayed for caffeine and paraxanthine using bioanalytic methods (James et al., 1989). Research has found caffeine concentrations in saliva and blood to be highly correlated with comparable elimination of caffeine half-life (Setchell et al., 1987; Zygler-Katz et al., 1984). Further, significant correlations have been found between bioanalytic and self-reported levels of caffeine use (James et al., 1988), with saliva and blood proving to be equally reliable sources of such biological information.

Diagnostic Criteria. Heavy or problematic caffeine use can lead to caffeine-related conditions. The DSM-5 recognizes two caffeine-related disorders, Caffeine Intoxication and Caffeine Withdrawal, which are summarized in Tables 6 and 7, respectively (APA, 2013). Among individuals who report symptoms of caffeine withdrawal, 13 percent experience clinically significant distress or impairment (Juliano & Griffiths, 2004). Meredith and colleagues (2013) recently reviewed studies of caffeine-related disorders and found the number of individuals meeting criteria ranged from nine percent in a sample of US adults (Hughes et al., 1998) to 79 percent in adults seeking treatment for caffeine use (Juliano et al., 2012).

Table 6. DSM-5 Criteria for Caffeine Intoxication

A. Recent consumption of caffeine (typically a high dose well in excess of 250 mg).
B. At least five of the following signs or symptoms developing during, or shortly after, caffeine use: 1. Restlessness. 2. Nervousness. 3. Excitement. 4. Insomnia. 5. Flushed face. 6. Diuresis. 7. Gastrointestinal Disturbance 8. Muscle twitching. 9. Rambling flow of thought and speech. 10. Tachycardia or cardiac arrhythmia. 11. Periods of inexhaustibility. 12. Psychomotor agitation. C. The signs or symptoms in Criterion B cause clinically significant distress or impairment in social, occupational, or other important areas of functioning.
D. The signs or symptoms are not attributable to another medical condition and are not better explained by another mental disorder, including intoxication with another substance.

Note. Adapted from *DSM-5*, p. 503-504, APA, 2013, Arlington.

Table 7. DSM-5 Caffeine Withdrawal Diagnostic Criteria

A. Prolonged daily use of caffeine.
B. Abrupt cessation of or reduction in caffeine use, followed within 24 hours by at least three of the following signs or symptoms: 1. Headache. 2. Marked fatigue or drowsiness. 3. Dysphoric mood, depressed mood, or irritability. 4. Difficulty concentrating. 5. Flu-like symptoms (nausea, vomiting, or muscle pain/stiffness).
C. The signs or symptoms in Criterion B cause clinically significant distress or impairment in social, occupational, or other important areas of functioning.
D. The signs or symptoms are not associated with the physiological effects of another medical condition (e.g., migraine, viral illness) and are not better explained by anther mental disorder, including intoxication or withdrawal from another substance.

Note. Adapted from *DSM-5*, p. 506, APA, 2013, Arlington.

While DSM-IV did not recognize caffeine dependence as a mental disorder, many researchers sought to determine whether some chronic caffeine users displayed symptoms of a substance use disorder. Caffeine dependence, including Caffeine Withdrawal has thus been the focus of much caffeine-related research. Such research continues for DSM-5 where it is categorized as a condition that merits further study. Despite such efforts, little attention has been paid to the actual tools used to make such a diagnosis. The Composite International Diagnostic Interview-Substance Abuse Module (SAM, DSM-IV version) (Cottler, Robins, & Helzer, 1989)

is the only psychometrically sound structured interview that has been developed to assess caffeine dependence (Compton et al., 1996). This measure collects additional clinical information regarding caffeine use (e.g., caffeine withdrawal and age of first regular use) (Cottler, Robins, & Helzer, 1989). The latest version of the SAM (version 4.1) operationalizes caffeine dependence according to DSM-IV criteria (Striley et al., 2011). To date, there is no published version for DSM-5.

Juliano et al. (2012b) recently created the Caffeine Withdrawal Symptom Questionnaire, a 23-item measure in which participants rate on a 5-point scale, how they are currently feeling regarding each withdrawal symptom (Juliano and Griffiths, 2004). While initial testing found good sensitivity (Juliano et al., 2012b), further research is needed to better establish psychometric properties of the questionnaire. Caffeine withdrawal has also been assessed without the use of a formalized measure by asking participants if they have ever experienced withdrawal symptoms, if they have used caffeine to avoid withdrawal symptoms, and if these symptoms have ever interfered with their functioning (Juliano et al., 2012a).

Epidemiology and Problems in College Students. The high rates of caffeine use among college students have been associated with various problems. McIlvain et al. (2011) found that college students consumed, on average, 850 mg of caffeine daily, which is three to five times the recommended amount (200 to 300 mg). Additionally, 83 percent of participants reported experiencing at least one symptom of caffeine intoxication during their lifetime and over half (51 percent) reported at least one symptom of caffeine withdrawal (McIlvain et al., 2011).

In particular, ED consumption in college students has been associated with use of many other drugs and problematic use. College students who consume EDs tend to be greater consumers of alcohol than non ED users (Arria et al., 2010; Terry-McElrath, O'Malley, &

Johnston, 2014). Furthermore, frequent ED users are more likely than infrequent and non-ED users to meet criteria for alcohol dependence (Arria et al., 2011). ED use in college students has also been linked to cigarette and tobacco use, illicit drug use, and misuse of prescription medications (Arria et al., 2010; Terry-McElrath, O'Malley, & Johnston, 2014; Miller, 2008; Hull et al., 2011; Arria, 2009). Other risky behaviors, such as driving while intoxicated and riding with a drunk driver are also associated with ED use (Spierer et al., 2014), as is sensation seeking (Arria et al., 2010; Azagba et al., 2014), particularly among college men (Miller, 2008). Additionally, heightened symptoms of depression have been linked to increased energy drink use (Azagba et al., 2014).

Assessment of Caffeine Use. Standardized, psychometrically sound tools for assessing caffeine use do not exist. Because of the wide array of sources of caffeine and variance by brand and size, measurement of caffeine consumption, particularly quantity consumed, has proven difficult. Instead, researchers have often focused solely on frequency of use, asking about daily versus nondaily use, average days of use per week, or number of days caffeine was used over a set time period (e.g., number of days caffeine consumed in the past 30 days). Other investigators have relied upon multiple choice response options with different categories of use (e.g., none, less than once a month, 1 to 3 times a month, 1 to 3 times a week, and most days) (Miyake & Marmorstein, 2014).

***Assessing several types of caffeine*.** Researcher efforts to measure caffeine use for specific types of caffeine-containing beverages and other sources have typically relied upon non-standardized survey items. As such, these items may be similar across studies, but they often differ in format and specific response options, thereby not always accurately determining the amount of caffeine being consumed. For instance, Terry-McElrath, O'Malley, and Johnston

(2014) chose to provide psychoeducation as part of their efforts to measure both energy drink and caffeinated soda consumption. To do so, they used the following questions: "Energy drinks are non-alcoholic beverages that usually contain high amounts of caffeine, including such drinks as Red Bull, Full Throttle, Monster, and Rockstar. They are usually sold in 8- or 16-ounce cans or bottles. About how many (if any) energy drinks do you drink PER DAY, on average?"; "Energy drinks are also sold as small 'shots,' that usually contain just 2 or 3 ounces. How many (if any) energy drink shots do you drink PER DAY, on average?"; "Regular (non-diet) soft drinks include Coke, Pepsi, Mountain Dew, Dr. Pepper, etc. How many (if any) 12-ounce cans or bottles (or the equivalent) of regular (non-diet) soft drinks do you drink PER DAY, on average?"; and finally "How many (if any) 12-ounce cans or bottles (or the equivalent) of diet soft drinks (like Diet Coke, Diet Pepsi, etc.) do you drink PER DAY, on average?" Response options for each question included none, less than 1, one, two, three, four, five or six, 7 or more (Terry-McElrath, O'Malley, & Johnston, 2014). Arria et al. (2010) did not include the level of informative detail in questions, but did allow for many different types of caffeine through the use of an open-ended response to the following question: "What types of caffeinated products do you consume?" Participants were also given a reference card for caffeine types (e.g., coffee, tea, soft drinks). They additionally attempted to assess quantity, examining all sources summed together by asking participants to estimate the typical, maximum, and minimum number of caffeinated beverages consumed in an average week, referencing a standard serving at being about one 8 oz cup of coffee (Arria et al., 2010). The accuracy of such methods for quantifying caffeine use is not yet known.

As an alternative, since ED consumption has been associated with a higher risk of substance use when compared to use of other caffeine-containing beverages (Terry-McElrath,

O'Malley, & Johnston, 2014), some researchers have measured only ED use, ignoring other sources of caffeine. Such studies typically ask about prevalence of past year energy drink use with response options ranging from at least once, three or more times, once or more in a month, and finally more than once in a month have been used in several surveys in youth populations (Attila and Cakir, 2011; Berger et al., 2011; Malinauskas et al., 2007; Azagba et al., 2014). Similarly, Miller (2008) included a single question asking participants how frequently in the past month they had consumed Red Bull or a similar energy drink, with response options 0 days, 1–2 days, 3–5 days, 6–9 days, 10–19 days, 20–29 days, and all 30 days. Other single question formats have asked about the number of energy drinks consumed on average per week, with response options from 0 (less than one) to 5 (15 or more) (Skewes et al., 2013). Trapp et al. (2014) included a scale assessing ED use without a time period of assessment (response options included never, <1/month, 1 day/month, 2 days/month, 3 days/month, 1 day/week, 2 days/week, 3 days/week, 4 days/week, 5 days/week, 6 days/week and every day) and an item about the usual amount of energy drinks (total number of cans) consumed per day on a day that they consumed an ED. In a younger sample, our lab recently found moderate/heavy energy drink users were most likely to report use of alcohol, tobacco, and xx other classes of drugs, followed by light energy drink users and finally non-users (Polak et al., under review). Taken together, the majority of these assessment measures have not been psychometrically tested, with little reliability or validity data.

The Caffeine Exposure Questionnaire (CEQ) (Harrell & Juliano, 2009; Svikis et al., 2005) was designed to obtain information on quantity and frequency of caffeinated products (e.g., coffee, tea, soft drinks, chocolate, etc.) consumed in a typical day. This measure is well established and has been used in many published studies (e.g., Juliano et al., 2012b, Svikis et al.,

2005, Harrell and Juliano, 2009 and Huntley and Juliano, 2011). Additionally, the Caffeine History Questionnaire was designed to collect other relevant caffeine information, including past caffeine-related advice given by health providers, previous attempts to quit or reduce caffeine use, and current desire and confidence to quit/reduce caffeine use (Juliano et al., 2012a).

Alcohol

Introduction. Alcohol is a depressant that comes in several forms, the most common of which are beer, wine, and liquor. There are many short-term effects of alcohol, such as slurred speech, nausea/vomiting, impaired judgement, and decreased coordination (Brown University, n.d.). High doses of alcohol are dangerous and can lead to a significant decrease in respiratory activity, coma or death (Brown University, n.d.).

The personal, social, and economic consequences of excessive alcohol use can be devastating for individuals, families, and communities. In 2006, excessive alcohol consumption cost the US 223.5 billion dollars (Epstein & McCrady, 2009). Individuals with alcohol use disorders (AUDs) often have other substance use disorders as well as psychiatric comorbidities (Hasin, Stinson, Ogburn, & Grant, 2007). Organic complications of AUDs include medical problems (e.g., nutritional deficits, pancreatitis, and liver disease), cognitive deficits, and chronic pain, with higher mortality rates across all age groups (Epstein & McCrady, 2009; Bates, Bowden, & Barry, 2002). Additionally, legal problems and crime are elevated in this population, with approximately one-fourth of all violent offenders having consumed alcohol at the time their crimes were committed (Bouchery, Harwood, Sacks, Simon, & Brewer, 2011).

Prevalence. Approximately 87 percent of individuals 18 years or older have reported lifetime alcohol use, 71 percent drank in the past year, and 56 percent have consumed alcohol in

the past month (SAMHSA, 2013d). For the past 50 years, beer has been the most popular alcoholic beverage, followed by liquor and wine (Greenfield, Midanik, & Rogers, 2000).

Youth is associated with heightened risk for substance use. The majority of individuals in the US have consumed at least one alcoholic drink by late adolescence (SAMHSA, 1999a). Specifically, 80 percent of 12th graders, over two-thirds of 10th graders, and over half of 8th graders reported lifetime alcohol use (Johnston et al, 1998). Among the emerging adult population, alcohol is the most widely abused substance (DHHS, 2007), with the highest rates of use peaking in young adulthood (Johnston et al, 2008; Merikangas et al., 2010; Swendsen et al., 2012). Within the emerging adult population, college students are particularly vulnerable to alcohol use and binge drinking (SAMHSA, 2013b).

Heavy Alcohol Use. Historically, definitions for such terms as heavy or binge drinking have varied considerably. Such inconsistency led to confusion among both researchers and clinicians. In an effort to reach consensus, in 2004, the National Institute on Alcohol Abuse and Alcoholism (NIAAA) convened an Advisory Council Task Force to better define the term binge drinking. In their report, the group recommended that binge drinking be defined as that which increases blood alcohol concentration (BAC) to at least 0.08 gm% (NIAAA, 2005). For men, this BAC is typically achieved when at least five drinks have been consumed in approximately two hours (NIAAA, 2005). Since women are physiologically prone to significantly higher concentrations of alcohol than men (Epstein & McCrady, 2009), this binge drinking BAC is typically reached by women after four drinks (NIAAA, 2005). Nonetheless, many continue to use the terms heavy and binge drinking interchangeably.

Epidemiologically, heavy alcohol consumption is particularly concerning, with 75 percent of the economic costs of alcohol use linked to binge drinking (Epstein & McCrady,

2009). In 2013, almost one-fourth (23 percent) of individuals over 12 years old were past month binge drinkers (SAMHSA, 2013c). Specifically, almost one-third (29 percent) of individuals 18 to 20 years old and almost half (43 percent) of those age 21 to 25 reported lifetime binge drinking (SAMHSA, 2013c). Additionally, nine percent of 18 to 20 year old individuals and 13 percent of those ages 21 to 25 reported current heavy drinking (SAMHSA, 2013c). Among emerging adults, males are more likely to be current binge drinkers than females (44 percent and 31 percent, respectively) (SAMHSA, 2013c). Additionally, college students are more likely than non-college attending counterparts to engage in past-month binge drinking (39 percent versus 33 percent, respectively) and heavy drinking (13 percent versus 9 percent, respectively) (SAMHSA, 2013b).

Diagnostic Criteria. In the DSM-IV, alcohol-related disorders (and other substance use disorders) were separated into two distinct categories (abuse and dependence), each with its own set of diagnostic criteria. With the transition from DSM-IV to DSM-5 in 2013, substance use disorder diagnostic criteria shifted. The DSM-5 integrates and re-conceptualizes DSM-IV criteria, taking a dimensional approach by placing substance use disorders on a continuum from mild to severe. Using this diagnostic system, severity level is determined by the number of criteria that have been met. Several other adjustments were made in the transition to DSM-5, including language modification and addition/deletion of certain criteria (APA, 2013). DSM-5 Alcohol Use Disorder diagnostic criteria are listed in Table 1. In addition to Alcohol Use Disorder, the DSM-5 recognizes two other alcohol-related disorders, Alcohol Intoxication and Alcohol Withdrawal, which are outlined in detail in Tables 2 and 3, respectively (APA, 2013).

Alcohol use disorders are among the most common psychiatric diagnoses in the US (Grant et al., 2004). Overall, almost one-third (30 percent) of the population meets lifetime

alcohol abuse or dependence criteria and approximately nine percent of individuals have a current alcohol-related disorder (Hasin, Stinson, Ogburn, & Grant, 2007). Specifically, the 12-month prevalence rate of alcohol-related disorders in US adults is about 12 percent (Merikangas & McClair, 2012). Additionally, approximately one-fourth of all patients in mental health settings have an alcohol-related disorder (Hasin, Stinson, Ogburn, & Grant, 2007). Compared to earlier data on prevalence rates, alcohol-related disorders appear to be increasing in number (Hasin, Stinson, Ogburn, & Grant, 2007). Among US youth, the prevalence of alcohol-related disorders is approximately eight percent, which is slightly lower than the overall average (Merikangas & McClair, 2012). Rates of alcohol-related disorders are significantly higher among males than females (Compton et al., 2007). Additionally, this gender difference is greater in adults than adolescents (Merikangas et al., 2010).

Heavy alcohol use can lead to alcohol use disorders. For example, binge drinking has been linked to alcohol use disorders, with about 25 percent of binge drinkers meeting criteria for an alcohol use disorder (NIAAA, n.d.b). Additionally, as the number of past month binge drinking days increases, so does risk for an alcohol use disorder (NIAAA, n.d.b).

Table 8. DSM-5 Criteria for Alcohol Use Disorder

A. A problematic pattern of alcohol use leading to clinically significant impairment or distress, as manifested by at least two of the following, occurring within a 12-month period:
    1. Alcohol is often taken in larger amounts or over a longer period than was intended.
    2. There is a persistent desire or unsuccessful efforts to cut down or control

alcohol use.

3. A great deal of time is spent in activities necessary to obtain alcohol, use alcohol, or recover from its effects.
4. Craving, or a strong desire to urge to use alcohol.
5. Recurrent alcohol use resulting in a failure to fulfill major role obligations at work, school or home.
6. Continued alcohol use despite having persistent or recurrent social or interpersonal problems caused or exacerbated by the effects of alcohol.
7. Important social, occupational, or recreational activities are given up or reduced because of alcohol use.
8. Recurrent alcohol use in situations in which it is physically hazardous.
9. Alcohol use is continued despite knowledge of having a persistent or recurrent physical or psychological problem that is likely to have been caused or exacerbated by alcohol.
10. Tolerance, as defined by either of the following:
    a. A need for markedly increased amounts of alcohol to achieve intoxication or desired effect.
    b. A markedly diminished effect with continued use of the same amount of alcohol.
11. Withdrawal, as manifested by either of the following:
    a. The characteristic withdrawal syndrome for alcohol (refer to Criteria A and B of the criteria set of alcohol withdrawal).
    b. Alcohol (or a closely related substance, such as a benzodiazepine) is

taken to relieve or avoid withdrawal symptoms.
Specify if: **In early remission:** After full criteria for alcohol use disorder were previously met, none of the criteria for alcohol use disorder have been met for at least 3 months but for less than 12 months (with the exception that Criterion A4, "craving, or a strong desire or urge to use alcohol," may be met). **In sustained remission:** After full criteria for alcohol use disorder were previously met, none of the criteria for alcohol use disorder have been met at any time during a period of 12 months or longer (with the exception that Criterion A4, "Craving, or a strong desire or urge to use alcohol," may be met).
Specify if: **In a controlled environment:** This additional specifier is used if the individual is in an environment where access to alcohol is restricted.
Specify current severity: **Mild:** Presence of 2-3 symptoms. **Moderate:** Presence of 4-5 symptoms. **Severe:** Presence of 6 or more symptoms.

Note. Adapted from *DSM-5*, p. 490-491, APA, 2013, Arlington.

Table 9. DSM-5 Criteria for Alcohol Intoxication

A. Recent ingestion of alcohol.
B. Clinically significant problematic behavior or psychological changes (e.g., inappropriate sexual or aggressive behavior, mood lability, impaired judgement) that developed during, or shortly after, alcohol ingestion.
C. One (or more) of the following signs or symptoms developing during, or shortly after, alcohol use: 1. Slurred speech, 2. Incoordination. 3. Unsteady gait. 4. Nystagmus. 5. Impairment in attention or memory. 6. Stupor or coma.
D. The signs or symptoms are not attributable to another medical condition and are not better explained by another mental disorder, including intoxication with another substance.

Note. Adapted from *DSM-5*, p. 497, APA, 2013, Arlington.

Table 10. DSM-5 Criteria for Alcohol Withdrawal

A. Cessation of (or reduction in) alcohol use that has been heavy and prolonged.
B. Two (or more) of the following, developing within several hours to a few days after the cessation of (or reduction in (alcohol use described in Criterion A: 1. Autonomic hyperactivity (e.g., sweating or pulse rate greater than 100 bpm). 2. Increased hand tremor. 3. Insomnia. 4. Nausea or vomiting. 5. Transient visual, tactile, or auditory hallucinations or illusions. 6. Psychomotor agitation. 7. Anxiety. 8. Generalized tonic-clonic seizures.
C. The signs or symptoms in Criterion B cause clinically significant distress or impairment in social, occupational, or other important areas of functioning.
D. The signs and symptoms are not attributable to another medical condition and are not better explained by another mental disorder, including intoxication or withdrawal from another substance.
Specify if: **With perceptual disturbances**: This specifier applies in the rare instance when hallucinations (usually visual or tactile) occur with intact reality testing, or auditory, visual, or tactile illusions occur on the absence of delirium.

Note. Adapted from *DSM-5*, p. 499-500, APA, 2013, Arlington.

Epidemiology and Problems in College Students. The high prevalence rates of alcohol use among emerging adults are associated with many adverse consequences. Of the total annual economic cost of alcohol use, underage drinking accounts for 62 billion dollars (Foster et al., 2003). Recent annual estimates in college students suggest that over half a million (599,000) are injured under the influence of alcohol, more than 150,000 develop health problems as a result of alcohol use, and 1,825 die as a result of problems related to alcohol use (NIAAA, n.d.a). Over 690,000 students are assaulted and more than 97,000 are sexually assaulted by another student who is under the influence of alcohol (NIAAA, n.d.a). Additionally, one in four college students reports academic problems as a result of alcohol use (NIAAA, n.d.a). These alarming rates of alcohol misuse among emerging adults prompted the Centers for Disease Control and Prevention to label alcohol abuse a priority health risk behavior (CDC, 2008).

Assessment of Alcohol Use. While 25 years ago, alcohol-related measures were few and far between (Alanko, 1984; Room, 1990; Sobell & Sobell 1992; Sobell & Sobell 1995), the past few decades have seen a growing literature describing a plethora of measures focused on quantity and frequency of alcohol use with psychometric evaluation across diverse study samples (Sobell & Sobell, 2004). There are different ways of characterizing drinking, such as recent (typically past 30 days), past 3 months, past year, and lifetime (any) use. Additionally, the intention of alcohol instruments ranges from assessment of general use, screening for problems, and diagnostic assessment.

***Quantity and Frequency Instruments*.** In the measurement of alcohol use, two recall approaches have been examined: usual quantity/frequency (QF) and graduated frequency (GF) (Dawson, 2003). Measures based on the QF approach include questions that assess for overall alcohol use frequency and quantity on a typical drinking occasion (Dawson, 2003). Responses to

typical quantity questions usually fall between the mean and mode (Gruenewald et al. 1996). Instruments that adhere to the GF approach ask how often during a period of time participants consumed amounts of standard drinks (e.g., 1 to 2 drinks, 3 to 4 drinks, etc.), with the usual format consisting of a question about largest amount followed by a question of frequency of use of the applicable quantity categories (from the maximum reported amount and all those falling below). Most large-scale current information about alcohol use is obtained using a small number of QF questions (Fishburne & Brown, n.d.). This reliance on the QF approach is potentially problematic as it appears to have certain methodological flaws.

Comparisons between QF and GF measures have revealed that GF approaches typically result in greater amounts of drinks per drinking day reported (Fishburne & Brown, n.d.). One possible reason for this discrepancy is that the GF approach is more specific, requiring less averaging and consolidating of alcohol use when compared to the QF approach (Fishburne & Brown, n.d.). This difference also potentially indicates that college students underestimate the average number of drinks consumed when reporting alcohol use on a QF measure (Fishburne & Brown, n.d.).

Specific QF Instruments. Straus and Bacon (1953) created the first QF measure (Quantity-Frequency) that assessed the average quantity and frequency of alcohol use over the past year. Participant responses were used to classify individuals by "typical" drinking pattern. As a result of these overgeneralizations during a large timeframe, this first attempt lacked the sensitivity needed to capture variability in drinking (Room, 1990).

The Volume-Variability (VV) Index (Cahalan and Cisin, 1968) and the Quantity-Frequency Variability (QFV) Index (Cahalan et al., 1969) attempted to address this problem of capturing variability of drinking patterns (Alanko, 1984; Room, 1990). These measures asked

about frequency of alcohol use and the "proportion of drinking occasions" for the numbers of drinks listed (Cahalan et al. 1969). The QFV also included a question about the maximum amount per occasion (Cahalan et al. 1969). Answers on the QFV were used to place individuals into one of 11 categories of quantity and variability of use (Cahalan et al. 1969). The QFV Index score would then be attained with consideration of both the category and frequency of use. Using the QFV classifications, participants would be labeled by type of drinker (heavy, moderate, light, infrequent, and abstainers) (Cahalan et al. 1969).

The VV Index uses a classification system composed of eight categories (see Cahalan et al. 1969, p. 215) derived from the aggregate volume (Q x F) and maximum quantity (Cahalan and Cisin 1968). The VV Index also calculates the average daily volume and whether or not an individual had at least five drinks on any occasion (Cahalan et al. 1969). The VV Index is more sensitive to the middle range of alcohol use than the QFV Index (Khavari and Farber 1978). However, for both the QFV and VV Index, the upper range drinking category (five or more drinks) creates a low ceiling, rendering these measures insensitive to extremely heavy drinking.

