Development and Initial Validation of the Caffeine Consumption Questionnaire-Revised

Jessica G. Irons, PhD,1 Drew T. Bassett, BS,2 Caroline O. Prendergast, BS,3 R. Eric Landrum, PhD,4 and Adrienne J. Heinz, PhD5,6

Objective: We sought to outline the rationale and procedures for revising the Caffeine Consumption Questionnaire (CCQ) and to examine the utility and validity of the CCQ-revised (CCQ-R). Given that caffeine is a widely used drug with myriad potential health effects, it is important to measure and understand its use; however, to date, there are no reliable and valid biological tests readily available and existing self-report measures do not have standardized administration or scoring procedures.

Method: The CCQ-R incorporates pictures, visual cues of relative sizes, sample products, and details about products to aid in responding. Additionally, a standardized scoring system is provided.

Results and Conclusion: Validity data are provided and suggest that the CCQ-R offers a valid approach for measuring self-reported caffeine use.

Introduction

Caffeine is the most commonly consumed psychoactive substance worldwide1,2 and occurs naturally in a variety of foods and beverages (e.g., chocolate, coffee, tea). Caffeine is also added to many consumables, including sodas, energy drinks, some chewing gums, and mints, as well as various over-the-counter and prescription drugs. In the United States, ~80% of adults3 and 75% of adolescents4,5 report regular caffeine consumption. Low-to-moderate dosages of caffeine (50–300 mg) can typically be consumed without adverse consequences6 and short-term effects can include reduced fatigue, increased wakefulness, improved performance on simple and complex attention tasks, perceived thought clarity, and analgesic effects for headaches.3,7,8 Negative effects are associated with higher doses of caffeine (500–1000 mg) and can include hypertension, cardiac arrhythmia, anxiety, tachycardia, insomnia, and detrimental effects on working memory at high load levels.3,9

Habitual caffeine consumers may develop tolerance and experience withdrawal symptoms upon cessation following prolonged repeated use. Withdrawal symptoms include drowsiness, headaches, fatigue, and negative mood.3,10 Only recently have caffeine intoxication and withdrawal been recognized in the Diagnostic and Statistical Manual-5 (DSM-5) following review of substantial evidence from preclinical and clinical studies that support the validity, reliability, and clinical significance of caffeine withdrawal.11 Indeed, accumulating data on the impact of caffeine withdrawal and other detrimental outcomes associated with caffeine use (e.g., intoxication) led to the inclusion of caffeine use disorder as a condition for further study in the DSM-5.11 In addition to withdrawal and potential caffeine-related disorders, long-term caffeine use may lead to increased risk of cardiovascular disease12,13 and has been shown to be associated with depression among children and adolescents14 and with increased anxiety and depression in adults.15,16 Additionally, caffeine use has clinical implications related to sleep as well as drug interaction effects that are not yet well characterized (e.g., alcohol)17,18 and are in need of further study.

Despite estimated use prevalence among adults and children and potential negative effects of use, no validated measure of caffeine use exists to date. Furthermore, many...
individuals who consume caffeinated products are unaware of the amount of caffeine they consume or the associated effects of their use patterns.\textsuperscript{19} Reasons for a lack of awareness regarding caffeine consumption levels may include a perception that caffeine is harmless because it is legal or a lack of understanding of its potential effects\textsuperscript{20} and/or a lack of user-friendly, reliable, and valid methods of assessment. In addition, the U.S. Food and Drug Administration (FDA) does not currently require that caffeine content be included on product labels such that caffeine content information is not readily available for many items.\textsuperscript{21} Because consumers are often unaware of their own caffeine consumption habits (in terms of quantifiable servings and caffeine levels) and no standard clinical assessment exists for caffeine use, a sound measurement tool is needed both for clinical and experimental purposes.

Biological testing is the standard and preferred method of assessing substance use for most drugs (e.g., cocaine, marijuana, heroine).\textsuperscript{22,23} Just as with other drugs, there are several biological methods of assessment of caffeine consumption, including breath tests, urinary tests, and salivary tests—each with varying levels of validity. Even among the most accurate of these available tests (urinalysis and salivary tests),\textsuperscript{23,24} biological tests only allow for determination of the amount of caffeine currently present in the body at a given time. Given that the half-life of caffeine is between 4 and 6 hours for healthy adults, the information gained from a biological measurement at a single time point is limited to recent acute use. As such, actual patterns and amounts of caffeine use are virtually impossible to discern over longer time periods.\textsuperscript{17} Given the restricted data available from biological tests and the costs typically associated with such tests, it is important to consider alternative methods for measuring caffeine consumption. Use of other substances is commonly measured through self-reports in both clinical and research settings; however, there are few self-report options for caffeine use. The self-report measures that exist are not ideal from the client/patient/perspective or from the practitioner/researcher perspective given that there are no standardized scoring systems and the measures rely solely on respondents’ knowledge of caffeinated products.

Although self-report measures may be subject to reporting inaccuracies, researchers have shown that self-report measures of substance use are as high as 95% consistent with biological measures.\textsuperscript{25} For example, Timeline Follow-Back (TLFB) is a widely used, calendar-based self-report measure of substance use frequency and quantity. Several researchers have found high percentages of agreement between the TLFB and biological measures for illicit substances, as high as 95.7% for cannabis,\textsuperscript{26} 84.1% for cocaine, and 94.0% for opiates.\textsuperscript{25} Thus, self-report measures can reduce the need for biological tests if shown to be valid measures of behavior. As such, an effective self-report instrument is both possible and needed for measuring caffeine use.