Many distinct QF measures have been created using these original measures as a framework. Khavari and Farber (1978) created the Khavari Alcohol Test, which is a 12–question derivation of the QFV. For three beverage types, individuals report typical drinking frequency, typical quantity, maximum quantity, and frequency of maximum quantity. These answers are used to both categorize individuals into one of 11 frequency categories and provide a norm-based comparison. Bowman and colleagues (1975) created the Volume-Pattern Index, which differed from previous QF measures in that individuals were placed on a continuum of volume and pattern of drinking, as opposed to categories. This measure has been criticized for its complicated data manipulation (Khavari and Farber 1978) and lengthy administration time of 30-

60 minutes. Polich and Orvis (1979) created the Composite QF Index, which asks about past month alcohol consumption across beverage categories and frequency of heavy drinking from the past year. The Rand QF (Polich et al., 1981) attempted to assess typical drinking for each beverage type as well as heavy drinking (i.e., asked about number of days characterized by high levels of use). The Lifetime Drinking History (LDH) (Skinner and Sheu, 1982) differed from previous measures by adding questions about lifetime alcohol use. The Cognitive Life Drinking History (Russell et al., 1997) was found to be reliable and assesses lifetime alcohol use and includes a "floating" interval in which individuals are asked to report when their patterns change and specific questions about beverages/sizes (Russell et al., 1997).

More recently, with a focus on translational research and the increased emphasis on brief screening to identify those at risk for alcohol problems, the form of the QF approach has shifted from lengthy assessment and analysis to one or two question screeners. With this approach, the results of brief QF screening are then used to determine if further assessment (e.g., administering the AUDIT-C) and necessary action (e.g., brief intervention) are needed (Strobbe, 2014). These brief QF items typically take the following form "How many times in the past year have you had X or more drinks in a day?" (where X is 5 for men and 4 for women), with an answer over one designated as positive (Smith et al., 2009). These one or two item screeners have been shown to be valid and relatively sensitive, and offer clinicians something practical and easier to implement (Strobbe, 2014).

Problems with screeners and diagnostic assessment. Despite the significant time and attention devoted to the development and implementation of alcohol-related measures, recent reports found methodological issues, with Devos-Comby and Lange (2008) noting that a strong theoretical framework has not yet been established. The three domains for comprehensive

assessment of problematic use include Alcohol Abuse, Alcohol Dependence (or Alcohol Use Disorder), and risky drinking (Devos-Comby & Lange, 2008). Devos-Comby & Lange (2008) reviewed current measures and found that most do not include all domains.

Additionally, the classic screeners contain questions that treat individuals with alcohol problems as a homogeneous group. As result, the measures designed to capture problematic use may lack sensitivity to differences that are crucial in identification of those at risk. Recently, increased attention has been paid to variations in temporal patterns of alcohol use (e.g., Del Boca, Darkes, Greenbaum, & Goldman, 2004), showing that for erratic heavy drinking, a focus on assessment periods that are too short can contribute to misclassification of individuals (Chen & Kandel, 1995; Lemmens, Tan, & Knibbe, 1992).

Additional factors have been shown to affect the psychometric properties of these measures. The use of supplemental memory aids during the assessment process (Midanik & Hines, 1991; L.C. Sobell & Sobell, 1992; Hammersley, 1994; Single & Wortley, 1994) and structuring measure items to facilitate both recall and mitigation of socially desirable responses have been shown to improve reliability and validity (Embree & Whitehead, 1993).

***Interview-Based Measures*.** In contrast to QF measures (and particularly the recently emphasized brief QF approach), interview-based measures require more resources, including administration time and clinician attention. However, this type of assessment often provides the most comprehensive picture of alcohol use and associated problems (Samet et al., 2007). As such, interview-based measures have generally been found to be psychometrically sound (Samet et al., 2007). There are a number of structured and unstructured interview-based measures that collect some quantity and frequency data, but focus primarily on alcohol-related problems. They include the Addiction Severity Index (ASI), Composite International Diagnostic Interview

(CIDI), Structured Clinical Interview for DSM-IV (SCID), Alcohol Use Disorders and Associated Disabilities Interview Schedule (AUDADIS), Psychiatric Research Interview for Substance and Mental Disorders (PRISM), and Semi-Structured Assessment for Drug Dependence and Alcoholism (SSADDA) (Samet et al., 2007).

*TLFB***.** For more detailed collection of quantity and frequency of use data, however, the TLFB has long been the gold standard. The TLFB is a widely-used, semi-structured interview that uses a calendar to retrospectively collect daily information about alcohol and other drug use. TLFB daily behavior reporting is facilitated through the use of a trained interviewer and memory aids, such as calendars and anchor dates. It was originally developed by Sobell and Sobell (1992) to collect data on alcohol consumption and was shown to have good reliability and validity. Subsequently, it has been modified to focus on other drug use, across different populations, and over extended time intervals, ranging from 30 to 360 days (Fals-Stewart et al., 2000; Robinson et al., 2014; Norberg, Mackenzie, & Copeland, 2012; DeMarce et al., 2007; Carey et al., 2004). For alcohol use, the TLFB has been found to significantly correlate with the Addiction Severity Index Alcohol Use Severity subscale (Fals-Stewart et al., 2000; DeMarce et al., 2007), Michigan Alcohol Screening Test (Fals-Stewart et al., 2000), and collateral reports (DeMarce et al., 2007). Additionally, the TLFB has been adapted to focus on other behaviors (e.g., sexual risk behavior, exercise, and work) (Carey et al., 2001; Panza et al., 2012; Svikis et al., 2012).

Comparing Survey and Interview**.** Brief surveys assessing alcohol problems have been linked with possible over-identification of alcohol use disorders when compared to more in-depth interviews and young heavy drinkers appear to be at greatest risk for this misclassification (Caetano and Babor, 2006; Midanik, Greenfield, & Bond, 2007). Potential reasons for this over-identification through survey items include participants misinterpreting questions (Slade et al.,

2013) or misperceiving the intoxicating effects or symptoms of alcohol use disorders (Caetano and Babor, 2006). Karriker-Jaffe, Witbrodt, and Greenfield (2015) recently conducted a study in which follow up questions were asked of individuals who previously completed in-depth interviews. They found a particular susceptibility for reporting errors for both questions about withdrawal and drinking larger amounts/longer than intended (Karriker-Jaffe, Witbrodt, & Greenfield, 2015).

Alternatives to Self-Report. Biological measures of alcohol use have been used to increase the validity of self-report information (Wish et al., 1997). Alcohol levels can be estimated through collection and analysis of breath, blood, and urine. Alcohol biomarkers include direct and indirect biomarkers (SAMHSA, 2006). Indirect biomarkers identify alcohol's effects on organ systems or body chemistry (i.e., mean corpuscular volume [MCV], gamma-glutamyltransferase [GGT], aspartate aminotransferase [AST], alanine aminotransferase [ALT], carbohydrate-deficient transferrin [CDT], total serum sialic acid [TSA], 5-hydroxytryptophol [5-HTOL], N-acetyl-beta-hexosaminidase [Beta-Hex], plasma sialic acid index of apolipoprotein J [SIJ], and salsolinol). Direct biomarkers identify alcohol or parts of the alcohol metabolism process (i.e., acetaldehyde, acetic acid, fatty acid ethyl ester [FAEE], ethyl glucuronide [EtG], ethyl sulfate [EtS], and phosphatidylethanol [PEth]). These methods have proven useful in detecting problems with retrospective measures; for instance, de Beaurepaire et al. (2007) found that 37 to 56 percent of patients underreported alcohol consumption based on comparison with biological measures.

Caffeine Combined with Alcohol

Introduction. The combined use of alcohol and caffeine (CAC) refers to the following types of beverages: prepackaged, premixed, and spontaneously prepared (e.g., combining Red

Bull and vodka right before drinking it). Typically, prepackaged CAC has undisclosed caffeine contents, is malt- or distilled-spirits based, and has a higher alcohol content than usually found in beer; five to 12 percent is the average ABV for prepackaged CAC, while four to five percent is the average ABV for beer (CDC, 2014). CAC use has been linked to dehydration, bad hangover, vomiting, heart palpitations, and increased CAC use (Brache et al., 2012).

Prevalence. The rate of CAC use increased following the introduction of prepackaged CAC in 2002 (M. Shanken Communications, Inc., 2009). Popularity of prepackaged CAC grew, experiencing a growth from 337, 500 gallons sold in 2002 to 22,905,000 gallons sold in 2008 (M. Shanken Communications, Inc., 2009). During this time, these drinks were heavily marketed toward youth (e.g., using youth-oriented images) (Simon & Mosher, 2007). In 2010, the FDA told four companies that sold prepackaged CACs that these drinks could not stay on the market because the added caffeine was an “unsafe food additive" (FDA, 2015). This FDA warning only targeted seven beverages in total and did not include alcoholic beverages that contain caffeine as a natural component of an ingredient (e.g., coffee flavor) (FDA, 2015). After this warning was issued, these companies removed the caffeine from the targeted beverages (FDA, 2015).

Despite the reduction in prepackaged CACs available for purchase, CAC use has become increasingly prevalent. Among college students, about one-fourth to one-half report past month CAC use (Brache & Stockwell, 2011; MacKillop et al., 2012; Miller, 2008). Additionally, rates of regular CAC use have been found to range from approximately one-fourth (O'Brien, McCoy, Rhodes, Wagoner, & Wolfson, 2008) to over one-third (34 percent) of college students (Mintel International Group Ltd., 2007).

Epidemiology and Problems in College Students. In addition to high prevalence rates, concern regarding CAC use among college students stems from the associated adverse health

and safety consequences of such use. While relatively little research has been done on CAC use, findings suggest CAC use may exacerbate the negative consequences of problematic drinking. Compared to those consuming the same amount of alcohol by itself, those consuming CAC are: more likely to drive while impaired, and be hurt or injured (Brache et al., 2012); twice as likely to report being taken advantage of sexually, taking advantage of someone else sexually, and riding with a driver who was under the influence of alcohol (O'Brien et al., 2008); three times more likely to binge drink (Thombs et al., 2010); more likely to report alcohol dependence, increased drug use (marijuana, ecstasy, and cocaine), sex under the influence of alcohol/drugs, unsafe sexual activities, and other risk-taking behavior (Arria et al., 2010; Arria et al., 2011); and they are more likely to leave a bar when intoxicated (Thombs et al., 2011).

One possible reason for the aforementioned increased risks and harms associated with CAC use is that individuals have a decreased awareness of the extent to which they are intoxicated. The effects of alcohol appear to be masked by the caffeine in CACs; individuals drink longer and consume more alcohol because of this effect (Arria et al., 2012). While the caffeine found in CACs can mask the depressant effects of alcohol, it has no effect on the metabolism of alcohol and thus does not reduce breath alcohol concentrations or the risk of alcohol-attributable harms (Ferreira et al., 2006).

Assessment of CAC Use. Research into the amount of knowledge and awareness about CAC use has been sparse. Recent findings have shown that participants often do not understand what beverages contain caffeine (Polak, Hancock, & Svikis, n.d.a; Polak, Hancock, & Svikis, n.d.b); a barrier to assessing caffeine use and thus CAC use is a lack of knowledge about what contains caffeine.

Similar to caffeine by itself, the assessment of CAC use is often done through the use of a limited number of unstandardized questions. Cobb et al. (2013) assessed CAC use (including prepackaged and premixed) focused on three different time points (lifetime, past year, and past 30 days). Those reporting CAC use, completed additional detailed items about CAC use (e.g., "What [alcoholic beverage, caffeinated beverage in combination with alcohol or caffeinated alcoholic product] do you prefer?") (Cobb et al., 2013). Prior to CAC item administration, participants were given brief psychoeducation about common CAC drinks (e.g., rum and Coke) (Cobb et al., 2013). The assessment also asked about reasons for CAC use, with questions adapted from O'Brien et al. (2008).

Additionally, CAC items often focus specifically on alcohol combined with energy drinks, leaving out other caffeinated beverages. Varvil-Weld et al. (2013) modified items from the Daily Drinking Questionnaire (DDQ); they asked participants to report the typical quantity of CAC use for each day of the week. Other researchers have assessed alcohol and energy drink use by asking about number consumed during a typical session where CAC was used (past three months) through free response option (Droste et al., 2014). Past year CAC use has previously been assessed with a single item, "In the last 12 months, have you had alcohol mixed or pre-mixed with an energy drink such as Red Bull, Monster, Rock Star, or another brand?" with response options "I have never done this," "I did not do this in the last 12 months," "yes," and "I don't know" (Reid et al., 2015).

To date, the only published attempt to develop a standardized measure of CAC use was the Caffeine + Alcohol Combined Effects Questionnaire (CACEQ), and this survey does not focus on quantity and frequency of CAC use. Rather, it asks nine questions about expectancies of CAC use (MacKillop et al., 2012). Participants rate these expectancies on a 5-point scale from

strongly disagree to strongly agree. Through this assessment, participants are asked about whether or not they drank CACs (separately for premixed or mixed ad hoc) and the frequency of premixed and ad hoc CAC consumption in the past 30 days (MacKillop et al., 2012). While reliability and validity have been established during initial measure development process (MacKillop et al., 2012), there are some apparent issues; for instance, this measure is not one of general CAC use expectancies, but rather that of energy drinks combined with caffeine use as indicated by the question instructions.

Method

Participants

Participants were $N = 50$ students 18 years or older recruited from the Virginia Commonwealth Student Health Center on the Monroe campus. Research staff posted flyers (see Figure 1) in the clinic waiting area and were seated in the clinic waiting room ready to discuss the project with students who express interest in learning more about the study. Study eligibility was determined pre- or post-clinic visit using the criteria summarized below. Students who met criteria and wanted to participate in the study were escorted to a semi-private area where they provided informed consent and completed the study. Additionally, students who were eligible and interested in participating, but did not have time that day, were scheduled to come to the Svikis lab within the following few days for enrollment. All participants were current VCU students, as this is a requirement to receive services through the Student Health Center.

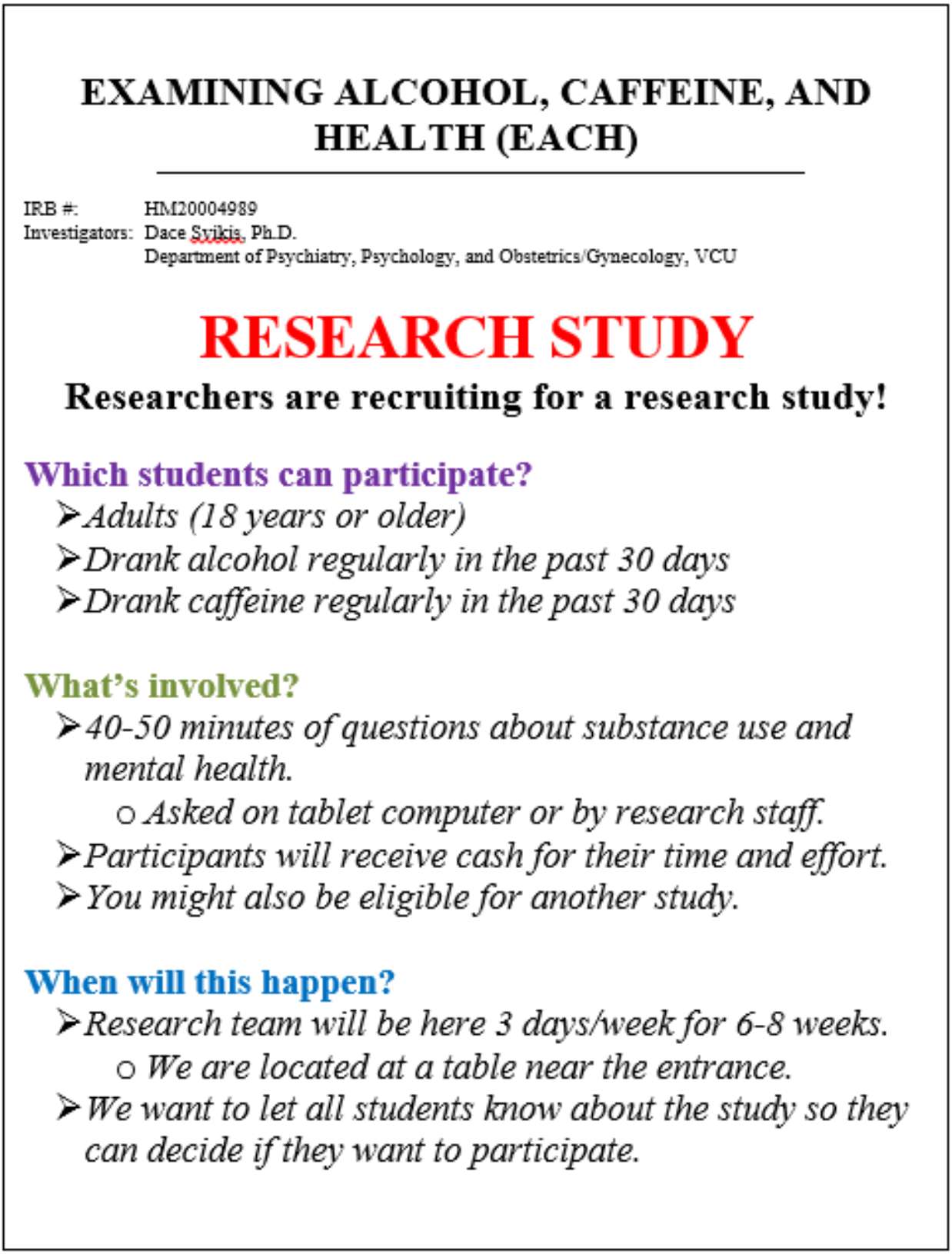

Figure 1. EACH study flyer.

Phase 1: Quantitative (Survey and Timeline Follow-Back).

Inclusion Criteria. Students were eligible to participate in the study if they met the following criteria:

a) 18 years of age or older.
b) Current VCU student (i.e., eligible for University Student Health Center services)
c) Seeking services or information at the Student Health Center (though having a visit that day is not required).
d) Reporting recent (past 30 day) regular use (at least one day per week) of alcohol (beer, wine or liquor).
e) Reporting recent (past 30 day) regular use (at least two days per week) of caffeine (coffee, tea, energy drinks, and soft drinks).
f) Able to read and understand English.
g) Able to provide informed consent.

Exclusion Criteria. Students were excluded from study participation if they met any of the following criteria:

a) Previously enrolled in the study.

Phase 2: Qualitative Interview.

Inclusion Criteria. Phase 1 participants were eligible for Phase 2 of the study if they reported recent combined use of caffeine and alcohol during Phase 1 data collection.

***Setting*.** Recruitment took place in the Virginia Commonwealth University Student Health Center on the Monroe campus. The center serves any student currently enrolled in classes. Services include outpatient primary care (e.g., allergy shots, blood and/or body fluid

exposures, immunizations, mental health, nutrition consults, travel health care, and women's health), prescription needs through an on-site pharmacy, and laboratory services.

Study Procedures

Recruitment. Study staff were stationed in the Student Health Center waiting room with resources available to answer study-related questions from students entering or exiting the clinic. IRB-approved fliers were placed on waiting room tables. The study staff approached students in the clinic waiting area, asking if they would be interested in the study. Students were told that the survey was about use of both caffeine and alcohol as well as other health behaviors in college students.

Screening. Students who expressed an interest in the study were screened to determine if they were eligible for the study. To avoid interrupting patient care, this screening process was relatively brief and structured using an IRB-approved script (see Figure 2). During this screening, students were asked about the following (corresponding to eligibility criteria): age, status as both a VCU student and an individual seeking services at the University Student Health Center, and recent (past 30 day) regular use of alcohol and caffeine. Recruitment information was maintained through an electronic database via tablet computer. For each day of recruitment, the number of patient refusals and reasons for non-participation were also recorded.

Research staff will utilize the following semi-structured recruitment script when approaching students in the waiting area or when responding to students who approach the research table in the waiting area.

"Good (morning/afternoon). I'm (insert name), a research assistant at VCU. Before I tell you more about the study, may I ask you a few questions to see if you are eligible?"

(If no, please thank the person for their time and go to the next person).

If yes,

"Ok, thank you.

Are you a current VCU student seeking services here today? **[Yes = eligible]**

How old are you? **[≥18 = eligible]**

In a typical week during the past month, how often did you drink caffeinated beverages (e.g., sodas, energy drinks, coffee)? **[≥2 times = eligible]**

In a typical week during the past month, how often did you drink alcohol?" **[≥1 time = eligible]**

(If ineligible, please thank the person for their time)

If eligible,

We are conducting an anonymous survey about substance use and health behaviors. It takes about 30-45 minutes to complete and you will receive $20 for your time. If it is ok, I would like to leave this Information Sheet with you to read over. Would this be something you would be interested in participating in?"

(If no, please thank the person for their time)

If yes,

"Great! Because we don't want to interfere with what you are doing here today, please see me or my associate (point to the person) (name) after you are finished. Thank you!"

Figure 2. EACH recruitment script.

Informed Consent.

Phase 1. Students who met inclusion criteria and wanted to participate in the study were asked to read an IRB-approved information sheet. This form describes all Phase 1 study elements, as well as the voluntary nature of the study; and limits of confidentiality. It emphasized that their decision about study participation would not impact their care at the health center or as a student at VCU. Research staff emphasized that the study was anonymous. Their names and other identifying information (e.g., date of birth) were not collected at any time during study participation. Additionally, participants were told that the survey could take between 25-45 minutes and that they could stop at any point without negative repercussions. After they read the information sheet, the study staff answered any questions and made sure each student understood what they were being asked to do. If the student elected to participate, he/she provided verbal consent and the study staff proceeded with random assignment. If a student chose not to participate, study staff recorded this in the electronic database via tablet computer. Information about students who did not meet study criteria were also be recorded in this database.

Randomization. Students who consented to Phase 1 were randomly assigned to one of two groups:

Group 1. TLFB Interview followed by the Computer-Assisted Survey (CAS).

Group 2. CAS followed by TLFB Interview.

To determine group assignment, study staff opened envelopes prepared by Dr. Svikis in advance. Each envelope contained an index card that said Group 1 or Group 2. The schedule for randomization was determined using a random number generating application for the iPhone. The opaque envelopes were numbered sequentially and study staff used them in this order when they assigned participants to groups.

Groups 1 and 2 were identical except for order in which the CAS and TLFB interview were administered. Counterbalancing the order of administration controlled for order effects and the possibility of test sensitization.

Phase 2: Qualitative Interview. Students who reported consuming caffeine combined with alcohol (past 30 days) in Phase 1 data collection were eligible for Phase 2. Similar to Phase 1, eligible students were given an IRB-approved information sheet describing Phase 2 study participation. Again, the voluntary nature of the study was emphasized and limits of confidentiality were explained.

It is important to note that participants who qualified for Phase 2 were only informed about Phase 2 after they completed Phase 1. This procedure was not only practical but also integral to the research design. From a practical perspective, only students who reported recently consuming alcohol combined with caffeine were eligible for Phase 2. Therefore, Phase 1 data were needed to determine Phase 2 eligibility. In addition, however, if students were told in advance about Phase 2, this could easily influence or even change their responses during Phase 1 data collection.

Computer-Assisted Survey (CAS). In the CAS, participants were asked about current (past 30 days) alcohol, caffeine, and CAC use. This period of assessment is congruent with the TLFB assessment period (past 30 days) and thus enabled a comparable evaluation. They also completed surveys on mood and personality. These items helped to both identify correlates of substance use as well as to aid in masking the primary purpose of the study.

The Spit for Science survey was used as a template for many of the substance use questions that were included in the CAS. The inclusion of these questions enabled a comparison of the two surveys. We were able to assess this study's representativeness through a comparison

with the large sample of students who answered the Spit for Science survey. Additionally, any interesting qualitative information gathered about these questions will potentially help inform future research, including the Spit for Science survey.

Participants completed the survey on a tablet computer, thereby eliminating need for research staff data entry. Research staff explained use of the device and provided an overview of survey procedures (e.g., it will self-administered, they can ask questions as needed). There were several occasions where the internet was disrupted during CAS completion, preventing the participants from completing the CAS on their own. As a result, study staff asked these participants the remainder of CAS questions, in order to collect complete CAS information.

Participant Compensation. Participants were compensated 20 dollars for completing Phase 1 and those eligible who completed Phase 2 received an additional 10 dollars for their time and effort.

Measures

Assessment measures for this proposal were carefully selected, based on domains to be studied, psychometric properties of existing measures, and available resources. In addition, total time for study participation was taken into consideration, with priority given to the primary aims of the study. For alcohol, many standardized measures were available and the proposed study focused on replication and extension of previous findings. For caffeine, in contrast, standardized measures were often unavailable, and prompted the use of alcohol measures modified to focus instead on caffeine consumption. In addition, personal experiences from preliminary studies with the target population, also guided development of the assessment battery. Assessment measures can be found in the Appendix.

Phase 1: Quantitative (Survey and Timeline Follow-Back).

***Computer-Assisted Survey (CAS) Measures*.** The CAS was administered through REDCap and took between 10-20 minutes to complete. While the primary focus of the survey was on caffeine, alcohol, and CAC use, it also included questions about demographics, other substance use, mood (depression), problems related to substance use, ADHD, family history, and impulsivity.

Demographics. Items included participant age, gender, race and ethnicity, education level, student status (full or part time), current employment (full time, part time, or none), and fraternity/sorority membership.

Substance Use Questions.

Recent Alcohol Use (Past 30 days). To determine recent quantity and frequency of alcohol use, participants were asked to estimate: total number of drinking days; average number of drinks consumed on drinking days; and total number of days on which 5 or more drinks for men (4 for women) were consumed. These measures are based on the QF measure by Cahalan and Cisin (1968) and have shown high reliability across heterogeneous subject groups (Babor, Stevens and Marlatt, 1987). Consistent with the QF measure, participants were also given a handout with information about what constitutes one standard drink (12oz beer; 5oz wine; shot of liquor; see Figure 3) (Fishburne & Brown, n.d.).

Figure 3. Standard Drink Card.

Recent Caffeine use (Past 30 Days).

Overall Quantity and Frequency Measure: Total amount of caffeine consumed (past 30 days) was measured using questions analogous to those described for alcohol. Participants were asked to estimate the number of days they drank at least one caffeinated beverage (range 0-30) and the average number of beverages they consumed on caffeine use days (across all caffeine sources). Additionally, participants were asked to estimate caffeine servings (1 serving = 1, 8oz cup of regular coffee) from a typical day in the past 30 days.

Beverage-Specific QF Measure: Since caffeine is found in so many products and the amount of caffeine (mgs) varies widely, participants were also asked to estimate their frequency of use (days per week on a typical week in past month) separately for the following beverage types: coffee, tea, sodas and energy drinks/shots. In addition, they were asked to estimate the average number of beverages they consumed on the days they had caffeine. For these estimates, they had the chance to designate the unit of measure.

Overall Use of Caffeine and Alcohol Combined (CAC) (Recent– Past 30 Days). Participants were asked to estimate their frequency of CAC use (number of days in the past month). In addition, they were asked to estimate the average number of beverages they consumed on the days they consumed at least one CAC beverage.

Caffeine and alcohol combined (CAC) (Past Year). Two structured questions assessed past year quantity and frequency of CAC use.

Alcohol Problems. Alcohol problems were assessed using the 3-item Alcohol Use Disorders Identification Test Consumption screening tool (AUDIT-C) which measures quantity and frequency of alcohol use (past year) as well as heavy/binge drinking (6+ drinks). This shortened version of the 10-item Alcohol Use Disorders Identification Test (AUDIT) identifies

persons at risk for hazardous/problem drinking and can be administered via electronic format (Graham et al., 2007). It was found to have good reliability across a variety of settings with different populations (Bush et al., 1998, Bradley et al., 2007, Dawson et al., 2006, Reinert et al., 2002). AUDIT-C scores range from 0-12, with men considered positive at a score of $\geq$ 4 and women at a score of $\geq$ 3. Participants were also asked about lifetime and past 30 days blackouts, derived from the blackout item from the TWEAK screener, a recommended screener for risk drinking in this population (Russell, 1994; NIAAA, 2005).

Other Drug Use. Participants were asked to estimate number of days they used different types of drugs in the past 30 days.

Substance Use Problems and Treatment History. Participants were asked about lifetime history of substance abuse treatment, including AA/NA and whether they ever had problems due to their use of alcohol or other drugs.

Nicotine Use. Participants were asked whether or not they currently smoke cigarettes, using formatting from the Fagerstrom Test for Nicotine Dependence (Heatherton et al., 1991). Additionally, participants were asked to estimate the number of days they used e-cigarettes/personal vaporizers in the past 30 days.

Prescribed Medications. Participants were asked if they were taking medications that were prescribed to them and whether these medications were psychiatric or non-psychiatric.

ADHD. Symptoms of ADHD were assessed using two questions about inattention and hyperactive symptoms developed by Dr. Heather Jones. Participants were also asked about diagnosis of ADHD, with response options: "Yes, I have been diagnosed with ADHD or ADD"; "I've never been diagnosed with ADHD or ADD, but I think I might have ADHD or ADD"; and "No, I've never been diagnosed with ADHD or ADD, and I don't think I have ADHD or ADD".

Depression. Recent depression (past 2 weeks) was assessed using the Patient Health Questionnaire-2 (PHQ-2). This brief, self-administered depression screen consists of two items focused on symptoms of depression. The PHQ-2 is composed of the first to items of the better known PHQ-9. The PHQ-2 has been shown to be both reliable and valid (Kroenke, Spitzer, & Williams, 2003; Zhang et al., 2013). Items ask about symptoms of depression over the past two weeks. The full questionnaire will be scored based on total points, ranging from 0 to 6, with a cutoff score of 3 for indication of possible depression.

Family history. For first degree biological relatives (i.e., mother, father, brothers, and sisters), participants were asked if they think the family member ever had a problem with alcohol or other drugs. Response options will include yes, no, and don't know or not applicable. Items were based on the Spit for Science survey and reflect those asked in the family history module of the Addiction Severity Index (ASI) (McClellan et al., 1992).

Impulsivity. The eight-item Barratt Impulsiveness Scale–Brief (Steinberg et al., 2013) is based on the 30-item Barratt Impulsivity Scale (BIS; Barratt, 1959). Responses are scored on a four-point scale (1 = rarely/never, 2 = occasionally, 3 = often, 4 = almost always/always) and higher scores indicate greater impulsivity. This measure has been found to be reliable and valid (Steinberg et al., 2013; Fields et al., 2015).

Personal Interview: Timeline Follow-Back (TLFB). For the purposes of this study, a TLFB was used to assess caffeine, alcohol, and other drug use in the past 30 days (in that order). The 30-day time period was selected because: a) it is short enough to keep study participation to a reasonable period of time (e.g. participant time, study time), b) it directly maps onto the timeframe of the survey questions of interest (i.e., questions about use over the course of the past month), and c) it is congruent with published research and other substance use measures.

Caffeine. Caffeine was the first substance assessed using TLFB procedures. To the best of our knowledge, there are no standardized methods for collecting caffeine use data using TLFB interview guidelines. Study procedures relied upon those used previously to collect other substance use data in a reliable and valid way. The number, type, and volume of caffeinated beverage were elicited in order to accurately calculate both quantity and frequency of use.

Use/Abstinence. For each day, participants were first be asked what caffeinated beverages they consumed.

Number of servings. The participants were asked how many of each listed beverage they consumed.

Type. The participants were asked about the beverage classification across liquid categories (including brand and how prepared, if known). Since a goal of this study was to examine potential misconceptions based on existing knowledge, if participants asked research staff about caffeine content in beverages, the research staff responded by informing them that they could not provide any information until after they finished the study. At that time, however, the researcher would be happy to answer questions. Until then, however, participants were instructed to provide their best guess about the caffeine content of different beverages.

Volume. Participants were asked how much of the beverage they drank in terms of fluid ounces, through the use of either direct report of ounces or elicitation of container used (e.g., was it in a mug, in a soda can, etc.).

To avoid confounding of subsequent reporting, participants did not receive any information about what did/did not contain caffeine during this assessment. Instead, the researcher recorded all items a participant described as containing caffeine (e.g., ginger ale).

Information obtained about non-caffeinated beverages were summarized to reflect the overall number of misconceptions as well as the total number of drinks according to the participant's estimate. Caffeine use frequency and quantities were calculated; for data analysis, quantities were converted to milligrams, which is consistent with previous literature.

Alcohol. For alcohol, the number, type, and volume of alcoholic drinks consumed were recorded and then converted into standard drink units (SDUs).

Use/Abstinence. For each day, participants were first be asked what alcoholic beverages they consumed.

Number. Next, participants reported how many of each listed beverage they consumed.

Type. The participants were asked about the beverage category, which includes beer, liquor, and wine.

Volume. The participants were asked how much of the beverage they drank in terms of fluid ounces, through the use of either direct report of ounces or elicitation of container used (e.g., shot glass, solo cup, etc.). For reported mixed drinks, participants were asked to approximate the amount of liquor in each drink (e.g., one shot).

Combined Use of Alcohol and Caffeine. After collecting caffeine use information and while collecting the alcohol use data, participants were asked about type of alcohol consumed (beer, wine, liquor) and whether it was mixed with other beverages. Asking participants about mixers served two purposes. First, it enabled both general information on combined consumption of alcohol and caffeine to be obtained, while masking intent to ask about combined use. Second, this was a way to obtain information on caffeinated drinks being consumed that the participants either assumed didn't contain caffeine or they didn't consider because it was mixed with alcohol; this was determined as the caffeine use TLFB was completed before asking about combined use.

Participants were also asked if their listed caffeine and alcohol consumption occurred at the same/different points in the day in an effort to thoroughly address any potential mixing.

Use of the TLFB might not have captured all caffeine use misconceptions. It is reasonable to assume that asking participants to list the types of caffeine used is the best procedure to capture the misconceptions associated with drinks they believe to be caffeinated. However, assessing their use of caffeinated beverages that they don't realize contain caffeine is not as straightforward. The researcher asked participants what they mixed with alcohol. This procedure enabled obtaining information on the caffeinated drinks they consumed that they did not know contained caffeine and thus didn't previously report; this process additionally helped obtain general information on mixing of alcohol and caffeine for survey comparison purposes.

Phase 2: Qualitative Interview. The final portion of the visit consisted of a qualitative interview. This served the dual purposes of clearing up/double-checking inconsistencies between the TLFB and survey responses and understanding individual perceptions about alcohol and caffeine use questions in Phase 1. The Phase 2 interview contributed to data accuracy and will help guide future survey item development. Phase 2 also avoided the need to rely exclusively on TLFB data, and allowed participants to remember more accurately when and how much caffeine and alcohol they consumed in past 30 days. The addition of a qualitative component also enhanced the design richness in an attempt to better capture the full picture (Yoshikawa et al., 2008).

Participants provided informed consent in order to participate in the qualitative interview portion (see Informed Consent section). Prior to the start of the qualitative interview, participants had a 3 to 5 minute break while data was compiled and reviewed to identify potential discrepancies in preparation for the interview. The qualitative interview used participants' TLFB

and survey results to inform the process, serving as a guide to the specific questions that were asked. Specifically, participants were asked about any discrepancies, how they came to the survey answers that were discrepant with the TLFB, thoughts about what these survey questions meant, etc. Additional Phase 2 topics discussed with participants included whether or not they received any recent information about substance use that could potentially have created noise in the data collected, including if they had any discussion/intervention regarding substance use with their medical provider during the visit or at any other point recently; their opinion about blackout definition versus term; and if they mixed prescription medication with alcohol. Participants were also asked both how honest and accurate they were in responding to the survey and interview.

This semi-structured interview format allowed for format modification, participant follow up, and detail consideration (Nelson & Quintana, 2005; Madill & Gough, 2008); each qualitative interview adhered to a general outline, but were tailored based on participant-specific responses. The inclusion of this interview also served to begin explorations that will help to inform the next phase of research in this area of research.

Methodological Considerations

Computerized Assessment. Research using computer-based data collection procedures is becoming increasingly common and the integration of computerized assessment was an important addition to the current study for several reasons. Computerized assessment can reduce scoring errors and ensure more reliable administration of the measure (e.g., no skipping). Each completed TLFB was directly entered into an electronic database via tablet computer. Using computerized assessment allowed for the accurate and time-efficient compilation of the information that was necessary for comparison with survey information. This comparison allowed for identification of discrepancies between TLFB and survey. By incorporating direct

data entry via tablet computer, noise in the data (e.g., data entry error) was potentially decreased. This method also increased the saving of resources, reducing the time needed to recruit the target number of study participants.

Data Analysis Plan

Order Effects. Independent *t*-tests were used to determine if differences existed between randomization groups for caffeine, alcohol, and CAC quantity and frequency. Dependent variables in these analyses were past 30-day TLFB and CAS frequency and quantity of caffeine, alcohol, and CAC use.

Aim 1. Assess descriptive and associative features of quantity and frequency of alcohol, caffeine, and CAC use.

Descriptive statistics were used to examine overall characteristics of alcohol, caffeine, and caffeine combined with alcohol use. Pearson product-moment correlation coefficients were used to characterize the relationship between alcohol, caffeine, and CAC use and other continuous variables in the study. These other variables included demographic information, impulsivity score, AUDIT-C score, and PHQ-2 score. Independent *t*-tests were used to determine if differences existed between alcohol, caffeine and CAC use and other categorical variables (gender, smoke cigarettes, used drugs, and family history).

Aim 2. Examine the agreement between survey and TLFB methods, identifying areas where there is lack of agreement.

To assess if lack of knowledge about what contains caffeine was responsible for some reporting errors, overall number and proportion of caffeine errors were examined. Additionally, the number of people with caffeine errors was calculated.

To determine if discrepancies existed between survey and TLFB reports, we compared caffeine and alcohol eligibility criteria, TLFB and survey results and calculated the number of inconsistencies for caffeine, alcohol and CAC use. For each type of substance, we compared CAS and TLFB quantity and frequency of use. We also identified total number of inconsistencies, magnitude of these differences, and direction of the inconsistency (i.e., CAS under or overestimate compared to TLFB). This was done both per participant and overall across caffeine, alcohol, and CAC use. Associations between these lack of agreement variables and other categorical and continuous variables were also examined. Individuals with missing data for specific questions were not included in applicable analyses.

Aim 3. Examine qualitative personal interview data and identify themes and patterns with respect to inconsistencies, in particular focusing on discrepancies and misconceptions.

Qualitative interview data were examined to identify patterns of responses and common themes, thereby contributing to the construct validity of the study (Nelson & Quintana, 2005). This preliminary review of qualitative interview data was used to augment quantitative findings and better inform future research on caffeine, alcohol and CAC use in college students.

Results

Recruitment and Enrollment

A total of 232 students were offered screening for the EACH study at the VCU Student Health Center on the Monroe Campus. As summarized in Figure 1, *N* = 91.38% of those made aware of the study proceeded to the study screen. To determine study eligibility, those interested were screened for past month regular alcohol (once per week) and caffeine (twice per week) use using open-ended interview-format questions (see Method Figure 2. EACH recruitment script). Among those screened, 50.47% met eligibility criteria (endorsed past month regular alcohol and caffeine use) and were informed about the study. Of these, 46.73% provided verbal consent to Phase 1 and were randomized into either TLFB followed by CAS (46%) or CAS followed by TLFB (54%). All persons randomized completed Phase 1. Of these, *N* = 29 (58%) reported CAC use in Phase 1 and were eligible for Phase 2. Over half of those eligible (*N* = 16) consented to Phase 2 and completed the qualitative interview.

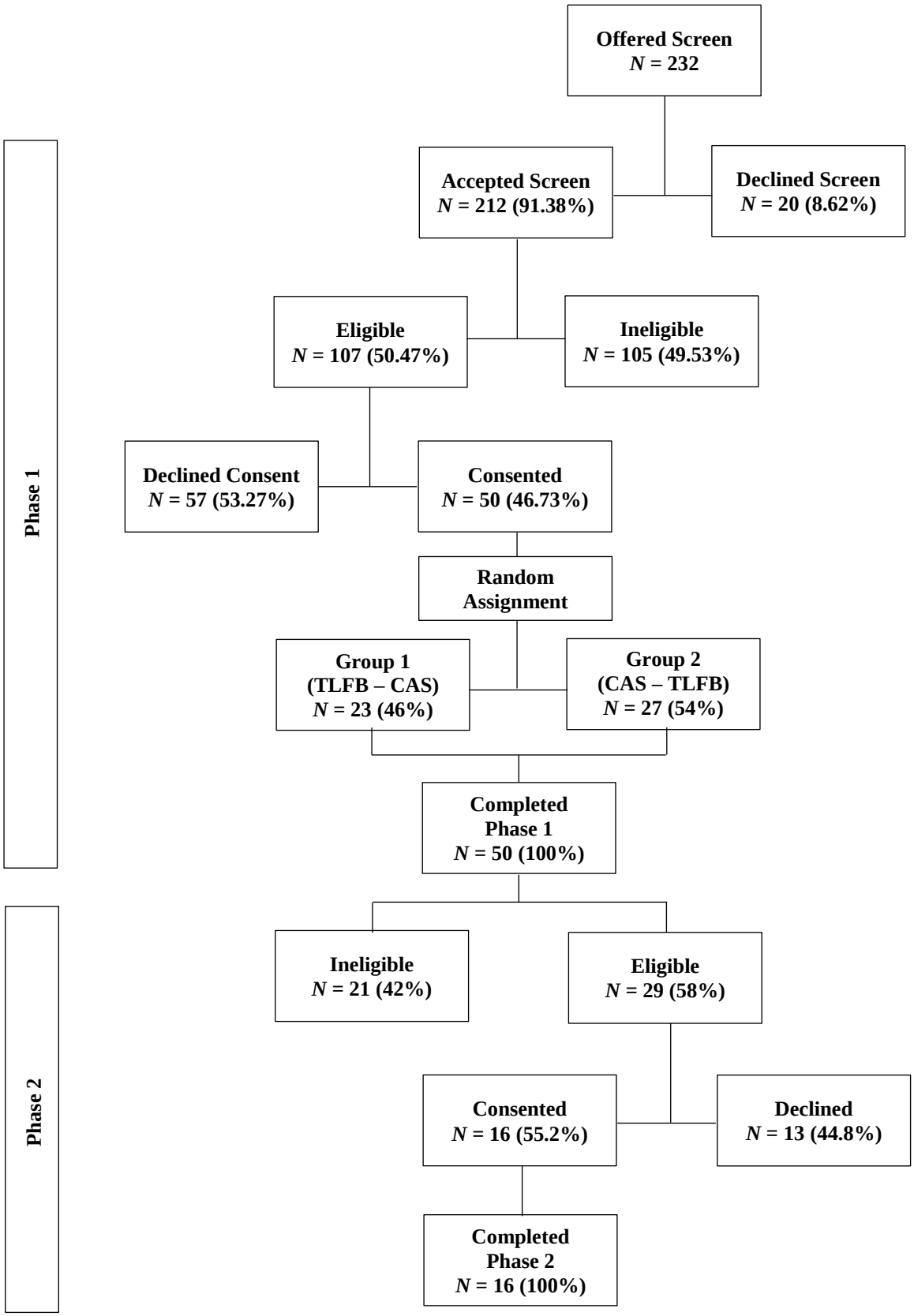

Figure 4. EACH recruitment and enrollment consort diagram.

Sample Demographics

Demographically, nearly all participants (98%) were full-time VCU students, and mean number of years of education was 14.3 (*SD* = 1.7). Less than one-tenth (8%) belonged to a fraternity or sorority. Mean age was 21.2 years (*SD* = 3.4) and almost three-fourths (72%) were female. The sample was predominantly White (48%), followed by Black/African American (22%), Asian/Pacific Islander (14%), Hispanic (8%), Mixed (6%), and Don't Know (2%).

Order Effects

To examine whether the order of CAS and TLFB administration influenced participant responding, students randomized to complete TLFB followed by CAS (Group 1) were compared to those completing CAS followed by TLFB (Group 2) for quantity and frequency of caffeine, alcohol and CAC use. TLFB frequency of alcohol use differed significantly by randomization group, $t(48) = -2.09$, $p = .04$. None of the other *t*-test comparisons of difference scores for quantity and frequency of caffeine, alcohol and CAC use by randomization group were significant. Findings are summarized in Table 1.

Table 11. Quantity and frequency of caffeine, alcohol, and CAC use by randomization group.

Substance Type	**Specific QF Variable**		**Mean (SD)**		***t*-value (*p*-value)**
			Group 1 (*N* = 23)	**Group 2 (*N* = 27)**	
Caffeine	Frequency (Days of Use)	TLFB	20.43 (10.68)	21.37 (9.15)	-.33 (.74)
		CAS	19.87 (9.70)	22.78 (7.79)	-1.18 (.25)
	Quantity (# Drinks on Typical Day)	TLFB	1.59 (1.02)	2.88 (3.89)	-1.66 (.11)
		CAS	1.35 (.82)	2.43 (2.89)	-1.73 (.09)
Alcohol	Frequency (Days of Use)	TLFB	5.57 (4.17)	9.07 (7.06)	-2.09 (.04)*
		CAS	7.24 (5.44)	10.33 (7.10)	-1.71 (.10)
	Quantity (# Drinks on Typical Day)	TLFB	4.65 (2.30)	5.52 (5.39)	-.72 (.47)
		CAS	4.07 (2.89)	3.74 (2.71)	.41 (.68)
CAC	Frequency (Days of Use)	TLFB	.35 (.78)	1.52 (5.37)	-1.04 (.31)
		CAS	3.52 (6.82)	4.67 (8.46)	-.52 (.61)
	Quantity (# Drinks on Typical Day)	TLFB	.63 (1.61)	1.89 (5.17)	-1.13 (.27)
		CAS	1.30 (1.66)	1.78 (2.52)	-.77 (.45)

*Denotes a statistically significant *t*-value ($p < .05$).

Aim 1. Assess descriptive and associative features of quantity and frequency of alcohol, caffeine, and CAC use

Caffeine Use. Recent (past month) caffeine use prevalence by CAS and TLFB are summarized in Table 2. As required for study eligibility, all participants reported some recent caffeine use. When CAS and TLFB methods were compared, similar prevalence rates were found; over three-fourths reported coffee consumption (78% and 80%), approximately half reported tea (54% and 42%) and soda (56% and 46%), and around one-fourth reported ED/shots (28% and 20%). For all four caffeine beverage types, percent of participants categorized as "users" did not differ by method of administration (CAS or TLFB) (All McNemar's tests $p > .05$).

Table 12. Recent caffeine use prevalence by CAS and TLFB.

Caffeine Category	**Number (%) of Users**		**χ^2 *p*-Value**
	CAS	**TLFB**	
Overall	50 (100%)	50 (100%)	N/A
Coffee	39 (78%)	40 (80%)	1.00
Caffeinated Tea	27 (54%)	21 (42%)	.07
Caffeinated Soda	28 (56%)	23 (46%)	.23
Caffeinated Energy Drink/Shot	14 (28%)	10 (20%)	.34

As shown in Table 3, on both CAS and TLFB, participants consumed caffeine on about 2 out of every 3 days, around 2 drinks per day on the days they used, with total consumption (number of beverages) estimated between 45-47 drinks. Greater variability was found for estimated number of servings per day (with one serving equivalent to 1, 8oz cup of regular coffee), with mean number of servings and total servings consumed over twice as high on the CAS compared to the TLFB.

When individual beverage types were examined separately, CAS means were higher than TLFB means on all but one variable, quantity of caffeinated tea. Looking across the 2 methods, most coffee use occurred about 4-5 days per typical week. However, the number of caffeinated tea, soda and ED use days in a typical week were at least twice as high on the CAS compared to the TLFB.

Table 13. Quantity and frequency of recent caffeine use by beverage type and research administration method.

Caffeine Beverage Type	Caffeine Use Domain	Mean (SD) of Users	
		CAS	TLFB
Total	Frequency (Days of Use)	21.44 (8.75)	20.94 (9.79)
	Quantity (# Drinks on Typical Day)	1.93 (2.24)	2.29 (2.99)
	Servings (Typical Day)	3.39 (6.72)	1.56 (1.58)
	Mg (Typical Day)	Not Assessed	156.39 (158.01)
	Heavy Use Days (≥ 500 mg)	Not Assessed	1.26 (5.16)
	Total Drinks Consumed (Past Month)*	46.95 (67.78)	45.06 (70.94)
	Total Servings Consumed (Past Month)*	76.97 (166.44)	37.59 (50.03)
Coffee	Number of Use Days (Typical Week)	5.18 (4.60)	3.93 (2.44)
	Number of Drinks (Typical Day)	1.35 (.80)	1.30 (.59)
Caffeinated Tea	Number of Use Days (Typical Week)	6.67 (9.04)	2.71 (2.71)
	Number of Drinks (Typical Day)	1.61 (1.39)	1.89 (2.68)
Caffeinated Soda	Number of Use Days (Typical Week)	4.21 (4.43)	2.20 (2.07)
	Number of Drinks (Typical Day)	1.59 (.99)	1.33 (.77)
Caffeinated Energy Drink/Shot	Number of Use Days (Typical Week)	3.71 (2.92)	.75 (.61)
	Number of Drinks (Typical Day)	1.21 (.58)	1.08 (.18)

*Product of past month caffeine frequency and quantity.

Alcohol use. As required for study eligibility, all participants reported some recent (past month) alcohol use. Table 4 displays quantity and frequency of recent alcohol use and problems for both administration methods. Mean days of alcohol use ranged from 7.46 (TLFB) to 8.91 (CAS). Past month mean number of drinks on a typical drinking day varied from 3.89 (CAS) to 5.12 (TLFB), with past month total drinks ranging from 40.20 (CAS) to 49.63 (TLFB).

Table 14. Quantity and frequency of recent alcohol use and problems by administration method.

Alcohol User Measure	**Mean (SD)**	
	CAS	**TLFB**
Frequency (Days of Use)	8.91 (6.52)	7.46 (6.11)
Quantity (# Drinks on Typical Day)	3.89 (2.77)	5.12 (4.24)
Total Drinks Consumed (Past Month)*	40.20 (62.78)	49.63 (121.45)
Number of Blackouts	.93 (1.44)	.15 (.42)

*Product of past month alcohol frequency and quantity.

As shown in Table 5, for both the CAS and TLFB, approximately four-fifths (80% and 78%, respectively) of participants reported at least one binge drinking day in the past 30 days. However, participants were twice as likely to report blackouts on the CAS (26%) as compared to the TLFB (13%).

Table 15. Recent (past 30 days) indicators of risky alcohol use by CAS and TLFB.

Alcohol User Domain	Number (%) of Users	
	CAS	**TLFB**
Binge Drinking* (at least one binge day)	40 (80%)	39 (78%)
Recent Blackout(s)	13 (26%)	6 (13%)

* Binge is defined as 4 standard drink units/drinking occasion for women and 5 standard drink units/drinking occasion for men.

CAC use. As shown in Table 6, CAC use was reported by almost half (48%) of participants by CAS and less than one-third (30%) by TLFB. The majority of CAC mixers were caffeinated soda (66.67%), followed by ED (20%), caffeinated tea (20%), and finally, coffee (6.67%). A McNemar's test determined no differences existed in number of CAC users based on CAS and TLFB.

Table 16. Prevalence and types of CAC use by administration method.

CAC Category	Number (%) of Users	
	CAS	**TLFB**
Overall*	24 (48%)	15 (30%)
Coffee	Not Assessed	1 (6.67% of CAC users)
Caffeinated Tea	Not Assessed	3 (20% of CAC users)
Caffeinated Soda	Not Assessed	10 (66.67% of CAC users)
Caffeinated Energy Drink/Shot	Not Assessed	3 (20% of CAC users)

*McNemar's test nonsignificant (p = .06).

Table 7 summarizes quantity and frequency of recent CAC use by administration method. Reports on the CAS and TLFB varied, with participants reporting more than twice as many days of CAC use by CAS (8.63) as compared to TLFB (3.27). Quantity of CAC use (number of drinks per occasion) reports were more similar, ranging from 3.25 (CAS) to 4.37 (TLFB). Mean total CAC drinks consumed (past month) was almost twice as high by TLFB (52.53) than by CAS (30.46).

Table 17. Quantity and frequency of recent CAC use by administration method.

CAC Category	CAC Use Domain	Mean (*SD*)	
		CAS	**TLFB**
Overall (out of CAC Users)	Frequency (Days of Use)	8.63 (9.26)	3.27 (6.90)
	Number of CACs (Typical Day)	3.25 (2.05)	4.37 (6.40)
	Mg (Typical Day)	Not Assessed	122.78 (235.65)
	Total Drinks Consumed (Past Month)*	30.46 (38.03)	52.53 (185.23)
Coffee (out of Coffee Users) *N* = 1	Number of Use Days	Not Assessed	27.00 (N/A)
	Number of Drinks	Not Assessed	2.07 (N/A)
Caffeinated Tea (out of Tea Users) *N* = 3	Number of Use Days	Not Assessed	9.33 (14.43)
	Number of Drinks	Not Assessed	3.56 (4.44)
Caffeinated Soda (out of Soda Users) *N* = 10	Number of Use Days	Not Assessed	1.60 (1.08)
	Number of Drinks	Not Assessed	1.10 (.32)
Caffeinated ED/Shot (out of ED/shot users) *N* = 3	Number of Use Days	Not Assessed	1.00 (0.00)
	Number of Drinks	Not Assessed	1.00 (0.00)

*Product of past month CAC frequency and quantity.

Other Substance Use and Related Problems.

Substance Use Variables. Based on CAS data, almost one-fifth (18%) of participants categorized themselves as current smokers, with one-fifth (20%) reporting recent use of e-cigarettes/personal vaporizers (mean = 7.70; *SD* = 8.96). Also, 14% of the sample used e-cigarettes/personal vaporizers to consume marijuana (mean = 7.00; *SD* = 4.20).

Recent use of one or more drugs was reported by over half of study participants (60% by CAS and 56% by TLFB) and did not differ for the two administration methods ($p = .63$). Rates of self-reported use of other drugs (past 30 days) are summarized in Table 8. For specific drugs, with the exception of cocaine, more students reported use on the CAS as compared to the TLFB.

Table 18. Recent other substance use prevalence by CAS and TLFB.

Other Substance Use Variables	**Number (%) of Users**	
	CAS	**TLFB**
Any Drug*	30 (60%)	28 (56%)
Cannabis	24 (48%)	23 (46%)
Sedatives	5 (10%)	1 (2%)
Stimulants	7 (14%)	5 (10%)
Cocaine	3 (6%)	3 (6%)
Opioids	3 (6%)	1 (2%)
Other Drugs	0 (0%)	0 (0%)

*McNemar's test nonsignificant ($p = .63$).

Alcohol/Drug Problems. The mean AUDIT-C score was 4.92 (*SD* = 2.21) with over four-fifths (84%) of participants screening at risk for problematic alcohol use (score of 4 or more for men and 3 or more for women). For specific AUDIT-C items, the mean scores were 2.46, (with 2 = "2-4 times a month" and 3 = "2-3 times a week"; *SD* = .84) for "How often do you have a drink containing alcohol?"; 1.02 (with 1 = "3 or 4"; *SD* = .94) for "How many standard drinks containing alcohol do you have on a typical day?"; and 1.44 (with 1 = "less than monthly" and 2 = "monthly"; *SD* = 1.01) for "How often do you have six or more drinks on one occasion?". Over one-tenth (14%) reported lifetime substance use problems, only 4% of the sample indicated that they have received substance abuse treatment, and over one-third (36%) reported alcohol or drug problems in at least one first degree relative.

Mental Health Variables. As displayed on Table 9, the mean PHQ-2 score was 1.56 (*SD* = 1.49) and 16% of participants scored at or above the clinical cutoff for depression (score of 3 or above). The mean BIS score was 15.50 (*SD* = 3.72), with almost one-fourth (24%) of participants reporting use of at least one prescribed psychiatric medication.

Over two-fifths (42%) of the sample reported possible ADHD; with 16% stating they had been diagnosed with the disorder and an additional 26% reporting no diagnosis, but the belief they have the disorder. Over two-fifths (44%) had at least one of these ADHD symptoms and over one-fifth (22%) endorsed both symptoms. Almost one-third (32%) of participants said they make careless mistakes, have trouble keeping their attention focused, or have difficulty organizing/planning most of the time. Over one-third (34%) reported that they interrupt other people when they are talking, talk a lot, or feel like they have a lot of energy most of the time.

Table 19. Prevalence and means (SDs) of mental health variables.

Mental Health Measure		***N* (%)**
ADHD Variables	Make careless mistakes, have trouble keeping your attention focused, or have difficulty organizing/planning	16 (32%)
	Interrupt other people when they are talking, talk a lot, or feel like they have a lot of energy	17 (34%)
	At least one ADHD symptom	22 (44%)
	Both ADHD symptoms	11 (22%)
	No ADHD diagnosis, but believe they have it	13 (26%)
	Diagnosis of ADHD from health care professional	8 (16%)
PHQ-2 Score $\geq$ 3		8 (16%)
		Mean (SD)
PHQ-2 Score		1.56 (1.49)
BIS score		15.50 (3.72)

Caffeine and Other Variables. Correlations between CAS and TLFB quantity and frequency variables and other continuous measures are summarized in Table 10. TLFB caffeine use frequency increased with age (r = .34; p = .02) and BIS score (r = .33; p = .02). No other significant correlations were found.

Table 20. Demographic, alcohol and mental health variables and measures of caffeine use.

Other Variables	**Caffeine Frequency *r* (*p*-value)**		**Caffeine Quantity *r* (*p*-value)**	
	CAS	**TLFB**	**CAS**	**TLFB**
Age	.24 (.098)	.34 (.02)*	.03 (.85)	.02 (.91)
Education Completed	.09 (.52)	.25 (.09)	.12 (.43)	.03 (.85)
PHQ-2 Score	.17 (.24)	.16 (.27)	.11 (.44)	-.06 (.66)
AUDIT-C Score	.01 (.96)	-.03 (.85)	.21 (.15)	.27 (.06)
BIS Score	.26 (.07)	.33 (.02)*	.08 (.60)	-.06 (.70)

*Denotes a statistically significant correlation coefficient ($p < .05$).

Table 11 summarizes *t*-test comparisons of caffeine quantity and frequency measures for categorical variables, such as gender and current smoker/non-smoker. All were non-significant except for gender, with females reporting greater frequency of recent caffeine use (23.33 days) than males (16.57 days), $t(48) = 2.59$, $p = .01$. No other significant differences were found.

Table 21. Caffeine use by demographic and other substance use risk groups

Other Variables	**Caffeine QF Variables**		**Mean (SD)**		***t* (*p*-value)**
Gender	Caffeine Frequency (Days)	CAS	**Male** (*N* = 14)	**Female** (*N* = 36)	2.59 (.01)*
			16.57 (10.55)	23.33 (7.26)	
		TLFB	16.21 (11.93)	22.78 (8.30)	1.89 (.08)
	Caffeine Quantity (Drinks/Typical Day)	CAS	1.89 (2.37)	1.94 (2.23)	.07 (.94)
		TLFB	2.64 (4.12)	2.15 (2.47)	-.51 (.61)
Current Smoker (CAS)	Caffeine Frequency (Days)	CAS	**Non-smoker** (*N* = 41)	**Smoker** (*N* = 8)	-.34 (.73)
			21.07 (8.84)	22.25 (8.80)	
		TLFB	20.54 (10.01)	23.38 (9.46)	-.74 (.46)
	Caffeine Quantity (Drinks/Typical Day)	CAS	1.93 (2.45)	1.94 (.94)	-.01 (.99)
		TLFB	2.19 (3.25)	2.94 (1.28)	-.64 (.52)
Other Drug Use (TLFB or CAS)	Caffeine Frequency (Days)	CAS	**Non-user** (*N* = 19)	**User** (*N* = 31)	-.18 (.86)
			21.16 (9.67)	21.61 (8.30)	
		TLFB	20.63 (10.63)	21.13 (9.41)	-.17 (.86)
	Caffeine Quantity (Drinks/Typical Day)	CAS	1.45 (.96)	2.23 (2.73)	-1.20 (.24)
		TLFB	2.42 (3.24)	2.21 (2.90)	.24 (.81)
Family History of Alcohol/ Drug Problems	Caffeine Frequency (Days)	CAS	**FH Negative** (*N* = 32)	**FH Positive** (*N* = 18)	1.01 (.32)
			22.38 (8.27)	19.78 (9.56)	
		TLFB	21.84 (9.77)	19.33 (9.90)	.87 (.39)
	Caffeine Quantity (Drinks/Typical Day)	CAS	1.88 (2.32)	2.03 (2.16)	-.23 (.82)
		TLFB	2.29 (2.63)	2.29 (3.62)	-.01 (.99)
ADHD Diagnosis by practitioner or self-report	Caffeine Frequency (Days)	CAS	**No ADHD Diagnosis** (*N* = 29)	**ADHD Dx+** (*N* = 21)	-.27 (.79)
			21.72 (9.31)	21.05 (8.12)	
		TLFB	19.97 (10.91)	22.29 (8.05)	.87 (.39)
	Caffeine Quantity (Drinks/Typical Day)	CAS	2.21 (2.87)	1.55 (.74)	-1.03 (.31)
		TLFB	2.69 (3.79)	1.73 (1.12)	-1.13 (.26)

*Denotes a statistically significant *t*-value ($p < .05$).

Alcohol and Other Variables. As summarized in Table 12, for both CAS and TLFB, no significant relationships were found between alcohol quantity and frequency measures and other continuous variables.

Table 22. Correlations between demographic and mental health variables and measures of alcohol use.

Other Variables	**Alcohol Frequency *r* (*p*-value)**		**Alcohol Quantity *r* (*p*-value)**	
	CAS	**TLFB**	**CAS**	**TLFB**
Age	-.03 (.83)	.02 (.88)	-.20 (.17)	-.16 (.26)
Education completed	.13 (.36)	.07 (.62)	.04 (.81)	-.003 (.98)
PHQ-2 score	.07 (.63)	-.07 (.64)	-.04 (.77)	.01 (.94)
BIS Score	.20 (.16)	.12 (.40)	.18 (.22)	.18 (.21)

Table 13 summarizes *t*-test findings for alcohol and other categorical variables. Two significant differences were found. Participants indicating recent drug use reported more frequent use of alcohol on CAS than those who did not use drugs (mean = 10.42 and 6.45 days, respectively), $t(48) = -2.17$, $p = .04$. Also, males reported consuming more drinks per drinking day on CAS than females (mean 5.14 and 3.40, respectively), $t(48) = 2.06$, $p = .045$.

Table 23. Alcohol use by demographic and other substance use risk groups.

Other Variables	**Alcohol QF Variables**		**Mean (SD)**		***t* (*p*-value)**
Gender	Alcohol Frequency (Days)	CAS	**Male** (*N* = 14)	**Female** (*N* = 36)	.44 (.66)
			9.57 (7.01)	8.65 (6.40)	
		TLFB	9.36 (7.51)	6.72 (5.42)	1.38 (.17)
	Alcohol Quantity (Drinks/Typical Day)	CAS	5.14 (3.21)	3.40 (2.46)	2.06 (.045)*
		TLFB	7.82 (6.70)	4.07 (2.10)	2.06 (.06)
Current Smoker (CAS)	Alcohol Frequency (Days)	CAS	**Non-smoker** (*N* = 41)	**Smoker** (*N* = 8)	-.14 (.89)
			8.43 (6.03)	8.75 (4.98)	
		TLFB	7.32 (6.17)	6.25 (2.82)	.48 (.64)
	Alcohol Quantity (Drinks/Typical Day)	CAS	3.96 (2.99)	3.75 (1.49)	.20 (.85)
		TLFB	4.93 (4.58)	6.12 (2.16)	-.72 (.48)
Other Drug Use (TLFB or CAS)	Alcohol Frequency (Days)	CAS	**Non-user** (*N* = 19)	**User** (*N* = 31)	-2.17 (.04)*
			6.45 (5.22)	10.42 (6.84)	
		TLFB	5.53 (5.71)	8.65 (6.14)	-1.79 (.08)
	Alcohol Quantity (Drinks/Typical Day)	CAS	3.18 (1.50)	4.32 (3.27)	-1.42 (.16)
		TLFB	3.81 (2.12)	5.92 (4.99)	-1.75 (.09)
Family History of Alcohol/Drug Problems	Alcohol Frequency (Days)	CAS	**FH Negative** (*N* = 32)	**FH Positive** (*N* = 18)	-1.12 (.27)
			8.02 (4.89)	10.50 (8.65)	
		TLFB	6.34 (3.87)	9.44 (8.59)	-1.45 (.16)
	Alcohol Quantity (Drinks/Typical Day)	CAS	3.64 (2.60)	4.33 (3.09)	-.85 (.40)
		TLFB	4.41 (2.26)	6.38 (6.31)	-1.28 (.22)
ADHD Diagnosis by practitioner or self-report	Caffeine Frequency (Days)	CAS	**No ADHD Diagnosis** (*N* = 29)	**ADHD Dx+** (*N* = 21)	1.14 (.26)
			8.02 (5.92)	10.14 (7.23)	
		TLFB	7.34 (6.61)	7.62 (5.51)	.16 (.88)
	Caffeine Quantity (Drinks/Typical Day)	CAS	3.85 (2.79)	3.95 (2.82)	.13 (.89)
		TLFB	5.18 (5.40)	5.04 (1.78)	-.11 (.91)

*Denotes a statistically significant *t*-value ($p < .05$).

CAC and Other Variables. For both CAS and TLFB, none of the relationships between CAC quantity and frequency of use and other (continuous) variables were significant (see Table 14). Similarly, as shown in Table 15, no differences were found between other (categorical) variables and CAC quantity and frequency measures, regardless of administration method. Lastly, Table 16 shows that those who did and did not endorse CAC use by CAS and/or TLFB did not differ regarding ADHD variables.

Table 24. Correlations between CAC variables and other substance use risk groups.

Other Variables	**CAC Frequency *r* (*p*-value)**		**CAC Quantity *r* (*p*-value)**	
	CAS	**TLFB**	**CAS**	**TLFB**
Age	.05 (.75)	.02 (.92)	-.15 (.29)	.14 (.62)
Education Completed	.19 (.19)	.08 (.61)	-.07 (.62)	.03 (.92)
PHQ-2 Score	-.17 (.24)	.11 (.45)	-.21 (.14)	.02 (.94)
BIS Score	.23 (.12)	.15 (.32)	-.01 (.97)	.06 (.84)

Table 25. CAC use by demographic and other substance use risk groups.

Other Variables	**CAC QF Variables**		**Mean (SD)**		***t* (*p*-value)**
Gender	CAC Frequency (Days)	CAS	**Male** (*N* = 14)	**Female** (*N* = 36)	.12 (.90)
			4.36 (7.14)	4.06 (7.99)	
		TLFB	2.29 (7.42)	.47 (.94)	.91 (.38)
	CAC Quantity (Drinks/Typical Day)	CAS	2.43 (3.06)	1.22 (1.62)	1.40 (.18)
		TLFB	8.07 (10.52)	2.52 (1.74)	1.17 (.31)
Current Smoker (CAS)	CAC Frequency (Days)	CAS	**Non-smoker** (*N* = 41)	**Smoker** (*N* = 8)	-1.01 (.34)
			2.93 (5.43)	7.13 (11.46)	
		TLFB	1.10 (4.39)	.38 (.74)	.46 (.65)
	CAC Quantity (Drinks/Typical Day)	CAS	1.51 (2.19)	1.63 (2.20)	-.13 (.90)
		TLFB	4.35 (6.94)	6.08 (5.20)	-.33 (.75)
Other Drug Use (TLFB or CAS)	CAC Frequency (Days)	CAS	**Non-user** (*N* = 19)	**User** (*N* = 31)	.39 (.70)
			4.68 (8.78)	3.81 (7.08)	
		TLFB	.26 (.56)	1.42 (5.03)	-1.00 (.33)
	CAC Quantity (Drinks/Typical Day)	CAS	1.21 (1.72)	1.77 (2.39)	-.89 (.38)
		TLFB	2.10 (.60)	5.19 (7.37)	-.82 (.43)
Family History of Alcohol/Drug Problems	CAC Frequency (Days)	CAS	**FH Negative** (*N* = 32)	**FH Positive** (*N* = 18)	-1.05 (.30)
			3.28 (6.66)	5.67 (9.26)	
		TLFB	.53 (.98)	1.78 (6.56)	-.80 (.43)
	CAC Quantity (Drinks/Typical Day)	CAS	1.28 (1.84)	2.06 (2.62)	-1.22 (.23)
		TLFB	3.26 (2.87)	6.60 (10.74)	-.68 (.53)
ADHD Diagnosis by practitioner or self-report	Caffeine Frequency (Days)	CAS	**No ADHD Diagnosis** (*N* = 29)	**ADHD Dx+** (*N* = 21)	1.59 (.12)
			2.59 (5.88)	6.29 (9.40)	
		TLFB	1.24 (5.21)	.62 (.87)	-.54 (.59)
	Caffeine Quantity (Drinks/Typical Day)	CAS	1.52 (2.52)	1.62 (1.60)	.16 (.87)
		TLFB	1.22 (4.79)	1.44 (2.56)	.19 (.85)

*Denotes a statistically significant *t*-value ($p < .05$).

Aim 2. Examine the agreement between survey and TLFB methods, identifying areas where there is lack of agreement

Study Eligibility: Screening vs Assessment. Discrepancies were found between student responses to initial screening questions about regular caffeine and alcohol use and subsequent CAS and/or TLFB reports. Overall, more than two-fifths (44%) of participants enrolled in the study later responded to TLFB and/or CAS items such that they would no longer meet criteria set for regular use of one or both substances. Over one-third (36%) of participants fell below the cut off for either caffeine (at least twice a week) or alcohol (at least once a week) and 8% were below the criterion for both substances on one or both methods for data collection. As shown in Table 17, almost one-fifth (18%) of participants would not have met criteria for caffeine (regular use) and over one-third (34%) would not have met criteria for alcohol by the CAS and/or TLFB. One-tenth (10%) of participants would not have met caffeine criteria and over one-fifth (22%) would not have met alcohol criteria by both assessment methods. Finally, 6% of the sample would be ineligible by both CAS and TLFB data for both caffeine and alcohol.

Table 26. Sample eligibility based on interview and survey results.

Screening Criterion	Number of Participants Ineligible by CAS (% of *N*)	Number of Participants Ineligible by TLFB (% of *N*)	Number of Participants with at least one Ineligible Report (% of *N*)	Number of Participants Ineligible Based on CAS and TLFB (% of *N*)
Used caffeine regularly (average of 2 times per week) in the past 30 days	6 (12%)	8 (16%)	9 (18%)	5 (10%)
Used alcohol regularly (average of 1 time per week) in the past 30 days	11 (22%)	17 (34%)	17 (34%)	11 (22%)

TLFB Caffeine Inconsistencies. To determine rates of misreporting (e.g., non-caffeinated beverages reported as caffeinated ones) and inconsistent caffeine reporting (e.g., reporting a caffeinated beverage as part of a CAC consumption that was not previously reported on the caffeine TLFB), caffeine use reports on caffeine TLFB and CAC use on the alcohol TLFB were examined.

Caffeine Overall. Over half (56%) of participants had at least one caffeine misreport with almost half (46%) involving whether a particular product did/did not contain caffeine. When misreports were examined across the 30-day assessment window, among those with at least one caffeine misreport, the mean number of such misreports was 13.30 (*SD* = 33.82) or 3-4 per week. About one-fourth (26%) of participants reported on a non-caffeinated beverage that they believed to contain caffeine, such as ginger ale and Sierra Mist. The average number of such misreports

was 7.31 (*SD* = 9.54). In addition, 30% of participants classified a non-caffeinated beverage as one that contained caffeine. As shown on Table 18, beverage types most frequently associated with misconceptions (what does/does not have caffeine) were sodas, followed by tea, and finally other beverages.

Table 27. Beverage types most frequently associated with misconceptions (what does/does not have caffeine).

Beverage Type	Type of Misconception	Percentage
Soda	Both*	18 (52.9%)
Light Sodas (e.g., Sprite, Ginger Ale, Sierra Mist)	NC as C**	10 (29.4%)
Dark/Cola Sodas (e.g., Coca Cola, Dr. Pepper, Pepsi)	C as NC***	6 (17.6%)
Fruit Sodas (e.g., Orange soda, Fanta)	Both*	2 (5.9%)
Tea (e.g., Green Tea, Sweet Tea)	Both*	10 (29.4%)
Other (e.g., Sparkling Water, Chocolate Beverage, Fruit Punch)	Both*	6 (17.6%)

*Both = non-caffeinated beverage listed as caffeinated and caffeinated beverage listed as non-caffeinated.
**NC as C = non-caffeinated beverage listed as caffeinated.
***C as NC = caffeinated beverage listed as non-caffeinated.

CAC. Inconsistencies in CAC reporting were found among almost half (44.8%) of participants who reported CAC use on either the CAS and/or TLFB. Among those with inconsistencies, the mean number was 13.46 (*SD* = 41.94). In these cases, 92.3% reported using a caffeinated beverage in combination with alcohol on the alcohol TLFB that they had not previously noted when reporting caffeine use on the caffeine TLFB.

Discrepancies between CAS and TLFB reports. Table 19 summarizes *t*-test comparisons of CAS and TLFB measures of caffeine, alcohol and CAC use and problems. For

caffeine, no differences were found for quantity, frequency and servings/day. For alcohol, however, differences were found, with greater frequency of use reported by CAS as compared to TLFB ($t(48) = 3.24$, $p = .002$), but higher quantity reported by TLFB as compared to CAS ($t(48) = -2.76$, $p = .01$). For CAC, a difference was found only for frequency of use, with CAS reports exceeding those of TLFB ($t(48) = 3.15$, $p = .003$).

Table 28. Comparison of CAS and TLFB data for quantity and frequency of caffeine, alcohol and CAC use and problems.

Substance Use Domain		***t*-value (*p*-value)**
Caffeine	Frequency (Days of Use)	.65 (.52)
	Quantity (# Drinks on Typical Day)	-.85 (.40)
	Servings (Typical Day)	1.94 (.058)
Alcohol	Frequency (Days of Use)	3.24 (.002)*
	Quantity (# Drinks on Typical Day)	-2.76 (.01)*
	Number of Blackouts	1.77 (.14)
CAC	Frequency (Days of Use)	3.15 (.003)*
	Number of CACs (Typical Day)	-1.58 (.14)

*Denotes a statistically significant *t*-value ($p < .05$).

CAS and TLFB inconsistencies in reports of alcohol, caffeine and CAC use were examined in three ways. First, the number of inconsistencies was recorded and evaluated. In this case, any difference between the two assessment methods was tallied as an inconsistency, regardless of magnitude. For example, on a frequency of use question, a CAS-TLFB difference of one day counts the same as a 30 day inconsistency. This tally of CAS-TLFB inconsistencies was examined overall, as well as separately for caffeine, CAC, and alcohol use. Second, magnitude of CAS-TLFB differences were examined, looking at the absolute value of the difference between CAS and TLFB reports. Third, inconsistencies were characterized as over- or underestimates (for CAS compared to TLFB) and summarized. Overall, all participants had at least one discrepancy between CAS and TLFB reporting with a mean of 7.90 total discrepancies (*SD* = 2.77) across all 3 beverage types (caffeine, alcohol, and CAC).

Caffeine. As shown on Table 20, every participant had at least one caffeine discrepancy with mean number of discrepancies at 4.48 (*SD* = 1.84; range: 1-9). For frequency of use, about three-fourths (74%) of participants had at least one caffeine use discrepancy with the mean discrepancy magnitude being 4.62 (*SD* = 4.32) days. Just over half the sample (51.4%) had survey overestimates for frequency of caffeine use. Almost half (48%) had a discrepancy between CAS and TLFB quantity of caffeine use reports and 56% had a discrepancy regarding number of caffeine servings (100 mg) on a typical day. While mean discrepancy magnitude for number of drinks reported was 2.33 (*SD* = 3.65), the mean discrepancy magnitude for caffeine servings was 4.14 (*SD* = 8.48). Additionally, whereas half (50%) of the quantity of caffeine use discrepancies were survey overestimates, almost three-fourths (71.4%) of caffeine serving discrepancies were survey overestimates.

Discrepancies between reports on type of caffeinated beverage questions were also examined. In no case did magnitude of TLFB estimates exceed survey estimates, with survey overestimates ranging from 50% (quantity of caffeine and quantity of coffee use) to 100% (frequency of ED use). Of the 42 individuals with coffee discrepancies, 5 (11.9%) indicated use on only one of the two administration methods (CAS or TLFB). Over one-fourth (28.6%) of those with tea discrepancies reported such use on either the CAS or TLFB, but not both. More than one-third (35.5%) of those with caffeinated soda discrepancies endorsed use on only the CAS or TLFB. Over half (58.8%) of those with ED use discrepancies indicated use on only CAS or TLFB.

Table 29. CAS and TLFB discrepancy information for caffeine quantity and frequency questions.

Comparison Question	Number (%) with Discrepancies	Mean (SD) Number Discrepancies	Mean (SD) Magnitude of Discrepancies	Number (%) Survey Overestimates
Overall	50 (100%)	4.48 (1.84)	N/A	N/A
Frequency of Caffeine Use	37 (74%)	N/A	4.62 (4.32)	19 (51.4%)
Quantity of Caffeine Use	24 (48%)	N/A	2.33 (3.65)	12 (50%)
Frequency of Coffee Use (Typical Week)	16 (32%)	N/A	3.03 (5.32)	10 (62.5%)
Quantity Coffee Use	10 (20%)	N/A	1.23 (.97)	5 (50%)
Frequency of Tea Use (Typical Week)	23 (46%)	N/A	5.33 (8.21)	21 (91.3%)
Quantity of Tea Use	17 (34%)	N/A	1.26 (1.08)	12 (70.6%)
Frequency of Caffeinated Soda Use (Typical Week)	26 (52%)	N/A	2.91 (3.74)	23 (88.5%)
Quantity of Caffeinated Soda Use	18 (36%)	N/A	1.24 (.87)	13 (72.2%)
Frequency ED Use (Typical Week)	12 (24%)	N/A	3.75 (2.92)	12 (100%)
Quantity of ED Use	13 (26%)	N/A	1.01 (.34)	9 (69.2%)
Caffeine Servings (Typical Day)	28 (56%)	N/A	4.14 (8.48)	20 (71.4%)

Alcohol. All (100%) of participants had at least one alcohol discrepancy and the mean number of discrepancies was 2.32 (*SD* = .91) (see Table 21). Almost three-fourths (74%) had a discrepancy on frequency of alcohol use reports, with a mean discrepancy magnitude of 2.91 (*SD* = 2.80) days. Additionally, about three-fourths (75.7%) of these discrepancies were CAS

overestimates. Over two-thirds (68%) had a quantity of alcohol use discrepancy with a mean discrepancy magnitude of 2.62 (*SD* = 3.15) drinks. Rates of CAS estimates exceeded TLFB estimates for all measures except quantity of alcohol use; where only 14.7% of the discrepancies were CAS overestimates.

Almost three-fourths (72%) of participants had a discrepancy regarding binge drinking, with a mean discrepancy magnitude of 2.64 (*SD* = 1.99) binge occasions. Over half (58.3%) of these binge drinking discrepancies were CAS overestimates, and one fourth (25%) of those with binge drinking discrepancies reported binge drinking on only the CAS or TLFB.

One-fifth (19.57%) of the sample had a discrepancy for number of past month blackouts. The mean discrepancy magnitude was 2.11 (*SD* = 1.36) blackout occasions with all discrepancies (100%) being CAS overestimates. Two-thirds (66.67%) of those with blackout drinking discrepancies reported blackouts on only either the CAS or TLFB.

Table 30. CAS and TLFB discrepancy information for alcohol quantity and frequency questions.

Comparison Question	Number (%) with Discrepancies	Mean (SD) Number Discrepancies	Mean (SD) Magnitude of Discrepancies	Number (%) Survey Overestimates
Overall	50 (100%)	2.32 (.91)	N/A	N/A
Frequency of Alcohol Use	37 (74%)	N/A	2.91 (2.80)	28 (75.7%)
Quantity of Alcohol Use	34 (68%)	N/A	2.62 (3.15)	5 (14.7%)
Binge Drinking	36 (72%)	N/A	2.64 (1.99)	21 (58.33%)
Blackout Drinking	9 (19.6%)	N/A	2.11 (1.36)	9 (100%)

CAC. As shown in Table 22, over half (58%) of the sample had CAC discrepancies and the mean number was 1.90 (*SD* = .31). Of the 29 individuals endorsing CAC use, 19 (65.5%) reported such use on only TLFB or CAS. For frequency of CAC use, 58% of the sample had a discrepancy between CAS and TLFB reports with a mean discrepancy magnitude of 6.34 (*SD* = 8.01) CAC use days. Four-fifths (79.3%) of these discrepancies were survey overestimates. About half (52%) of participants had a quantity of CAC use discrepancy with the mean discrepancy magnitude being 3.51 (*SD* = 3.96) CAC beverages. Additionally, almost three-fourths (73.1%) of these discrepancies were survey overestimates.

Table 31. CAS and TLFB discrepancy information for CAC quantity and frequency questions.

Comparison Question	Number (%) with Discrepancies	Mean (SD) Number Discrepancies	Mean (SD) Magnitude of Discrepancies	Number (%) Survey Overestimates
Overall	29 (58%)	1.90 (.31)	N/A	N/A
Frequency CAC Use	29 (58%)	N/A	6.34 (8.01)	23 (79.3%)
Quantity CAC Use	26 (52%)	N/A	3.51 (3.96)	19 (73.1%)

Discrepancies and Other Variables. Correlations were run for CAS and TLFB discrepancy variables and other continuous variables (see Table 23). Younger participants had more overall discrepancies (r = -.40; p = .004), caffeine discrepancies (r = -.31; p = .03), and alcohol discrepancies (r = -.31; p = .03). Participants with less years of education had greater caffeine discrepancies (r = -.33; p = .02). Those with higher AUDIT-C scores had a greater number of overall discrepancies (r = .40; p = .004), alcohol discrepancies (r = .40; p = .004), and CAC discrepancies (r = .40; p = .004). No other significant correlations were found.

Table 32. Correlations between discrepancy variables and other variables.

Other Variables	**Overall Discrepancies *r* (*p*-value)**	**Caffeine Discrepancies *r* (*p*-value)**	**Alcohol Discrepancies *r* (*p*-value)**	**CAC Discrepancies *r* (*p*-value)**
Age	-.40 (.004)*	-.31 (.03)*	-.31 (.03)*	-.25 (.08)
Education Completed	-.26 (.07)	-.33 (.02)*	-.02 (.88)	-.11 (.45)
PHQ-2 Score	-.02 (.89)	-.02 (.90)	-.06 (.68)	.03 (.83)
AUDIT-C Score	.40 (.004)*	.20 (.18)	.40 (.004)*	.40 (.004)*
BIS Score	.25 (.08)	.16 (.27)	.25 (.08)	.17 (.25)

*Denotes a statistically significant r-value ($p < .05$).

As shown on Table 24, when discrepancy data were compared across different categorical variable groups, *t*-test differences were found only for drug users and nonusers. Specifically, compared to non-drug users, participants reporting drug use on either TLFB or CAS had higher overall discrepancies ($t(48) = -2.96$, $p = .005$), caffeine discrepancies ($t(48) = -2.71$, $p = .009$), and CAC discrepancies ($t(48) = -2.14$, $p = .04$).

Table 33. *t*-tests between discrepancy variables and other variables.

Other Variables	**Discrepancy Variables**	**Mean (SD)**		***t* (*p*-value)**
Gender	Overall Discrepancies	**Male** (*N* = 14)	**Female** (*N* = 36)	.05 (.96)
		7.93 (3.41)	7.89 (2.53)	
	Caffeine Discrepancies	4.36 (1.87)	4.53 (1.86)	-.29 (.77)
	Alcohol Discrepancies	2.57 (.85)	2.22 (.93)	1.22 (.23)
	CAC Discrepancies	1.00 (.96)	1.14 (.99)	-.45 (.66)
Other Drug Use (TLFB or CAS)	Overall Discrepancies	**Non-user** (*N* = 19)	**User** (*N* = 31)	-2.96 (.005)*
		6.53 (2.29)	8.74 (2.72)	
	Caffeine Discrepancies	3.63 (1.54)	5.00 (1.84)	-2.71 (.009)*
	Alcohol Discrepancies	2.16 (.96)	2.42 (.89)	-.98 (.33)
	CAC Discrepancies	.74 (.93)	1.32 (.95)	-2.14 (.04)*
Family History of Alcohol/Drug Problems	Overall Discrepancies	**FH Negative** (*N* = 32)	**FH Positive** (*N* = 18)	-1.05 (.30)
		7.59 (2.83)	8.44 (2.64)	
	Caffeine Discrepancies	4.34 (1.89)	4.72 (1.78)	-.69 (.49)
	Alcohol Discrepancies	2.25 (.98)	2.44 (.78)	-.72 (.48)
	CAC Discrepancies	1.00 (.98)	1.28 (.96)	-.97 (.34)

*Denotes a statistically significant *t*-value ($p < .05$).

Aim 3. Examine qualitative personal interview data and identify themes and patterns

To examine Phase 2 sample representativeness, the caffeine, alcohol and CAC use as well as other measures were compared for eligible students who did or did not choose to complete Phase 2 of the study. No significant *t*-test differences were found for any variables summarized in Table 25.

Table 34. *t*-tests comparing caffeine, alcohol, and CAC use and other continuous variables by those who did/did not consent to Phase 2.

Variables		**Mean (SD)**		***t*-value (*p*-value)**
		Declined (*N* = 13)	**Consented** (*N* = 16)	
Caffeine	Frequency (Days of Use)	22.23 (8.30)	21.75 (10.49)	.13 (.89)
	Quantity (# Drinks on Typical Day)	1.56 (.84)	3.10 (3.91)	-1.53 (.14)
Alcohol	Frequency (Days of Use)	8.54 (5.74)	9.88 (8.29)	-.49 (.63)
	Quantity (# Drinks on Typical Day)	4.30 (1.53)	6.79 (6.61)	-1.46 (.16)
CAC	Frequency (Days of Use)	.92 (.95)	2.31 (6.93)	-.72 (.48)
	Quantity (# Drinks on Typical Day)	2.70 (1.90)	6.28 (9.13)	-1.02 (.35)
Age		19.92 (1.50)	20.81 (2.61)	-1.09 (.29)
Education Completed		14.08 (1.51)	14.28 (2.10)	-.28 (.78)
PHQ-2 Score		1.69 (1.38)	1.50 (1.67)	.33 (.74)
AUDIT-C Score		5.08 (1.93)	6.06 (2.29)	-1.23 (.23)
BIS Score		17.23 (3.77)	15.13 (2.78)	1.73 (.10)

Phase 2 qualitative interviews were examined for patterns of responses, with common themes regarding participant beliefs about discrepancies identified and presented below. All Phase 2 participants reported that they responded honestly during Phase 1 data collection.

For caffeine, over three-fifths (61.54%) of participants attributed discrepancies to having to estimate/average on the CAS, which they felt made it less accurate than the TLFB. Of those who attributed caffeine differences to estimating/averaging, 61.11% were survey overestimates. Almost one-fifth (15.38%) believed discrepancies were due to variability in their pattern of caffeine use (e.g., caffeine use is different when at home versus at college) and the measurement window was restricted to one month. Less than one-tenth (7.69%) attributed the difference to not reporting a caffeinated beverage as a result of not considering it to be caffeinated. Another 7.69% reported general confusion about the question(s) and an additional 7.69% stated they simply made a mistake.

For alcohol, the majority (85.71%) of participants believed discrepancies were the result of having to estimate/average on the CAS. Of those who said they had to estimate on alcohol questions, 57.14% of discrepancies were survey underestimates. Less than one-tenth (7.14%) stated they were confused about the alcohol question(s) and 21.43% reported that discrepancies were the result of making a mistake (e.g., forgot about specific drinking incidents).

Almost three-fourths (73.33%) of Phase 2 participants believed the CAS blackout definition was not accurate as it did not capture the full intensity of a "blackout." Specifically, they felt the inaccuracy stemmed from the fact it did not include characteristics such as “passing out” and “not remembering anything” in the description. The remainder (26.67%) thought the definition was accurate. Over one-tenth (12.5%) of Phase 2 participants reported a blackout on only the CAS.

For CAC, discrepancies were attributed by over one-third (35.71%) to believing the question(s) referred to using caffeine and alcohol in the same day rather than in the same beverage. All (100%) discrepancies associated with this perception of the question were survey overestimates. Another 35.71% attributed CAC discrepancies to caffeine errors, such as reporting non-caffeinated mixers and not including caffeinated beverages in reports. Of those who reported a CAC caffeine error, 60% were survey overestimates. Over one-fourth (28.57%) reported general confusion about the question and less than one-tenth (7.14%) stated they made a mistake.

Discussion

ED and CAC use in college students has garnered much media attention and research has begun to elucidate the negative consequences of CAC use, constituting a public health issue. The need for accurate measurement is critical to future research. A review of the relevant literature, however, found no standardized measures currently exist for CAC use. Since measurement of CAC use is limited to self-report, it is necessary to examine accuracy of college student reports. In addition, given the health consequences associated with such use, it is clinically important to develop brief tools to screen for heavy/problem caffeine and CAC use.

The present study used a mixed methods approach to compare alternate methods for collecting recent caffeine, alcohol and CAC use information from college students. In Phase 1, a sample of self-identified regular caffeine and alcohol using college students (N = 50) answered questions about their alcohol, caffeine and CAC use. Participants were randomly assigned to complete either CAS followed by TLFB or TLFB followed by CAS. Inconsistencies were identified and participants reporting CAC use were invited to complete Phase 2, a qualitative personal interview about the discrepancies, with a focus on identification of common themes and misperceptions.

Summary of Findings

While participants consistently reported greater frequency of use on the CAS as compared to the TLFB, quantity measures were more varied. Further examination of the agreement between CAS and TLFB reports at the individual level revealed consistent patterns of discrepancies within substance classification (alcohol vs caffeine vs CAC) (i.e., different patterns depending upon type of substance) and type of question assessing use (e.g., quantity versus frequency). The magnitude of differences and proportion of participants with CAS

overestimates (compared to TLFB) also varied depending on substance type (alcohol vs caffeine vs CAC). Significant associations were found between these discrepancies and other variables. Additionally, we found that the majority of participants had inconsistencies regarding their reporting of caffeine.

Phase 2 qualitative interviews identified themes in the discrepancies, which also varied based on substance type. For alcohol, the majority of participants attributed discrepancies to variability in drinking that was difficult for them to estimate in response to the past 30 days CAS alcohol items. For caffeine, participants indicated not only problems estimating average quantity consumed, but also inaccurate knowledge about the caffeine content of certain beverages. For CAC, the majority of participants attributed inconsistencies to either their misinterpretation of what was meant by "caffeine combined with alcohol" and/or limited knowledge about which beverages do and do not contain caffeine (e.g., Sprite). Additionally, associations were found between assessment method and CAS over/under estimates; with CAS overestimations for caffeine, but CAS underestimations for alcohol.

CAS and TLFB Discrepancies.

Overall participant self-reports of caffeine quantity and frequency of use on the CAS did not differ from those obtained with TLFB. For alcohol, however, significant CAS and TLFB differences were found for number of drinking days, number of alcoholic drinks/drinking day and number of CAC use days. Mean alcohol and CAC frequency was significantly higher on CAS than TLFB. However, average alcohol quantity was greater on the TLFB compared to CAS. There was also a general trend for CAS reports to be higher than those on the TLFB, with the exception of overall quantity measures for caffeine, alcohol, and CAC as well as caffeinated tea.

Alcohol. Whereas there has been no published research on caffeine and CAC discrepancies, previous studies have examined problems associated with alcohol assessment. While the present study found CAS overestimates (compared to TLFB reporting) for number of alcohol use days, the CAS alcohol quantity (typical day) item underestimated what was found on the TLFB. Previous studies with QF approach to alcohol assessment have found similar results; college students tend to underestimate the average number of drinks they consume (Fishburne & Brown, n.d.) with these responses typically falling between the mean and mode (Gruenewald et al. 1996). Together, these findings suggest that when the full TLFB is impractical, alternatives, such as the GF approach should be considered.

Binge drinking. The prevalence of binge drinking (past 30 days) in our sample was 78-80% (including both TLFB and CAS). This is much higher than the national binge drinking rate among college students of 44% (Wechsler & Nelson, 2008). One possible reason for our higher rate is that our sample only included regular alcohol users. Nearly three-fourths of participants had a CAS-TLFB discrepancy for number of binge drinking days. One-fourth of these discrepancies involved reporting binge drinking (at least one binge drinking episode) by only one method (i.e., either the CAS or TLFB), with over half (55.6%) by only CAS. The magnitudes of these discrepancies were also relatively high considering that this type of drinking is lower frequency. In contrast to our finding that typical alcohol quantity was generally underestimated on the CAS, the majority of discrepancies involving binge drinking were CAS overestimates.

Quantity of alcohol use is a gauge of problematic drinking and is often used to screen individuals for both research and clinical purposes (NIAAA, n.d.). It is recommended that one or two item screeners taking a similar form to our question, “How many times in the past year have you had X or more drinks in a day?” (where X is 5 for men and 4 for women) (Smith et al.,

2009) be used to determine risk drinking in clinical settings (Strobbe, 2014). The fact that individuals are overestimating binge episodes using a standard screening question could indicate that we may be overly identifying this specific alcohol risk behavior. These results point to the need for more accurate collection of this crucial information; measure sensitivity is particularly important in the assessment of problematic drinking as these types of alcohol use are clinically relevant at smaller frequencies.

At the broadest (federal) level, consolidation of what constitutes a binge episode is needed as there is currently no single definition of binge drinking that is consistently used in research/clinical work. The NIAAA defines binge drinking distinguished by sex ($\geq$4 drinks for women or $\geq$5 for men per occasion), while SAMHSA defines binge drinking as 5 or more alcoholic beverages per occasion, regardless of sex/gender (NIAAA, n.d.).

Furthermore, these existing binge definitions have received criticism for potentially misrepresenting problem drinking, indicating the possible need for modification of risk drinking definitions (Read et al., 2008). For example, studies have shown that many binge drinkers do not reach intoxication as determined by blood alcohol concentration levels (Beeirness, Foss, & Vogel-Sprott, 2004; Lange & Voas, 2001; Perkins, DeJong, Linkenbach, 2001). There are other important alcohol risk indicators that are typically not part assessments of risk drinking, including certain alcohol consequences (e.g., impaired control) found to be predictive of later alcohol use disorders (Chung & Martine, 2001; Nelson, Little, Heath, & Kessler, 1996; O'Neill & Sher, 2000). Many college students drink substantially more than the current binge standard, but are not distinguished from less-alcohol-consuming binge drinkers (White, Kraus, and Swartzwelder, 2006). Read et al. (2008) conducted a study examining separate cutoffs for quantity of use as it relates to risk drinking (non-binge - less than 4/5 drinks; binge - 4/5 drinks;

and heavy binge - 6/7 drinks or more) and found that, whereas binge drinkers were not different from the non-binge drinkers on important indicators (e.g., drinking frequency, total alcohol consequences), heavy binge drinkers were different. Additionally, they found that the heavy binge drinkers differed from the binge drinkers with regard to impaired control and blackout drinking (Read et al., 2008).

Placing individuals in the current binge categories, which are designated as at risk, possibly indicates these measures are overly sensitive (e.g., Gruenewald, Johnson, Light, & Salz, 2003), blunting the ability to identify those at greatest risk (Read et al., 2008). This is potentially exacerbated by the fact that, based on our results, participants tend to overestimate binge episodes when asked specifically about them.

Blackout drinking. Similar patterns were found for alcohol-related blackouts, where all inconsistencies between the two assessment methods were cases where participants reported blackouts on the CAS, but not on the TLFB. Several factors may account for this discrepancy. First, method of administration may play a role, with greater anonymity completing a computer survey than participating in a personal interview. Research has shown that for sensitive issues in particular, anonymity can provide a safer setting for disclosure of sensitive or potentially stigmatizing experiences (e.g., Whelan, 2007). Second, the definition of blackout varied for the CAS and TLFB. On the CAS, participants were asked about blackouts (lifetime and past 30 days) using a question from a 5-item standardized screening tool to identify problem drinking (Russell, 1994), which does not incorporate the term "blackout". In contrast, during the TLFB, participants were simply asked about days on which they consumed alcohol and a blackout occurred, with no further definition of the term. The intent of these methodological choices was to examine the agreement between the Russell (1994) definition of blackout and the term itself;

any discrepancies might speak to perceptual incongruence between the two. The fact that all discrepancies were CAS overestimates indicates the definition is possibly more sensitive than the term itself. This is supported by the fact that two-thirds of those with blackout drinking discrepancies reported blackouts on only the CAS. Additionally, Phase 2 responses indicated that most participants believed the CAS blackout definition was not accurate as it did not capture the full intensity of the term "blackout."

Regardless of the phrase/term used, methodology could also account for the difference in reporting blackouts. It is possible that participants are less likely to endorse blackouts during the TLFB interview compared to the self-administered computer survey. Participants have been shown to provide more socially desirable responses during face-to-face interviews than self-administered surveys (Bowling, 2005). However, even on computerized assessments, such misreports of sensitive information still occur (Tourangeau & Yan, 2007). Additionally, the lack of specificity of the recent blackout item (i.e., focus only on past 30 days) could have resulted in participants thinking about a broader window of time.

The definition of blackout drinking varies widely between measures. For instance, the Young Adult Alcohol Consequences Questionnaire (YAACQ) is a measure that assesses various domains of alcohol consequences (e.g., blackout drinking) to more comprehensively assess for risk (Read, Hahler, Stron, & Colder 2005). It has been shown to be reliable, valid and predictive of alcohol use and binge drinking frequency (Read et al., 2007). The YAACQ definition of blackout consists of 7 items (e.g., “I have had a hangover (headache, sick stomach) in the morning after drinking”; “I have passed out from drinking”; “I have not been able to remember large stretches of time while drinking”), none of which ask about “other” (e.g., friend or family member) reports (Read, Hahler, Stron, & Colder 2005).

Our results could be confounded by our choice to measure blackouts on the CAS using one question that is based on "other" reports. Investigating inconsistent blackout reporting was not the primary purpose of our study and thus was limited to only our single screener question; using a more comprehensive measure, such as the YAACQ, may have yielded different findings regarding discrepancies.

Consolidation of blackout definition and term is important as there is an apparent disconnect. This incongruence between term and definition of blackout could lead to inaccurate assessment of blackouts. Our findings could also be indicative of a fundamental misunderstanding of what a blackout actually is, which could potentially lead individuals experiencing blackouts to not weigh these experiences as problematic. Researchers and clinicians should be cautious when using the term "blackout" interchangeably with the definition-based question we used in this study. It will be important to conduct further research that more objectively examines this blackout question, the term "blackout", and participant responses/perceptions regarding both the question and term itself (particularly among college students).

There are also gender differences in blackout drinking as an indicator of later alcohol problems, which should be considered with regard to screening for risk. For example, Read, Wardell, Bachrach (2013) found that blackout drinking was linked to later increases in alcohol consumption in men, but that the opposite was true for women. Additionally, among males, externalizing behavior (e.g., physical fights) is less likely to indicate problems later in life compared to females (Kahler et al., 2004).

Caffeine inconsistencies. The present study was among the first to evaluate use of the TLFB to collect past month quantity and frequency data for caffeine and CAC. In general,

identifying reliable and valid tools to measure quantity and frequency of caffeine use may be more challenging than for alcohol. While there is a standard drink unit for alcohol (i.e., 5 oz wine, 12 oz beer, 1.5 oz liquor), caffeine has many sources, with widely varying caffeine content both within and across beverage types and minimal standardization (Carpenter, 2014). In addition, participants may not have accurate knowledge about which beverages do and do not contain caffeine, particularly when it comes to sodas (e.g., Sprite; Polak, Hancock, & Svikis, n.d.a.).

In this study, using the TLFB, participants appeared able to estimate their caffeine use on individual days and every participant was able to provide a response for all days in the assessment period. Since no standardized methods were available for collecting detailed caffeine use data, TLFB procedures were modeled after those established for alcohol use (Sobell & Sobell, 1992) These TLFB procedures have been used successfully to measure other drug use across diverse populations and over extended time intervals (Fals-Stewart et al., 2000; Robinson et al., 2014; Norberg, Mackenzie, & Copeland, 2012; DeMarce et al., 2007; Carey et al., 2004), as well as for other behaviors (e.g., sexual risk behavior, exercise, and work) (Carey et al., 2001; Panza et al., 2012; Svikis et al., 2012). The present study provided an opportunity to examine, using the TLFB, challenges specific to the measurement of caffeine quantity and frequency of use.

In the present study, the majority of participants had inconsistencies involving reports of caffeine use and almost half of these inconsistencies resulted from inaccurate reports as to what did/did not contain caffeine. Specifically, when asked to recall caffeinated drinks, one fourth of the sample reported non-caffeinated drinks; typically, these misreports were non-caffeinated sodas (e.g., Sierra Mist and Sprite), but other non-caffeinated beverages were also thought to

contain caffeine (e.g., herbal tea and sparkling water). Almost one-third of those with non-caffeinated reports had over 10 of these misreports, substantively impacting the overall quantity and frequency of use reported over the 30-day period. The reverse was also true, with about one-third of participants labeling caffeinated beverages (e.g., green tea and Pepsi) as non-caffeinated beverages they had consumed. So, a single error in knowledge/caffeine can lead to substantive misreports when repeated throughout the assessment period. This is cause for concern for studies where quantity and frequency of caffeine use are integral variables.

These inconsistencies in caffeine reporting suggest a potential knowledge gap regarding what does/does not contain caffeine; a large proportion of our sample did not have knowledge about caffeine content of the beverages they consumed in the past month. Caffeine misreporting appears to be associated with beverage categories that have both caffeinated and non-caffeinated versions, such as soda and tea, versus beverage categories that are predominantly/always caffeinated (e.g., energy drinks). This lack of knowledge about caffeine could result in survey reporting errors, especially for commonly used beverages that have both caffeinated and non-caffeinated types (e.g., sodas). Such reporting errors would disable both an accurate assessment of use and precise identification of those at risk for problems associated with use.

Of particular concern was the fact that almost half of CAC users had CAC misreports with over 90% of these individuals reporting a caffeinated beverage that they had not previously reported on the caffeine TLFB; the majority of these misreports were caffeinated sodas (e.g., Coca Cola), but other types of caffeinated beverages were also reported (e.g., tea, Red Bull, and coffee). This disconnect in reporting potentially indicates that individuals frequently do not consider CACs to be caffeinated beverages. The implications of this misperception include that it could engender inaccurate reporting on CAC and caffeine assessment, and that participants may

be less cognizant about their use of CACs. Additionally, such CAC reporting errors potentially lead to compounding of errors (i.e., alcohol and caffeine misreporting simultaneously).

There were differences in discrepancy patterns between quantity of caffeine use and caffeine servings (participants were told on the CAS that 1 caffeine serving = 1, 8oz cup of coffee). Compared to quantity of caffeine use responses, those for caffeine servings had increased number of participants with the discrepancy, discrepancy magnitude, and proportion of survey overestimates.

The inclusion of the "servings" question was done in part to check the agreement with more detailed information from the TLFB, compared to asking about quantity. The fact that this "servings" question had greater variability speaks to the fact that asking for servings might be less accurate than asking about quantity of caffeine use without referencing servings.

Comparison of quantity of caffeine use and caffeine serving responses is another gauge of participant understanding of caffeine amounts in beverages consumed. The fact that participants had greater discrepancies regarding servings than quantity of caffeine use could indicate participants do not have a firm understanding of caffeine contents, which is consistent with previous findings from our lab (Polak, Hancock, & Svikis, n.d.a.). This would imply that there is a general deficit in knowledge about caffeine. Additionally, this discrepancy could indicate the potential for problems in assessment of caffeine use; specifically, the use of questions asking about servings of caffeine use should not be expected to accurately reflect number of drinks consumed. However, our results could also be confounded by the questions themselves as caffeine and CAC use questions have not been previously validated.

Additionally, we found that while overall frequency of caffeine use was about 21 days for both CAS and TLFB, specific caffeinated beverage type CAS and TLFB frequencies were

discrepant. Potential reasons for this could be the way our questions were worded, such as asking about typical week use versus overall days of use.

Extending beyond issues of measurement, a lack of knowledge about caffeine poses a general health concern as individuals could be less likely to gauge healthy consumption amounts. Heavy caffeine consumption is defined as 500 to 600 mg per day (Mayo Clinic, 2014), with 8% of our sample meeting this criteria (one or more days of at least 500 mg of caffeine in the past 30 days). Within our sample, there were multiple individuals who had large numbers of misreports, potentially impacting their caffeine/CAC reporting and understanding of their use. Such misreporting becomes especially problematic when screening for/assessing unhealthy, heavy caffeine consumption. However, our data is preliminary and more research is needed to determine if this misreporting does negatively impact perceptions and reporting of use, warranting a health concern.

To our knowledge, this is first study to systematically examine consistency in self-reports of caffeine use with two methods of assessment (interview and survey). These results are consistent, however, with preliminary data from the Svikis lab, which indicated that many college students misreported non-caffeinated beverages as having caffeine when asked to report CAC beverages consumed (Polak, Hancock, & Svikis, n.d.). A recent study examining validity of a food frequency survey compared with 24-hour dietary recalls among premenopausal women also found variations in caffeine reporting; they found survey overestimates for caffeine and coffee drinks/cocoa, and survey underestimates for soda intake (Schliep et al., 2013). We did not find the same specific types of discrepancies, which could be due to our small *N* and/or the fact that we did not include daily prospective tracking as they did (Schliep et al., 2013).

Comparison of caffeine, alcohol, and CAC. All participants had caffeine and alcohol discrepancies and every individual who endorsed CAC use on either the CAS or TLFB had CAC discrepancies. Magnitudes/degree of discrepancies varied depending on substance; overall, alcohol had the lowest magnitude, followed by caffeine, then CAC. Additionally, rates of individuals reporting use on only CAS or TLFB varied by type of caffeinated beverage, with the lowest rate for coffee (11.9%) and the highest for CAC (65.5%). Within specific type of caffeine, ED use (59%) had the highest rate of this type of discrepancy.

These high rates of discrepancies speak to the need for more research, including prospective monitoring as an assessment method to be compared with survey reporting and further direct examination of caffeine knowledge (e.g., giving a quiz). Our findings about CAC discrepancies being larger compared to caffeine and alcohol by themselves suggest that CAC measurement is particularly vulnerable to discrepant reporting and thus increased inaccuracy of survey reports. This could be at least partially explained by the fact that CAC survey questions have not been validated as alcohol and, to some extent, caffeine questions have been. However, it is also important to be cautious when comparing across alcohol, caffeine, and CAC categories as there could be important base rate differences (e.g., if the number of caffeine use days is higher than alcohol use days, then there may be more opportunities for larger caffeine discrepancies).

Phase 2 participant responses help to elucidate other reasons for these discrepancies and indicate fundamental differences by type of substance assessed. The majority of participants believed the reason for alcohol discrepancies was that they had to estimate/average on the CAS, making it less accurate than the TLFB. Among these individuals, the majority of discrepancies

were CAS underestimates, possibly speaking to a tendency to underestimate alcohol use in the absence of specifically remembering such use.

Less (but still most) Phase 2 participants attributed caffeine discrepancies to having to estimate/average on the CAS. However, unlike alcohol, the majority of caffeine discrepancies among these participants were CAS overestimates. This indicates that individuals are possibly more likely to overestimate caffeine use when they do not recall specific use. Additionally, other participants attributed caffeine discrepancies to not considering beverages to be caffeinated and variable patterns of use. These elaborations provide further evidence for the existence of gaps in knowledge/awareness regarding caffeine use and the negative impact of such gaps on accuracy of survey reporting.

For CAC, over one-third of participants attributed discrepancies to believing the question(s) referred to using caffeine and alcohol in the same day, with all of these being CAS overestimates. Another 36% of participants indicated caffeine errors; the majority of these discrepancies were CAS overestimates. This again potentially indicates that caffeine errors are associated with these discrepancies, and may particularly lead to overestimation on survey assessment. These responses suggest a fundamental misunderstanding of what is being asked during CAC assessment; individuals possibly do not know what CACs are, indicating need for psychoeducation. These errors could also be due to how we worded the questions (i.e., "During the past 30 days, how many days did you drink either a caffeinated alcoholic product or a caffeinated beverage in combination with alcohol?" and "During the past 30 days, on days when you did drink caffeine combined with alcohol, how many drinks did you usually have?"), suggesting the need to improve these questions.

Relationship with other variables. We found that discrepancy variables were associated with several other variables. Younger participants had more overall discrepancies, caffeine discrepancies, and alcohol discrepancies; younger individuals may have more variable use patterns, making it harder to estimate/average over reference periods. Younger students might also be less cognizant of their substance use in general. This is perhaps a reason why this period of emerging adulthood is such a crucial time for development of substance use problems. Similarly, those with less education had greater caffeine discrepancies.

Those with higher AUDIT-C scores had a greater number of overall discrepancies, alcohol discrepancies, and CAC discrepancies. This pattern of associations between discrepancies and problematic alcohol use is potentially a result of greater alcohol use increasing the likelihood of discrepancies in reporting such use. It may also speak to problematic alcohol use being associated with decreased mindfulness of use. Additionally, heavy/problematic drinkers (AUDIT-C positive) could have poorer recall because their increased alcohol consumption could indicate they experienced intoxication and/or blackouts, potentially impacting their memory (White, 2004). Future research should look at those who report such problematic use and whether or not they have more discrepancies in their reports.

Lastly, we found that participants who reported drug use on either the TLFB or CAS had higher overall discrepancies than non-drug users. Similar to the association between AUDIT-C scores and discrepancies, this could be related to the fact that drug use has been shown to be associated with increased use of other substances (e.g., NIAAA, 2008). This difference could also suggest that problematic substance use is associated with less awareness of substance use in general.

To our knowledge, no other study has examined discrepancies and their associations with other variables. These discrepancies have broad implications for other substance use. For instance, rates of drug use followed a similar pattern to alcohol, caffeine and CAC use in that reports on the CAS were generally higher than on the TLFB; important associations between drug discrepancies and other variables may exist. Additionally, as indicated by our comparison between caffeine, alcohol, and CAC, fundamental differences appear to exist by substance type that help to explain discrepant reporting. Thus, other types of substance use reporting might have unique features contributing to discrepancies. One such example may be abuse of prescription medication; individuals who abuse prescription medications have been found to misperceive safety regarding such use (Volkow, 2010), which may have implications for how they report this use.

Eligibility. Interestingly, we found instances of CAS and TLFB reports conflicting with initial screening question responses for both regular alcohol (at least once per week for the past month) and caffeine (at least twice per week for the past month) use. We examined how well general, screening reports of past 30 days "regular use" of caffeine and alcohol compared with more detailed Phase 1 responses. To determine past month regular caffeine and alcohol use from CAS and TLFB reports, the number of use days was divided by 30 and multiplied by 7; those with $\geq$2 and $\geq$1 were counted as regular caffeine and alcohol users, respectively. For 44% of the sample, a mismatch was found between preliminary screening for regular alcohol and caffeine use, and subsequent CAS and/or TLFB reports of such use. Additionally, 10% would not have met the caffeine screening criteria based on both CAS and TLFB responses and 22% would not have met for alcohol. One implication of this finding is that our results may not necessarily be generalizable to regular alcohol and caffeine users. Another potential implication of our results is

that participants were reporting "usual" use as it applies to a reference period of greater than the past 30 days. Since substance use among college students has been shown to vary based on different contexts (e.g., weekend vs weekday alcohol use; Woodyard & Hallam, 2010), a participant's "usual" use will potentially vary substantially based on the specific contexts included within the assessment reference period.

This particular discrepancy issue has not, to our knowledge, been previously addressed in the literature. Examination of screening criteria/assessment is particularly important as studies investigating substance use often include criteria cutoffs using broad use terms, similar to our study eligibility criteria. For instance, laboratory studies on CAC use have traditionally incorporated these cutoffs into their inclusion criteria, such as including only "moderate social drinkers" (Azcona et al., 1995) or "moderate alcohol users" (Ferreira et al., 2004). A similar eligibility criterion was also used when establishing the psychometric properties of the YAACQ; participants included in the YAACQ validation study were those who self-reported alcohol use at least once per week in the past three months (Read, Hahler, Stron, & Colder, 2005).

Furthermore, these findings have implications for validity of research pools within the substance abuse field as well as potentially extending into any research domain that assesses eligibility using retrospective, self-report information. In particular, alcohol screening questions are widely used in both research and clinical work (Strobbe, 2014). Among the college population, the use of these questions are especially important as this group is particularly vulnerable to development of problems related to alcohol use. Additionally, studies have shown alcohol use among college students is significantly different on the weekends compared to the weekdays, such that researchers often examine these times of use separately (Woodyard &

Hallam, 2010). Thus, discrepancies between the interview-format screening questions and survey questions speak to the need to understand why these discrepancies occur.

Sample Representativeness. In terms of race, the sample is relatively consistent with VCU overall: 14% Asian, 19% African American, 8% Hispanic, 5% more than one race, and 48% White (Dick et al., 2014). Our sample has a disproportionate amount of females represented compared to males. However, this gender representation does appear similar to that of the Student Health Center patient breakdown as indicated by the most recent report available (for appointment count, the Student Health Center breaks down to 70% female and 30% male).

As determined by the PHQ-2, 16% of participants scored at or above the clinical cutoff for depression (score of 3 or above). This is consistent with prevalence rates of depression among college students (Ibrahim et al., 2013). We found prevalence of CAC use in our sample to be almost half of participants by CAS and less than one-third by TLFB, both of which are consistent with previous findings of CAC use rates (Brache & Stockwell, 2011; MacKillop et al., 2012; Miller, 2008). However, we expected that our CAC use rate would be higher than previously reported CAC prevalence rates given that our sample is not representative of college students as a whole (as we started with regular caffeine and alcohol users).

The rate of individuals that may have ADHD (42%) as well as the rate of official ADHD diagnosis (16%) were relatively high compared to previously reported prevalence rates of ADHD among college students (between 2-8%) (Green & Rabiner, 2012). However, these rates from other studies should be interpreted with caution as they are based on self-reported ADHD symptoms and/or diagnostic status rather than comprehensive assessment. Additionally, these reported rates of ADHD did not come from nationally representative samples, but from individual universities. We also based our rate of official diagnosis on diagnostic status rather

than completing a full evaluation. In addition, our CAS question about ADHD diagnosis included a response option for belief that one has ADHD, which was not included in previous studies.

Our higher rates of ADHD are possibly also attributable to the fact that our sample only included regular alcohol users. Previous studies have established a relationship between increased alcohol use and college students with ADHD compared to peers without ADHD (Green & Rabiner, 2012). However there have been mixed results regarding this relationship (e.g., Upadhyaya et al., 2005).

General Characteristics. We found several significant associations between the various substance use variables and the other variables assessed, which we did not expect to find given our small sample size. Participants with higher BIS scores had increased TLFB caffeine frequency, consistent with previous findings (Jones & Lejuez, 2005). Those who used drugs reported significantly higher CAS alcohol frequency than those who did not, which is congruent with the established robust relationship between alcohol and drug use (NIAAA, 2008). Additionally, we found that males reported significantly greater CAS alcohol quantities than females, which has been consistently found in past studies (e.g., Wilsnack et al., 2009).

We also found nonsignificant results that conflict with the current literature. For instance, we did not find a significant relationship between family history of substance problems and alcohol use, but we expected to find this relationship based on the consistency of this association across past studies (e.g., LaBrie et al., 2010). None of the relationships between CAC CAS and TLFB quantity and frequency and other variables were significant. We expected significance, however, as CAC use has been previously linked to gender differences (Amlung et al., 2013), problematic alcohol use (Thombs et al., 2010; Arria et al., 2010; Arria et al., 2011), impulsivity

(Amlung et al., 2013), and drug use (Arria et al., 2010; Arria et al., 2011). We suspect these nonsignificant findings are possibly the result of a small *N*, which did not provide the power necessary to detect these differences. Additionally, we started with regular alcohol and caffeine users so we do not have the heterogeneity of the larger population.

Lastly, we found that mean CAS caffeine frequency was significantly greater for females than males. While studies have not directly examined gender difference in caffeine consumption among college students, adolescent males have been found to be have increased use compared to female counterparts (Arria et al., 2014). Our finding might again be due to the fact that our study had a small *N* and, in particular, we had disproportionately more females than males.

We found an order effect for TLFB frequency of alcohol use, which could potentially be confounding results that include this variable. No significant differences between those who did/did not consent to Phase 2 were found across several other continuous variables (i.e., age, education, PHQ-2 score, BIS score, and AUDIT-C score). Further, across continuous variables, no significant differences between those who did/did not consent to Phase 2 were found.

Implications for Research/Clinical Work

Across our findings, there are several overarching implications for research and clinical work. There appear to be misperceptions and inaccuracies about what contains caffeine and caffeine quantities, which can potentially impact reporting of such use and identification of problematic/unhealthy use. Our results point to the need for development and provision of psychoeducation for caffeine and CAC, such as cheat-sheet guidelines for caffeine content to be used during assessment. Implementation of strategies that raise awareness/mindfulness about consumption patterns is also indicated. Our results also point to the need for measure improvement and consolidation of substance-related terms. Lastly, it is crucial that we have a

better understanding of those who are at risk and those whose discrepancies/inconsistencies impact their placement in risk categories.

Study Limitations, Strengths and Future Directions

Sample Characteristics. As this was an exploratory, pilot study, the sample *N* of 50 was too small to have enough power to run certain statistics. If the sample had been larger, regression analyses would have been performed; for example, linear regression analyses could be used to examine whether demographic (e.g., gender, age) and other substance use risk (e.g., ADHD diagnosis, current smoker) variables predict number of caffeine, alcohol, and CAC discrepancies. Furthermore, the sample consisted of VCU undergraduate or graduate students that were regular caffeine and alcohol users. As a result, our sample characteristics potentially underrepresent certain groups and results may have limited transferability.

The decisions about sample size and inclusion of regular users were intentionally made to ensure that we would have enough data given the pilot nature of the study; one strength of the study is that we allowed for sufficient data to look for patterns of inconsistencies and misperceptions. Additionally, beginning these investigations with college students who are regular users appeared to be the best place to start given that they are a high risk group for caffeine (e.g., ED) (e.g., McIlvain et al., 2011) and CAC use (O'Brien, McCoy, Rhodes, Wagoner, & Wolfson, 2008; Mintel International Group Ltd., 2007).

These decisions, however, potentially limited our ability to generalize from the sample more broadly to VCU and other college students. Specifically, we found nonsignificant results that conflict with the current literature, such as no relationship between family history of substance problems and alcohol use (e.g., LaBrie et al., 2010). This potentially indicates that we did not have the power necessary to detect important differences; while we could have closer

examined problem use, the intent of our study was to look at the broader base of regular caffeine and alcohol users. We also found that we had a higher rate of official ADHD diagnosis compared to other studies (Green & Rabiner, 2012), which could be an indicator that our sample is not representative of typical college students.

Another possible limitation of our study is that the fatigue factor is influencing our results. In support of the possibility of the fatigue factor is the fact that we did find an order effect for TLFB alcohol frequency. The influence of fatigue effects for the other variables examined cannot entirely be ruled out as we may not have comprehensively looked for it; there could be order effects in relation to other important variables.

Additionally, we originally intended to sample with an equal number of males and females. We were unable to accomplish this as there were less available/eligible/interested males; in the interest of adhering to our recruitment timeline, we decided to attempt to reach 50 participants total, regardless of gender breakdown. As a result, our sample has a disproportionate amount of females represented compared to males as well as an unequal *N* for each randomization group. However, this gender representation does appear similar to that of the Student Health Center patient breakdown as indicated by the most recent report available; our sample appears to accurately reflect clinic proportions.

Recruiting from the VCU Student Health Center was decided upon for several potential benefits. There is increased potential for diversity with respect to different types of students, enabling conclusions drawn to be more generalizable among the community of students as a whole. Additionally, recruiting in a health care setting has applicability to the general health care settings where these assessments are often used and where research efforts are focused for measurement development.

Measures. One unavoidable weakness of the study design is that all of the measures are based on retrospective, self-report information. The exclusive reliance on self-report measures is potentially problematic because chances of bias and distortion on the part of the participant are increased (e.g., social desirability). However, this type of measurement was suitable to the study aims.

Since the purpose of this study was to examine reporting issues, the measures were carefully considered and conservatively chosen. The selection of assessments was based on consideration for the reported psychometric properties, need for comprehensive data to form comparisons, costs of data collection in terms of participant/researcher time, and feasibility of completion within this healthcare setting. Therefore, assessment questions were limited to those that have demonstrated clinical value and could contribute directly to the objectives of this study.

Additionally, the time period on which we focused our examination was only 30 days, limiting how far back we could assess, the generalizability of use patterns, and comparison with commonly used measures that include time periods of 3 months, one year, etc. One specific possible problem associated with this limited reference period is that, for erratic heavy drinking, shorter assessment periods (such as the 30 day period used in this study) can misclassify individuals (Chen & Kandel, 1995; Lemmens, Tan, & Knibbe, 1992). However, 30 days was chosen instead of a longer period because this is a timeframe often used in substance use assessments. Also, the shorter length of assessment helped save resources and allowed for minimal remembering/fatigue on the part of the participant, ensuring as accurate recall as possible.

Furthermore, as is the case with all measures, neither the TLFB nor the CAS are perfectly reliable and valid; there is no way to get guaranteed correct information on caffeine, alcohol, and

CAC use. However, the TLFB has high reliability and validity, and the timeframe chosen for the TLFB was done so as it allowed for direct comparison with the chosen CAS questions. One of the reasons we chose to include the qualitative interview was to provide a way to double-check information we were given during Phase 1, strengthening our study design and results.

Assessment of Inconsistencies. Another issue with our study design is that this method will probably not capture all caffeine use and misconceptions. It is reasonable to assume that asking participants to list the types of caffeine used is the best procedure to capture the misconceptions associated with drinks they assume will be caffeinated, but are not. However, assessing their use of caffeinated beverages that they don't realize contain caffeine is not as straightforward. The procedures employed potentially underestimated caffeine inconsistencies because we do not have TLFB-level detail for what participants believe does not contain caffeine that actually does. For this same reason, the caffeine TLFB results inherently underestimate the use of caffeine. Additionally, for the sake of tightly controlling procedures, the data collection only assessed caffeinated beverages and thus missed caffeine consumed from other sources.

Several ways to address these issues have been employed. During the alcohol TLFB, the researcher asked participants what they mixed alcohol with in order to get mixing data and information about caffeine not previously reported. Though no full TLFB data are available, participants were also asked about sources of caffeine that did not contain caffeine in an effort to capture a fuller picture of caffeine misreports. The assessment of caffeine misconceptions is not exhaustive and is possibly missing misconceptions, but with respect to misconceptions as they relate to mixing alcohol and caffeine, the data is adequate.

Qualitative Interview. There are several limitations associated with the qualitative interview. This was not a rigorous mixed-methods design, and qualitative interviews were used

primarily to supplement quantitative analyses and inform future research. Additional analyses are planned, with qualitative coding using a scoring software program. In addition, other areas of interest emerged during the conduct of this study that would have benefitted from further exploration with study participants. For example, the interview did not include questions about discrepancies associated with eligibility criteria. Additionally, the same researcher who conducted Phase 1 also conducted the interview, which could have biased these procedures.

In general, the inclusion of a qualitative interview is a strength, regardless of the fact that all domains were not comprehensively assessed. These efforts were aimed at expanding the potential knowledge base when compared to using only one method (Kazdin, 2007). The addition of a qualitative component also enhanced the design complexity in an attempt to better capture the full picture (Yoshikawa et al., 2008). This exploratory study benefited from elicitation of a greater level of detail and individual elaboration about response patterns.

Reactivity. The measures for this study are obtrusive and thus the participants were aware of, and possibly influenced by, the fact that they participated in an investigation. This fact has to be considered as possibly having influenced responses.

We attempted to limit reactivity in several ways. The design included efforts to keep the purpose of the study obscured from the participants (e.g., included filler items on the survey, assessment of additional substance use on the TLFB, and procedures for assessing knowledge of caffeine). Additionally, the use of a computerized survey may be less reactive. Since this was a low risk study without follow-ups, we chose to ensure anonymity by not collecting participant names or other identifying information (e.g., date of birth), which decreased chances of participant bias and reactivity.

Procedural Issues. As we attempted to avoid researcher bias by having RAs transcribe Phase 2 interviews, we were unable to offer the opportunity to participate in Phase 2 to many eligible participants when RAs were not available. We also encountered several technological issues during the course of our study. For instance, the internet stopped working during 8 (16%) CAS administrations (with 75% of these being from Group 2), creating the need for the researcher to administer CAS questions. Our *N* was too small to examine any potential differences between those who exclusively answered CAS questions by themselves and those who required RA assistance with administration.

From the standpoint of decreasing risk of experimenter expectancies, it would have been ideal if the researcher did not have direct contact with participants. However, having a separate researcher posed its own set of problems; additional training, compliance testing, monitoring, and payment for services would be taxing on resources that were not available. We considered attempting to enhance descriptive validity through the use of monitoring Phase 2 interviews via audio recording for later transcription by an outsider to maintain compliance. We ultimately decided the implications this would have for perceived anonymity for participation (for which many efforts have been made) outweighed the benefit.

Future Directions. This study is unique in its focus on potential discrepancies between interview and survey methods for measurement of caffeine, alcohol and CAC use in a sample of college students. Our findings present benchmark data on methodological issues associated with assessment of caffeine, alcohol and CAC use. Since this study was exploratory in nature, the employed design could not address all of the questions involving this line of research (e.g., issues with other substance use classes). As such, this study represents a first step in what needs to be a long line of research in this area. In general, more research is needed to better understand

caffeine, alcohol, and CAC use and factors that contribute to risk for abuse and negative consequences. Such research will help inform subsequent prevention and intervention efforts.

Large-Scale Study. This pilot study should be a platform for a larger project. Based on what was learned in this study, there are several changes that should be made to the study methods.

There are several logistical changes that should be made for a larger-scale study. In order to ensure all eligible participants are offered the opportunity for interview participation, more research assistants should be available for transcription of qualitative interviews (otherwise recording through audio file for transcription later should be included). Additionally, assurances of completion of REDCap survey without the internet disruption should be made (e.g., external internet source). With a larger study conducted over a longer period of time, the chances of recruiting the same person multiple times becomes more realistic; using a system to check for duplicates while protecting anonymity will likely be necessary.

Additionally, now that we have established that discrepancies exist, the next phase of research should include further elucidation of these discrepancies. Sampling in this population would benefit from recruitment from multiple settings (e.g., MCV Student Health Center). A larger *N* would allow for the following: additional subgroup populations to be examined, including comparison of developmental groups (i.e., emerging adults and adults investigated as distinct, developmental periods); order effects to be controlled for in the applicable analyses; and determination of any differences associated with assessment procedural variation (e.g., RA administration of CAS questions vs CAS completion without RA assistance). Data analyses should include prediction (i.e., regression analyses), including influence of caffeine misreports on discrepancies. The development of an additional TLFB for non-caffeinated beverages (as

opposed to just asking for them) would also help expose the number of actual misreports and magnitude of discrepancies. Lastly, the qualitative interview should be expanded to include asking participants about eligibility discrepancies.

Other studies. From results of this pilot study, there are indications of areas that need additional exploration apart from a larger-scale study. This study could be expanded to include focus on other substance use (e.g., tobacco, drugs) and/or other types of caffeine (e.g., food, medication with caffeine). Based on the information about caffeine inconsistencies gathered here, a quiz could be developed to help as part of patient education. An important line of research to correct discrepancies will begin with furthering understanding of discrepancies by trying to determine causes for misconceptions. As the present study data focused on college students, a high risk group where much attention has been focused on EDs, it will be important in future research to look at other groups, such as adolescents and non-college student young adults.

Our results indicate that there are several measurement issues that should be further investigated. Overall, substance use assessments should be created to allow for more specificity/variability (akin to the TLFB, but without the impracticality of the TLFB). For example, a measure among college students might benefit from asking about quantity and frequency of weekend (Friday through Sunday) use separate from use during the remainder of the week, as opposed to students estimating/averaging for the entire reference period. TLFB procedures specifically developed/validated for caffeine use should be created after more extensive research has been done on caffeine assessment, including establishing caffeine unit standardization (e.g., something equivalent to alcohol standard drinking units), finding a way to comprehensively assess caffeine (i.e., how to address all types of caffeine intake). This could

also include development of a self-administered version of the TLFB, a design which has been successfully applied to other domains (e.g., HIV risk; Carey et al., 2001) and might be beneficial in terms of resources. Our study was based on retrospective information, and future research should include prospective designs. For instance, daily prospective tracking has previously been applied to caffeine use measurement (Schliep et al., 2012) and, while this is a more cumbersome collection process, it is also a more rigorous assessment method. The inclusion of daily prospective measurement could potentially help to provide more specific information on discrepancies in reporting (e.g., survey versus daily records), inconsistencies/lack of knowledge about caffeine, and actual amounts being consumed. Additionally, more accurate measures should be developed for both CAC and caffeine as these currently do not have well-established measures.

The significant discrepancies that were found here indicate additional, targeted research is necessary in several areas. Regarding problematic drinking, there appears to be a need to improve assessment and potentially develop brief education to help mediate the discrepancy. Further psychoeducation efforts should include developing a supplemental aid for CAC use to be used during CAC assessment (analogous to ones used for alcohol). Both across substance abuse research and outside of this field, further research is needed to determine how often eligibility discrepancies occur as this speaks to validity of results.

Conclusion

To our knowledge, this is the first study to investigate caffeine and CAC discrepancies. Using a mixed methods approach, patterns of discrepant reporting across and within type of substance were found and affirmed the need to focus more attention on measure development, determination of causes of discrepancies, and dissemination of psychoeducation and strategies

for raising substance use awareness. Additionally, further research is warranted to better understand caffeine, alcohol and CAC use patterns and the factors that contribute to risk for abuse and negative consequences. This study and future studies like it will inform and guide development and testing of screening and assessment measures that can aid clinicians and researchers alike in the identification of college students at risk for development of caffeine and CAC problems.

References

References

Alanko, T. (1984). An overview of techniques and problems in the measurement of alcohol consumption. In: Smart, R.G.; Cappell, H.; Glaser, F.; Israel, Y.; Kalant, H.; Popham, R.E.; Schmidt, W.; and Sellers, E.M., eds. *Research Advances in Alcohol and Drug Problems, 8*, New York: Plenum Press, 209–226.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Amlung, M. (2013). Impulsivity and alcohol demand in relation to combined alcohol and caffeine use. *Exp Clin Psychopharmacol*, *21*(6), 467-74.

Arria, A. (2009). The Association between Energy Drinks and Later Drug Use among College Students. Presentation Summary; Cortland, New York: SUNY Youth Sports Institute.

Arria, A. M. et al. (2010). Increased alcohol consumption, nonmedical prescription drug use, and illicit drug use are associated with energy drink consumption among college students. *Journal of Addiction Medicine, 4*(2), 74-80.

Arria, A. M. et al. (2011). Energy drink consumption and increased risk for alcohol dependence. *Alcoholism Clinical and Experimental Research, 35*(2), 365-375.

Arria, A. M. & O'Brien, M. C. (2011). The "high" risk of energy drinks. *JAMA, 305*(6), 600-601.

Arria, A.M. et al. (2014). Evidence and knowledge gaps for the association between energy drink use and high-risk behaviors among adolescents and young adults. *Nutr Rev, 72*(Suppl 1), 87-97. doi: 10.1111/nure.12129.

Attila, S., Çakir, B. (2011). Energy-drink consumption in college students and associated factors. *Nutrition, 27*(3), 316-22. doi: 10.1016/j.nut.2010.02.008.

Azagba, S., Langille, D., Asbridge, M. (2014). An emerging adolescent health risk: caffeinated energy drink consumption patterns among high school students. Prev Med., 62, 54-9. doi: 10.1016/j.ypmed.2014.01.019.

Azcona et al. (1995). Evaluation of the central effects of alcohol and caffeine interaction. *Br J Clin Pharmacol, 40*(4), 393–400.

Babor, T.F., Stephens, R.S., & Marlatt, G.A. (1987). Verbal report methods in clinical research on alcoholism: response bias and its minimization, *Journal of Studies on Alcohol, 48*, 410--424.

Barratt, E. S. (1959). Anxiety and impulsiveness related to psychomotor efficiency. *Perceptual and Motor Skills, 9*(2), 191-198.

Bates, M. E., Bowden, S. C., & Barry, D. (2002). Neurocognitive impairment associated with alcohol use disorders: implications for treatment. *Experimental Clinical Psychopharmacology, 10*(3), 193-212.

Behavioral Pharmacology Research Unit (BPRU) (2003). BPRU Fact Sheet: Caffeine Dependence. Johns Hopkins Medicine. Retrieved at http://www.hopkinsmedicine.org/psychiatry/research/BPRU/docs/Caffeine_Dependence_Fact_Sheet.pdf.

Beeirness, Foss, & Vogel-Sprott (2004). Drinking on campus: Self-reports and breath tests. *Journal of Studies on Alcohol, 64*, 600-604.

Berger, L.K., Fendrich, M., Chen, H.Y., Arria, A.M., Cisler, R.A. (2011). Sociodemographic correlates of energy drink consumption with and without alcohol: results of a community survey. *Addict Behav., 36*, 516–519.

Bouchery, E.E., Harwood, H.J., Sacks, J.J., Simon, C.J., & Brewer, R.D. (2011). Economic costs of excessive alcohol consumption in the U.S., 2006. *American Journal of Preventative Medicine, 41*(5), 516-24. doi: 10.1016/j.amepre.2011.06.045

Bowling, A. (2005). Mode of questionnaire administration can have serious effects on data quality. *Journal of Public Health, 27*(3), 281-291.

Bowman, R.S., Stien, L.I., Newton, J.R. (1975). Measurement and interpretation of drinking behavior. *J Stud Alcohol* 36, 1154–1172.

Brache, K. & Stockwell, T. (2011). Drinking patterns and risk behaviors associated with combined alcohol and energy drink consumption in college drinkers. *Addict Behav., 36*(12), 1133-40. doi: 10.1016/j.addbeh.2011.07.003.

Bradley, K.A. et al. (2007). AUDIT-C as a brief screen for alcohol misuse in primary care. *Alcohol Clin Exp Res., 31*(7), 1208-17.

Brown University (n.d.). Alcohol & Your Body. Retrieved at: http://www.brown.edu/Student_Services/Health_Services/Health_Education/alcohol,_tobacco,_&_other_drugs/alcohol/alcohol_&_your_body.php

Bush K. et al. (1998). The AUDIT Alcohol Consumption Questions (AUDIT-C): An effective brief screening test for problem drinking. *Arch Internal Med, 3*, 1789-1795.

Caetano, R. & Babor, T.F. (2006). Diagnosis of alcohol dependence in epidemiological surveys: an epidemic of youthful alcohol dependence or a case of measurement error? *Addiction, 101*, Suppl 1, 111-4.

Cahalan, D. & Cisin, I.H. (1968), American drinking practices: Summary of findings from a national probability sample. *Q J Stud Alcohol* 29, 130–151.

Cahalan, D., Cisin, I.H., Crossley, H.M. (1969). *American Drinking Practices*. New Brunswick, NJ: Rutgers Center of Alcohol Studies.

Carey, M.P. et al. (2001). Assessing Sexual Risk Behavior with the Timeline Followback (TLFB) Approach: Continued Development and Psychometric Evaluation with Psychiatric Outpatients. *Int J STD AIDS*, *12*(6), 365–375.

Carey, K.B. et al. (2004). Temporal Stability of the Timeline Followback Interview for Alcohol and Drug Use with Psychiatric Outpatients. *J Stud Alcohol*, *65*(6), 774–781.

Centers for Disease Control and Prevention (2008). Youth Risk Behavior Surveillance --- United States, 2007. Retrieved at: http://www.cdc.gov/mmwr/preview/mmwrhtml/ss5704a1.htm

Centers for Disease Control and Prevention (2014). Caffeine and alcohol. Retrieved at: http://www.cdc.gov/alcohol/fact-sheets/caffeine-and-alcohol.htm.

Center for Science in the Public Interest (2014). Caffeine Content of Food & Drugs. Retrieved at: http://www.cspinet.org/new/cafchart.htm.

Chen, K. & Kandel, D.B. (1995). The natural history of drug use from adolescence to the mid-thirties in a general population sample. *Am J Public Health, 85*(1), 41-7.

Chung, T. & Martin, C.S. (2002). Concurrent and discriminant validity of DSM-IV symptoms of impaired control over alcohol consumption in adolescents. *Alcoholism, Clinical and Experimental Research, 26*, 485-492.

Clauson, K.A., Shields, K.M., McQueen, C.E., & Persad, N. (2008). Safety issues associated with commercially available energy drinks. *Journal of American Pharmacists Association, 48*:e55-63.

Cobb, C.O. et al. (2013). The role of caffeine in the alcohol consumption behaviors of college students. *Substance Abuse*.

Compton, W.M. et al. (1996). Comparing assessments of DSM-IV substance dependence disorders using CIDI-SAM and SCAN. *Drug Alcohol Depend, 41*(3), 179-87.

Compton, W.M., Thomas, Y.F., Stinson, F.S., Grant, B.F. (2007). Prevalence, correlates, disability, and comorbidity of DSM-IV drug abuse and dependence in the United States: results from the national epidemiologic survey on alcohol and related conditions. *Arch Gen Psychiatry, 64*(5), 566-76.

Cottler, L.B., Robins, L.N., Helzer, J.E. (1989). The reliability of the CIDI-SAM: a comprehensive substance abuse interview. *Br J Addict., 84*(7), 801-14.

Dawson, D.A. (2003). Methodological Issues in Measuring Alcohol Use. NIAAA. Retrieved at: http://pubs.niaaa.nih.gov/publications/arh27-1/18-29.htm

Del Boca, F.K.. Darkes, J., Greenbaum, P.E., & Goldman, M.S. (2004). Up close and personal: Temporal variability in the drinking of individual college students during their first year. *Journal of Consulting and Clinical Psychology, 72*, 155-164.

DeMarce, J.M., Burden, J.L., Lash, S.J., Stephens, R.S., Grambow, S.C. (2007). Convergent validity of the Timeline Followback for persons with comorbid psychiatric disorders engaged in residential substance use treatment. *Addict Behav., 32*(8), 1582-92.

de Beaurepaire, R. et al. (2007). Comparison of self-reports and biological measures for alcohol, tobacco, and illicit drugs consumption in psychiatric inpatients. *European Psychiatry, 22*, 540–548.

Devos-Comby, L. & Lange, J.E. (2008). Standardized measures of alcohol-related problems: a review of their use among college students. Psychol Addict Behav., 22(3), 349-61. doi: 10.1037/0893-164X.22.3.349.

Droste, N. et al. (2014). Combined alcohol and energy drink use: motivations as predictors of consumption patterns, risk of alcohol dependence, and experience of injury and aggression. *Alcohol Clin Exp Res., 38*(7), 2087-95. doi: 10.1111/acer.12438.

Embree, B.G. & Whitehead, P.C. (1993). Validity and reliability of self-reported drinking behavior: dealing with the problem of response bias. *Journal of Studies on Alcohol, 54*(3), 334–344. Retrieved at: http://www.jsad.com/doi/abs/10.15288/jsa.1993.54.334

Epstein, E. E. & McCrady, B. S. (2009). *A Cognitive-Behavioral Treatment Program for Overcoming Alcohol Problems: Therapist Guide*. New York: Oxford University Press.

Fals-Stewart, W., O'Farrell, T.J., Freitas, T.T., McFarlin, S.K., Rutigliano, P. (2000). The timeline followback reports of psychoactive substance use by drug-abusing patients: psychometric properties. *J Consult Clin Psychol., 68*(1), 134-44.

Ferreira S.E. et al. (2006). Effects of energy drink ingestion on alcohol intoxication. *Alcohol Clin Exp Res*, *30*(4), 598–605. doi: 10.1111/j.1530-0277.2006.00070.x.

Fishburne, J.W. & Brown, J.M. (n.d.). How do college students estimate their drinking? Comparing consumption patterns among quantity-frequency, graduated frequency, and timeline follow-back methods. *Journal of Alcohol & Drug Education*. Retrieved at: https://rti.org/pubs/fishburne.pdf

Foster, S.E., Vaughan, R.D., Foster, W.H., and Califano, J.A., Jr. (2003). Alcohol consumption and expenditures for underage drinking and adult excessive drinking. *JAMA: Journal of the American Medical Association* 289(8), 989–995.

Frary, C.D., Johnson, R.K., Wang, M.Q. (2005). Food sources and intakes of caffeine in the diets of persons in the United States. *Journal of the American Dietetic Association, 105*(1), 110-3.

Fulgoni V.L. 3rd, Keast D.R., Lieberman H.R. (2015). Trends in intake and sources of caffeine in the diets of US adults: 2001-2010. *American Journal* of *Clinical Nutrition*, 101(5), 1081-7. doi: 10.3945/ajcn.113.080077.

Grant, B.S. & Dawson, D.A. (2004). Age at onset of alcohol use and its association with DSM-IV alcohol abuse and dependence in the United States: results from the National Epidemiologic Survey, *Journal of Substance Abuse, 9*, 103-110.

Green, A.L. & Rabiner, D.L. (2012). What do we really know about ADHD in college students? *Neurotherapeutics, 9*(3), 559-68. doi: 10.1007/s13311-012-0127-8.

Greenfield, T.K., Midanik, L.T., Rogers, J.D. (2000). A 10-year national trend study of alcohol consumption, 1984-1995: is the period of declining drinking over? *Am J Public Health, 90*(1), 47-52.

Gruenewald, P.J.; Treno, A.J.; & Mitchell, P.R. (1996). Drinking patterns and drinking behaviors: Theoretical models of risky acts. *Contemporary Drug Problems, 23*, 407–440.

Gruenewald, P.J., Johnson, F.W., Light, J.M., & Salz, R.F. (2003). Drinking to extremes: Theoretical and empirical analyses of peak drinking levels among college students. Journal of Studies on Alcohol, 64, 817–824.

Hammersley, R. (1994). A digest of memory phenomena for addiction research. *Addiction, 89*(3), 283-93.

Harrell, P.T. & Juliano, L.M. (2009). Caffeine expectancies influence the subjective and behavioral effects of caffeine. *Psychopharmacology (Berl)., 207*(2), 335-42. doi: 10.1007/s00213-009-1658-5.

Hasin, D. S., Stinson, F. S., Ogburn, E., & Grant, B. F. (2007). Prevalence, correlates, disability, and comorbidity of DSM-IV alcohol abuse and dependence in the United States: Results from the National Epidemiologic Survey on Alcohol and Related Conditions. *JAMA Psychiatry, 64*(7), 830-42.

Heatherton, T.F., Kozlowski, L.T., Frecker, R.C., Fagerström, K.O. (1991). The Fagerström Test for Nicotine Dependence: a revision of the Fagerström Tolerance Questionnaire. *Br J Addict., 86*(9), 1119-27.

Heckman, M.A., Weil, J., Gonzalez de Mejia E. (2010). Caffeine (1, 3, 7-trimethylxanthine) in foods: a comprehensive review on consumption, functionality, safety, and regulatory matters. *Journal of Food Science, 75*(3):R77-87. doi: 10.1111/j.1750-3841.2010.01561.x.

Hughes, J.R., Oliveto, A.H., Liguori, A., Carpenter, J., Howard, T. (1998). Endorsement of DSM-IV dependence criteria among caffeine users. *Drug Alcohol Depend.,52*(2), 99-107.

Hull, L., Dillon, P.M., O'Connell, M.M., Chitnavis, P., Svikis, D.S. (2011). Is use of caffeinated energy drinks associated with higher levels of tobacco, alcohol and other drug use in American students? *Eur Neuropsychopharm, 21*(2), S168–S169.

Huntley, E.D., Juliano, L.M. (2012). Caffeine Expectancy Questionnaire (CaffEQ): construction, psychometric properties, and associations with caffeine use, caffeine dependence, and other related variables. *Psychol Assess., 24*(3), 592-607. doi: 10.1037/a0026417.

Ibrahim, A.K. (2013). A systematic review of studies of depression prevalence in university students. *J Psychiatr Res, 47*(3), 391-400. doi: 10.1016/j.jpsychires.2012.11.015. Epub 2012 Dec 20.

James, J.E. et al. (1988). Biochemical validation of self-reported caffeine consumption during caffeine fading. *Journal of Behavioral Medicine, 10,* 15-30.

James, J.E. (1989). Self-report reliability and symptomatology of habitual caffeine consumption. *British Journal of Clinical Pharmacology, 27,* 507-514.

James, J.E. (1997). *Understanding caffeine: a biobehavioral analysis*. Behavioral Medicine and Health Psychology Series, Vol. 2 (ed. J.R. Turner). Sage Publications: Thousand Oaks, CA.

Johnston, L. D. et al. (1998). National Survey Results on Drug Use From Monitoring The Future Study, 1975-1997. Volume I: Secondary School Students. Volume II: College Students and Young Adults. (Vol. I NIH Publication no. 98-4345) (Vol. II NIH Publication no. 98-4346). Washington DC: U.S. Department of Health and Human Services.

Jones, H.A. & Lejuez, C.W. (2005). Personality correlates of caffeine dependence: the role of sensation seeking, impulsivity, and risk taking. *Exp Clin Psychopharmacol, 13*(3), 259-66.

Juliano, L.M., Evatt, D.P., Richards, B.D., Griffiths, R.R. (2012a). Characterization of individuals seeking treatment for caffeine dependence. Psychol Addict Behav., 26(4), 948-54. doi: 10.1037/a0027246.

Juliano, L.M., Griffiths, R.R. (2004). A critical review of caffeine withdrawal: empirical validation of symptoms and signs, incidence, severity, and associated features. *Psychopharmacology (Berl), 176*(1), 1-29.

Juliano, L.M., Huntley, E.D., Harrell, P.T., Westerman, A.T. (2012b). Development of the caffeine withdrawal symptom questionnaire: caffeine withdrawal symptoms cluster into 7 factors. *Drug Alcohol Depend, 124*(3), 229-234.

Kahler et al. (2004). Mapping the continuum of alcohol problems in college students: A Rasch model analysis. *Psychology of Addictive Behaviors, 18*, 322-333.

Karriker-Jaffe, K.J., Witbrodt, J., Greenfield, T.K. (2015). Refining measures of alcohol problems for general population surveys. *Alcohol Clin Exp Res., 39*(2), 363-70. doi: 10.1111/acer.12627.

Kazdin, A.E. (2007). Research Design in Clinical Psychology, Pearson, 4th Edition.

Khavari, K.A. & Farber, P.D. (1978). A profile instrument for the quantification and assessment of alcohol consumption: The Khavari Alcohol Test. *J Stud Alcohol* 39, 1525–1539.

Kroenke K, Spitzer R L, Williams J B (2001). The PHQ-9: validity of a brief depression severity measure. *Journal of General Internal Medicine, 16*(9), 606-613.

Kroenke K, Spitzer R L, Williams J B (2003). The Patient Health Questionnaire-2: validity of a two-item depression screener. *Med Care, 41*(11), 1284-92.

Kurtz, A.M., Leong, J., Anand, M., Dargush, A.E., Shah, S.A. (2013). Effects of caffeinated versus decaffeinated energy shots on blood pressure and heart rate in healthy young volunteers. *Pharmacotherapy, 33*(8):779-86. doi: 10.1002/phar.1296.

Labrie, J.W. et al. (2010). Family history of alcohol abuse associated with problematic drinking among college students. *Addict Behav, 35*(7), 721-5. doi: 10.1016/j.addbeh.2010.03.009. Epub 2010 Mar 16.

Lange, J.E. & Voas, R.B. (2001). Defining binge drinking quantities through resulting blood alcohol concentration. *Psychology of Addictive Behaviors, 15*, 310-316.

Lemmens, P., Tan, E.S., Knibbe, R.A. (1992). Measuring quantity and frequency of drinking in a general population survey: a comparison of five indices. *J Stud Alcohol., 53*(5), 476-86.

M. Shanken Communications, Inc. (2009). *The U.S. Beer Market: Impact Databank Review and Forecast*. New York, NY: M. Shanken Communications, Inc., 533.

MacKillop, J et al. (2012). Initial development of a measure of expectancies for combinations of alcohol and caffeine: the Caffeine + Alcohol Combined Effects Questionnaire (CACEQ). *Exp Clin Psychopharmacol., 20*(6), 466-72. doi: 10.1037/a0030539.

Malinauskas, B.M., Aeby, V.G., Overton, R.F., Carpenter-Aeby, T., Barber-Heidal, K. (2007). A survey of energy drink consumption patterns among college students. *Nutr J., 6*, 35.

Mayo Clinic (2014). Caffeine: How much is too much? Retrieved at: http://www.mayoclinic.org/healthy-lifestyle/nutrition-and-healthy-eating/in-depth/caffeine/art-20045678.

McIlvain, G.E., Noland, M.P., & Bickel, R. (2011). Caffeine Consumption Patterns and Beliefs of College Freshmen. *American Journal of Health Education, 42*(4), 235-244.

McLellan, A.T. et al. (1992). The Fifth Edition of the Addiction Severity Index. *J Subst Abuse Treat., 9*(3), 199-213.

McLellan, T.M., & Lieberman, H.R. (2012). Do energy drinks contain active components other than caffeine? *Nutrition Reviews, 70*(12), 730-744.

Meredith, S.E., Juliano, L.M., Hughes, J.R., Griffiths, R.R. (2013). Caffeine Use Disorder: A Comprehensive Review and Research Agenda. *Journal of Caffeine Research, 3*(3), 114-130.

Merikangas, K.R., et al. (2010). Lifetime prevalence of mental disorders in U.S. adolescents: results from the National Comorbidity Survey Replication–Adolescent Supplement (NCS-A). *J Am Acad Child Adolesc Psychiatry, 49*(10), 980-9. PMID: 20855043

Merikangas, K.R., McClair, V.L. (2012). Epidemiology of substance use disorders. Hum Genet., 131(6), 779-89. doi: 10.1007/s00439-012-1168-0.

Midanik, L.T. & Hines, A.M. (1991). 'Unstandard' ways of answering standard questions: protocol analysis in alcohol survey research. Drug Alcohol Depend., 27(3), 245-52.

Midanik, L.T., Greenfield, T.K., Bond, J. (2007). Addiction sciences and its psychometrics: the measurement of alcohol-related problems. *Addiction, 102*(11), 1701-10.

Miller, K.E. (2008). Energy drinks, race, and problem behaviors among college students. J Adolesc Health, 43(5), 490-7. doi: 10.1016/j.jadohealth.2008.03.003.

Mintel International Group Ltd. (2007). *Energy Drinks*. Chicago, IL: Mintel International Group Ltd.

Miyake, E.R., Marmorstein, N.R.. (2015). Energy drink consumption and later alcohol use among early adolescents. Addict Behav., 4, 60-5. doi: 10.1016/j.addbeh.2014.12.009.

National Institute on Alcohol Abuse and Alcoholism (NIAAA) (2005). Screening and Assessing. College Drinking – Changing the Culture. Retrieved at: http://www.collegedrinkingprevention.gov/niaaacollegematerials/trainingmanual/module_2.aspx

National Institute on Alcohol Abuse and Alcoholism (NIAAA) (2005). Module 1: Epidemiology of Alcohol Problems in the United States, Social Work Education for the Prevention and Treatment of Alcohol Use Disorders. Retrieved at: http://pubs.niaaa.nih.gov/publications/Social/Module1Epidemiology/Module1.html

National Institute on Alcohol Abuse and Alcoholism (NIAAA) (2008). Alcohol and other drugs, Alcohol Alert, NIAAA, *76*. Retrieved at: http://pubs.niaaa.nih.gov/publications/AA76/AA76.htm

National Institute of Alcohol Abuse and Alcoholism (NIAAA) (n.d.a). College drinking. Retrieved at: http://niaaa.nih.gov/alcohol-health/special-populations-co-occurring-disorders/college-drinking.

National Institute on Alcohol Abuse and Alcoholism (NIAAA) (n.d.b). What's "at-risk" or "heavy" drinking? Rethinking Drinking: Alcohol & your health. Retrieved at: http://www.rethinkingdrinking.niaaa.nih.gov/How-much-is-too-much/Is-your-drinking-pattern-risky/whats-At-Risk-Or-Heavy-drinking.aspx

National Institute on Alcohol Abuse and Alcoholism (NIAAA) (n.d.c). Drinking Levels Defined. Retrieved at: http://www.niaaa.nih.gov/alcohol-health/overview-alcohol-consumption/moderate-binge-drinking

National Institute of Health (2015). Caffeine in the diet. Retrieved at: http://www.nlm.nih.gov/medlineplus/ency/article/002445.htm.

Nelson, C.B., Little, R.J., Heath, A.C., & Kessler, R.C. (1996). Patterns of DSM-III-R alcohol dependence symptom progression in a general population survey. Psychological Medicine, 26 449–460.

Norberg, M.M., Mackenzie, J., Copeland, J. (2012). Quantifying cannabis use with the timeline followback approach: a psychometric evaluation. *Drug Alcohol Depend.*, *121*(3), 247-52. doi: 10.1016/j.drugalcdep.2011.09.007.

O'Brien, M.C., McCoy, T.P., Rhodes, S.D., Wagoner, A., Wolfson, M. (2008). Caffeinated cocktails: energy drink consumption, high-risk drinking, and alcohol-related consequences among college students. *Academy of Emergency Medicine*, *15*(5), 453-460.

O'Neill, S.E. & Sher, K.J. (2000). Physiological alcohol dependence symptoms in early adulthood: A longitudinal perspective. *Experimental and Clinical Psychopharmacology, 8*, 493–508.

Panza, G.A., Weinstock, J., Ash, G.I., Pescatello, L.S. (2012). Psychometric Evaluation of the Timeline Followback for Exercise among College Students. *Psychol Sport Exerc., 13*(6), 779-788.

Perkins, H.W., DeJong, W., & Linkenbach, J. (2001). Estimated blood alcohol levels reached by "binge" and "non-binge" drinkers: A survey of young adults in Montana. *Psychology of Addictive Behaviors, 15*, 317-320.

Polak, K.M., Hancock, L., Svikis, D. (n.d.a). [Misreporting of caffeine use - collected 2014]. Unpublished raw data.

Polak, K.M., Hancock, L., Svikis, D. (n.d.b). [Knowledge about caffeine use - collected 2015]. Unpublished raw data.

Polak, K.M. et al. (2016). *Energy Drink Patterns and Other Substance Use in Male and Female Eighth, Tenth, and Twelfth Graders*. Manuscript submitted for publication.

Polich, J.M., & Orvis, B.R. (1979). *Alcohol Problems: Patterns and Prevalence in the U. S. Air Force*. Santa Monica, CA: Rand.

Polich, J.M., Armor, D.J., and Braiker, H.B. (1981). *The Course of Alcoholism: Four Years After Treatment*. New York: John Wiley & Sons.

Read, J.P., Kahler, C.W., Strong, D.R. & Colder, C.R. (2006). Development and preliminary validation of the Young Adult Alcohol Consequences Questionnaire. *Journal of Studies on Alcohol, J Stud Alcohol, 67*(1), 169-177.

Read et al. (2007). Predicting functional outcomes among college drinkers: Reliability and predictive validity of the Young Adult Alcohol Consequences Questionnaire. *Addictive Behaviors, 32,* 2597-2610.

Read, J.P. et al. (2008). Beyond the "binge" threshold: Heavy drinking patterns and their association with alcohol involvement indices in college students. *Addictive Behaviors, 33*, 225-234.

Read, J.P., Wardell, J.D., Bachrach, R.L. (2013). Drinking consequence types in the first college semester differentially predict drinking the following year. *Addictive Behaviors, 38*, 1464-1471.

Reinert, D.F. & Allen, J.P. (2002). The Alcohol Use Disorders Identification Test (AUDIT): a review of recent research. *Alcohol Clin Exp Res., 26*(2), 272-9.

Reissig, C. J., Strain, E. C., & Griffiths, R. R. (2009). Caffeinated energy drinks-a growing problem. *Drug and Alcohol Dependence, 99*(1-3), 1-10.

Robinson, S. M., Sobell, L. C., Sobell, M. B., Leo, G. I. (2014). Reliability of the Timeline Followback for Cocaine, Cannabis, and Cigarette Use.. *Psychology of Addictive Behaviors, 28*(1), 154-162.
Available at: http://nsuworks.nova.edu/cps_facarticles/539

Room, R. (1990). Measuring alcohol consumption in the United States: Methods and rationales. In: Kozlowski, L.T.; Annis, H.M.; Cappell, H.D.; Glaser, F.B.; Goodstadt, M.S.; Israel, Y.; Kalant, H.; Sellers, E.M.; and Vingilis, E.R., eds. *Research Advances in Alcohol and Drug Problems, 10*. New York: Plenum Press, 39–80.

Russell, M. (1994). New assessment tools for drinking in pregnancy: T-ACE, TWEAK, and others. *Alcohol Health and Research World, 18*(1), 55-61

Russell, M. et al. (1997). Test–retest reliability of the Cognitive Lifetime Drinking History. *Am J Epidemiol, 146*(11), 975–981.

Samet, S. et al. (2007). Assessing Addiction: Concepts and Instruments. *Addict Sci Clin Pract., 4*(1), 19–31.

SAMHSA (1999). Summary of Findings From the 1999 National Household Survey on Drug Abuse. Rockville, MD: Department of Health and Human Services.

SAMHSA (2006). The role of biomarkers in the treatment of alcohol use disorders. Retrieved at: http://www.kap.samhsa.gov/products/manuals/advisory/pdfs/0609_biomarkers.pdf.

SAMHSA (2013a). Update on Emergency Department Visits Involving Energy Drinks: A Continuing Public Health Concer. The DAWN Report. Retrieved at: http://archive.samhsa.gov/data/2k13/DAWN126/sr126-energy-drinks-use.pdf.

SAMHSA (2013b). National Survey on Drug Use and Health (NSDUH). Table 6.88B—Alcohol Use in the Past Month among Persons Aged 18 to 22, by College Enrollment Status and Demographic Characteristics: Percentages, 2012 and 2013. Retrieved at: http://www.samhsa.gov/data/sites/default/files/NSDUH-DetTabsPDFWHTML2013/Web/HTML/NSDUH-DetTabsSect6peTabs55to107-2013.htm#tab6.88b

SAMHSA (2013c). Results from the 2013 National Survey on Drug Use and Health: Summary of National Finding. Center for Behavioral Health Statistics and Quality, US Department of Health and Human Services. Retrieved at: http://www.samhsa.gov/data/sites/default/files/NSDUHresultsPDFWHTML2013/Web/NSDUHresults2013.pdf

SAMHSA, 2013d. National Survey on Drug Use and Health (NSDUH). Table 2.41B—Alcohol Use in Lifetime, Past Year, and Past Month among Persons Aged 18 or Older, by Demographic Characteristics: Percentages, 2012 and 2013. Retrieved at: http://www.samhsa.gov/data/sites/default/files/NSDUH-DetTabsPDFWHTML2013/Web/HTML/NSDUH-DetTabsSect2peTabs1to42-2013.htm#tab2.41b

Schliep, K.C. et al. (2013). Validation of different instruments for caffeine measurement among premenopausal women in the BioCycle study. *Am J Epidemiol, 177*(7), 690-699.

Setchell, K.D.R. et al. (1987). Rapid high-performance liquid chromatography assay for salivary and serum caffeine following an oral load: An indicator of liver function. *Journal of Chromatography, 385*, 267-274.

Simon, M. & Mosher, J. (2007). *Alcohol, Energy Drinks, and Youth: A Dangerous Mix*. San Rafael, CA: Marin Institute.

Single, E. & Wortley, S. (1994). A comparison of alternative measures of alcohol consumption in the Canadian National Survey of alcohol and drug use. *Addiction, 89*(4), 395-9.

Skewes, M.C., Decou, C.R., Gonzalez, V.M. (2013). Energy drink use, problem drinking and drinking motives in a diverse sample of Alaskan college students. *Int J Circumpolar Health*, 72. doi: 10.3402/ijch.v72i0.21204.

Skinner, H.A. & Sheu, W.J. (1982). Reliability of alcohol use indices: The Lifetime Drinking History and the MAST. *J Stud Alcohol,* 43,1157–1170.

Smith, P.C. et al. (2009). Primary care validation of a single-question alcohol screening test. *Journal of General Internal Medicine, 24*(7), 783-788.

Sobell, L.C. and Sobell, M.B. (1992). Timeline Follow–back: A technique for assessing self–reported alcohol consumption. In: Litten, R.Z., and Allen, J.P., eds. *Measuring Alcohol Consumption: Psychosocial and Biological Methods*. Totowa, NJ: Humana Press, 41–72.

Sobell, L.C., and Sobell, M.B. (1995). Alcohol consumption measures. In: Allen, J.P., and Columbus, M., eds. *Assessing Alcohol Problems: A Guide for Clinicians and Researchers*. Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism, 55–73.

Sobell, L.C. & Sobell, M.S. (2004). Alcohol consumption measures. Retrieved at: http://pubs.niaaa.nih.gov/publications/AssessingAlcohol/measures.htm

Spierer, D.K., Blanding, N., Santella, A. (2014). Energy drink consumption and associated health behaviors among university students in an urban setting. J Community Health., 39(1), 132-8.

Spitzer R., Kroenke, K., Williams, J. (1999). Validation and utility of a self-report Version of PRIME-MD: the PHQ Primary Care Study. *Journal of the American Medical Association, 282*, 1737-1744.

Steinberg, L., Sharp, C., Stanford, M.S., Tharp, A.T. (2013). New tricks for an old measure: the development of the Barratt Impulsiveness Scale-Brief (BIS-Brief). *Psychol Assess., 25*(1), 216-26. doi: 10.1037/a0030550.

Straus, R. & Bacon, S.D. (1953). *Drinking in College*. New Haven, CT: Yale University Press.

Striley, C.W., Khan, S.R. (2014). Review of the energy drink literature from 2013: findings continue to support most risk from mixing with alcohol. *Current Opinion in Psychiatry, 27*(4), 263-8. doi: 10.1097/YCO.0000000000000070.

Striley, C.W. et al. (2011). Evaluating dependence criteria for caffeine. *J Caffeine Res. 1,* 219-225.

Strobbe, S. (2014). Prevention and screening, brief intervention, and referral to treatment for substance use in primary care. *Primary Care, 41*(2), 185-213. doi: 10.1016/j.pop.2014.02.002. Epub 2014 Mar 21.

Svikis, D.S., Berger, N., Haug, N.A., Griffiths, R.R.. (2005). Caffeine dependence in combination with a family history of alcoholism as a predictor of continued use of caffeine during pregnancy. *Am J Psychiatry*, *162*(12), 2344-51.

Svikis, D.S. et al. (2012). Randomized multi-site trial of the Job Seekers' Workshop in patients with substance use disorders. *Drug Alcohol Depend.*, *120*(1-3), 55-64. doi: 10.1016/j.drugalcdep.2011.06.024.

Swendsen J. et al. (2012). Use and abuse of alcohol and illicit drugs in US adolescents: results of the National Comorbidity Survey-Adolescent Supplement. *Arch Gen Psychiatry*, *69*(4), 390-8. doi: 10.1001/archgenpsychiatry.2011.1503.

Terry-McElrath, Y.M., O'Malley, P.M. & Johnston, L.D. (2014). Energy drinks, soft drinks, and substance use among United States secondary school students. *Journal of Addiction Medicine, 8*(1), 6-13.

Thombs D.L. et al. (2010). Event-level analyses of energy drink consumption and alcohol intoxication in bar patrons. *Addictive behaviors*, *35*(4), 325–330.

Thombs, D. et al. (2011). Is there a misplaced focus on AmED? Associations between caffeine mixers and bar patron intoxication. *Drug Alcohol Depend.*, *116*(1-3), 31-6. doi: 10.1016/j.drugalcdep.2010.11.014.

Tourangeau, R. & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin, 133*(5), 859-883.

Trapp, G.S., Allen, K.L., O'Sullivan, T., Robinson, M., Jacoby, P., Oddy, W.H.. (2014). Energy drink consumption among young Australian adults: associations with alcohol and illicit drug use. *Drug Alcohol Depend., 134*, 30-7. doi: 10.1016/j.drugalcdep.2013.09.006.

University of Virginia (n.d.) Red Bull in brief. *Darden Business Publishing*. Retrieved at: http://faculty.darden.virginia.edu/farrisp/jnjcourseii/general/documents/redbullm-0663_condensed.pdf.

Upadhyaya, H.P. (2005). Attention-deficit/hyperactivity disorder, medication treatment, and substance use patterns among adolescents and young adults. *J Child Adolesc Psychopharmacol*, *15*(5), 799-809.

US Food and Drug Administration (2012).Energy "Drinks" and Supplements: Investigations of Adverse Event Reports. Retrieved at: http://www.fda.gov/Food/NewsEvents/ucm328536.htm

US Food and Drug Administration (2015). *CFSAN Adverse Event Reporting System Voluntary Reports on Red Bull Energy Drink*. Food and Drug and Administration Center for Food Safety and Applied Nutrition. Retrieved at: http://www.fda.gov/downloads/AboutFDA/CentersOffices/OfficeofFoods/CFSAN/CFSANFOIAElectronicReadingRoom/UCM328525.pdf

US Food and Drug Administration (2014). FDA Consumer Advice on Powdered Pure Caffeine. Retrieved at: http://www.fda.gov/Food/RecallsOutbreaksEmergencies/SafetyAlertsAdvisories/ucm405787.htm.

US Food and Drug Administration (2015). Serious Concerns Over Alcoholic Beverages with Added Caffeine, US Department of Health and Human Services. Retrieved at: http://www.fda.gov/ForConsumers/ConsumerUpdates/ucm233987.htm

Volkow, N.D. (2010). Prescription Drug Abuse, Presentation at Congressional Caucus. Retrieved at: https://www.drugabuse.gov/about-nida/legislative-activities/testimony-to-congress/2010/09/prescription-drug-abuse

Wechsler H., Nelson T.F. (2008). What We Have Learned From the Harvard School of Public Health College Alcohol Study: Focusing Attention on College Student Alcohol Consumption and the Environmental Conditions That Promote It. *Journal of Studies on Alcohol and Drugs, 69*(4), 481-490.

Whelan, T. J. (2007). Anonymity and confidentiality: Do survey respondents know the difference? Poster presented at the 30th annual meeting of the Society of Southeastern Social Psychologists, Durham, NC.

White, A.M. (2004). What happened? Alcohol, memory blackouts, and the brain. Retrieved at: http://pubs.niaaa.nih.gov/publications/arh27-2/186-196.htm

White, A.M., Kraus, C.L. & Swartzwelder, H.S. (2006). Many college freshmen drink at levels far beyond the binge threshold. *Alcoholism, Clinical and Experimental Research, 30*, 1006–1022.

Wilsnack, R.W. (2009). Gender and alcohol consumption: patterns from the multinational GENACIS project. *Addiction, 104*(9), 1487-500. doi: 10.1111/j.1360-0443.2009.02696.x.

Wish, E. et al. (1997). The validity of self–reports of drug use at treatment admission and at followup: Comparisons with urinalysis and hair assay. *NIDA Res Monogr, 167*, 200–226.

Woodyard, C.D. & Hallam, J.S. (2010). Differences in college student typical drinking and celebration drinking. *J Am Coll Health, 58*(6), 533-8.

Yoshikawa, H. et al. (2008). Mixing qualitative and quantitative research in developmental science: Uses and methodological choices. *Developmental Psychology, 44*(2), 344-354.

Varvil-Weld, L. et al. (2013). Examining the relationship between alcohol-energy drink risk profiles and high-risk drinking behaviors. *Alcohol Clin Exp Res., 37*(8), 1410-6. doi: 10.1111/acer.12102.

Zhang et al. (2013). Validity and reliability of Patient Health Questionnaire-9 and Patient Health Questionnaire-2 to screen for depression among college students in China. *Asia Pac Psychiatry, 5*(4), 268-75. doi: 10.1111/appy.12103.

Zygler-Katz, E. et al. (1984). Relationship between caffeine concentrations in plasma and saliva. *Clin. Pharmac. Ther., 36*, 133-137.

Appendix

I. Computer-Assisted Survey (CAS) Items

Demographics.

1. How old are you? (free response option)
2. What is your gender? (response options: female, male, other [with free response option])
3. Of what race do you consider yourself? (response options: White, Black, American Indian, Alaskan Native, Asian/Pacific Islander, Hispanic, Mixed, Don't Know)
4. How much education have you completed (GED = 12 years): (free response option)
5. What is your current student status? (response options: Full time or Part time)
6. Do you belong to a fraternity or sorority? (response options: Yes or No)

Substance Use Questions.

Recent Alcohol Use (Past 30 days)

7. During the past 30 days, on how many days did you have any beverage containing alcohol (including beer, wine, or liquor)? (free response option)
8. During the past 30 days, on days when you did drink alcohol, how may drinks did you usually have? (free response option)
9. During the past 30 days, on how many days did you have X or more drinks (X = 4 for female, 5 for male)

Recent Caffeine use (Past 30 Days)

10. During the past 30 days, on how many days did you have any beverage containing caffeine? (free response option)

11. During the past 30 days, on days when you did drink caffeinated beverages, how may drinks did you usually have? (free response option)
12. In the past 30 days, how many days per week in a typical week did you have coffee? (free response option)
13. During the past 30 days, on days when you did drink coffee, how may drinks did you usually have? (free response option)
14. In the past 30 days, how many days per week in a typical week did you have caffeinated tea? (free response option)
15. During the past 30 days, on days when you did drink caffeinated tea, how may drinks did you usually have? (free response option)
16. In the past 30 days, how many days per week in a typical week did you have caffeinated soda? (free response option)
17. During the past 30 days, on days when you did drink caffeinated soda, how may drinks did you usually have? (free response option)
18. In the past 30 days, how many days per week in a typical week did you have caffeinated energy drinks/shots? (free response option)
19. During the past 30 days, on days when you did drink caffeinated energy drinks/shots, how may drinks/shots did you usually have? (free response option)
20. In the past 30 days, how many servings of caffeine did you have on a typical day (1 serving = 1, 8oz cup of regular coffee)? (free response option)

Recent caffeine and alcohol combined use (past 30 days)

21. During the past 30 days, how many days did you drink either a caffeinated alcoholic product or a caffeinated beverage in combination with alcohol? (free response option)

22. During the past 30 days, on days when you did drink caffeine combined with alcohol, how many drinks did you usually have? (free response option)

Caffeine and alcohol combined (past year)

23. In the past year, how often did you drink alcohol mixed with caffeine? (response options: none, less than once a month, 1 to 3 times a month, 1 to 3 times a week, and most days)
24. In the past year, how many combined caffeine and alcohol drinks did you have on a typical day? (free response option)

Alcohol Problems

25. How often do you have a drink containing alcohol? (response options: Never, Monthly or less, 2−4 times a month, 2−3 times a week, or 4 or more times a week)
26. How many standard drinks containing alcohol do you have on a typical day when drinking? (response options: 1 or 2, 3 or 4, 5 or 6, 7 to 9, or 10 or more)
27. How often do you have six or more drinks on one occasion? (response options: Never, Less than monthly, Monthly, Weekly, or Daily or almost daily)
28. In your lifetime, has a friend or family member told you about things you said or did while you were drinking that you could not remember? (response options: Yes or No)
29. In the past 30 days, how many days have you said or done things while you were drinking that you could not remember that a friend or family member told you about? (free response option)

Other Substance Use

30. In the past 30 days, how many days have you used cannabis (marijuana, pot, weed, hash, etc.)? (free response option)

31. In the past 30 days, how many days have you used sedatives (medications for anxiety or sleep such as Xanax, Ativan, Klonopin, Valium, GHB, Rohypnol/roofies, etc.)? (free response option)
32. In the past 30 days, how many days have you used stimulants (speed, diet pills, Adderall, Ritalin, methamphetamine, crystal meth, etc.)? (free response option)
33. In the past 30 days, how many days have you used cocaine (coke, crack, etc.)? (free response option)
34. In the past 30 days, how many days have you used opioids (heroin, morphine, methadone, codeine, buprenorphine, Suboxone, OxyContin, Dilaudid, Percocet, Vicodin, etc.)? (free response option)
35. In the past 30 days, how many days have you used other drugs (for example, synthetic marijuana (spice))? (free response option)
36. In the past 30 days, how many days did you mix drugs and alcohol? (free response option)
37. In the past 30 days, how many days did you mix ADHD/other prescription drugs and alcohol? (free response option)

Substance Use Problems and Treatment History

38. In your lifetime, have you ever had problems due to use of alcohol or other drugs? (response options: Yes or No)
39. In your lifetime, have you ever received substance abuse treatment, including AA/NA? (response options: Yes or No)

Nicotine/Tobacco Use

40. Do you currently smoke cigarettes? (response options: Yes or No)

41. In the past 30 days, how many days have you used e-cigarettes/personal vaporizers? (free response option)
42. In the past 30 days, how many days have you used e-cigarettes/personal vaporizers to use marijuana or chemical extracts of marijuana (wax, THC oil, etc.)? (free response option)

Prescribed Medications

43. Are you taking any medications that were prescribed to you? (response options: No, I'm not on any prescription medications; Yes, I take prescription medications that a psychiatrist or primary care doctor prescribed for a condition like ADHD/ADD, depression, anxiety, etc.; Yes, I take prescription medications daily for a non-psychiatric condition (e.g., allergies, diabetes, etc.); Yes, I take medications for both of the categories of reasons listed above)

ADHD

44. Do you make careless mistakes, have trouble keeping your attention focused, or have difficulty organizing/planning most of the time? (response options: Yes or No)
45. Do you interrupt other people when they are talking, talk a lot, or feel like you have a lot of energy most of the time? (response options: Yes or No)
46. Have you ever been officially diagnosed with ADHD or ADD? (Yes, I have been diagnosed with ADHD or ADD; I've never been diagnosed with ADHD or ADD, but I think I might have ADHD or ADD; No, I've never been diagnosed with ADHD or ADD, and I don't think I have ADHD or ADD)

Depression

Over the last 2 weeks, how often have you been bothered by any of the following problems?

47. Little interest or pleasure in doing things (response options: not at all, several days, more than half the days, or nearly every day)
48. Feeling down, depressed, or hopeless (response options: not at all, several days, more than half the days, or nearly every day)

Family history

49. Do you think any of your first degree biological relatives (i.e., mother, father, brothers, and sisters) ever had a problem with alcohol? (response options: Yes, No, and Don't Know).
50. Do you think any of your first degree biological relatives (i.e., mother, father, brothers, and sisters) ever had a problem with drugs? (response options: Yes, No, and Don't Know).

Impulsivity

51. I plan tasks carefully (response options: rarely/never, occasionally, often, almost always/always)
52. I do things without thinking (response options: rarely/never, occasionally, often, almost always/always)
53. I don't "Pay attention" (response options: rarely/never, occasionally, often, almost always/always)
54. I am self-controlled (response options: rarely/never, occasionally, often, almost always/always)
55. I concentrate easily (response options: rarely/never, occasionally, often, almost always/always)

56. I am a careful thinker (response options: rarely/never, occasionally, often, almost always/always)
57. I say thinks without thinking (response options: rarely/never, occasionally, often, almost always/always)
58. I act on the spur of the moment (response options: rarely/never, occasionally, often, almost always/always)