To date, there are only two self-report caffeine measures of which we are aware. The first, the Caffeine Consumption Questionnaire (CCQ),\textsuperscript{27} is a self-report questionnaire designed to circumvent the need for knowledge of the caffeine concentrations in each product. Patients or participants report whether they have ingested specific products, including coffee, tea, soda, cocoa, chocolate milk, and over-the-counter medications. A common serving size is designated for each type of product (e.g., one serving size per item category; 12 oz soft drink, 5 oz coffee). The practitioner or researcher then calculates how much caffeine was ingested. Although the measure allows participants to report caffeine ingestion in terms of the number of servings consumed, the CCQ is not intuitive for consumers to complete. Consumers often do not know the amount of a product they consumed (i.e., how many ounces of soda or coffee consumed), thus rendering their CCQ data a potentially unreliable assessment of caffeine use. In addition, the CCQ is laborious for practitioners and researchers to score and many newer caffeinated items (e.g., energy drinks, caffeinated foods) have been made commercially available since its original conception. Finally, no data are available to provide evidence for the validity of CCQ scores.

A more recent self-report measure, the Stanford Caffeine Questionnaire (SCQ), was developed based on the CCQ for the purposes of studying the effects of chronotype and caffeine on sleep.\textsuperscript{28} As part of the study, the researchers examined the validity of the SCQ using saliva tests for detecting caffeine use. Similar to the CCQ, using the SCQ, participants self-reported their consumption of caffeinated items, including cola, diet cola, pepper soda, citrus soda, other soda, tea (hot), iced tea, instant coffee, brewed coffee, other coffee, store-bought coffee, energy drink, other drink, food, and medications. Although the SCQ measure included common conversions for serving size (e.g., 1 can = 12 oz and Red Bull = 8.3 oz), it suffers from many of the same challenges affecting the CCQ—neither measure accounts for consumers not knowing the amount of caffeinated products they have consumed (i.e., how many ounces of soda or coffee). Furthermore, the SCQ requires participants to report their daily caffeine consumption in 6-hour blocks for 7 days. The use of time blocks may be beneficial for discerning temporal patterns of use; however, time blocks may pose a challenge with respect to reporting accuracy given that caffeine use is often ongoing throughout the day. Furthermore, the use of time blocks necessarily renders SCQ scoring more complex. Participants also provided saliva samples as a validity check for the SCQ; however, saliva tests for caffeine use provide a poor standard for testing validity of self-report measures designed to assess specific use details.\textsuperscript{17} Indeed, Nova et al. found that only 41% of the variance in salivary caffeine concentration level accounted for participant responses on the SCQ.\textsuperscript{28} Thus, it remains unclear if the SCQ is a
valid measure of caffeine use or if the saliva test is simply an insufficient tool for validity assessment.

The current study aims to determine the validity of a modified CCQ-revised (CCQ-R), based on Landrum’s initial measure, designed to be more easily completed by consumers and scored by practitioners and researchers. The validity of caffeine self-report measures may be challenged because (1) typical consumers lack awareness of caffeine content in consumables and (2) there is no standard unit of caffeine dosing (compared with alcohol standard drinks; 5 oz wine = 1.5 oz hard liquor = 12 oz beer) given the variety of methods of ingestion across myriad types of consumables. In an effort to improve self-report of caffeine consumption, the CCQ-R includes the following three advantages over extant measures of caffeine consumption. First, it provides a visual presentation of a variety of commonly consumed caffeinated products across a wide range of commonly available serving sizes. Second, it includes current commercially available items that were not in existence when previous measures were developed (e.g., energy drinks, caffeinated mints, and caffeinated gum). Third, the CCQ-R simplifies how participants report consumption of the caffeinated product by asking about how many times the caffeinated product is consumed, on average, per week. Previous measures ask consumers to report how many times caffeinated products are consumed on average daily, which is problematic because many consumers may not ingest caffeine on a daily basis, use may vary widely across days, and patterns of use may not be captured by examining average daily use. We aim to provide a new self-report measure of caffeine use as well as validity data to suggest its practical utility and potential viability for clinical and research purposes using a validity assessment plan that does not include biological testing.

Materials and Methods

Participants

Participants were 96 college students (57% female; 84% Caucasian, 9% African American, 5% Asian, and 2% Hispanic) from a mid-Atlantic university in the United States.

Measures

Demographic survey. A demographic survey was used to assess gender, class standing, race, and marital status. In addition to the previous items, participants were asked to report whether they had ever been advised to avoid caffeine, whether they were particularly sensitive to the effects of caffeine, and to provide their best estimates of how many milligrams of caffeine they believe they consume in an average day.

Caffeine Consumption Questionnaire-revised. The CCQ-R is a web-based self-report measure of caffeine consumption adapted from the CCQ. The measure instructs participants to report their average weekly consumption of several caffeinated beverages (e.g., coffee, tea, soda, energy drinks), caffeinated foods (e.g., chocolate, candy bars, baked goods), and over-the-counter caffeinated drugs (e.g., weight-loss drugs, headache powder). If an item is not consumed, participants may enter 0 or leave the item blank. Most items pair the caffeinated product with pictures of a variety of commonly available serving sizes of the products. For example, when asked how much coffee participants consume in an average week, they are presented with four pictures of cups of varying sizes that are typically available for purchase (e.g., 8, 12, 16, and 20 oz) and asked to indicate how many of each size of the beverage they consume in an average week. Over-the-counter drugs and caffeinated food products, excluding candy and chocolate bars, are the only products not paired with pictures. The reported number of each serving size is used to determine the total number of units (e.g., oz) of each item type consumed in an average week (to create item-type subscores). This value is then multiplied by the caffeine content for that item type to produce an item-type subscore. Subscores are then summed to reflect the total amount of caffeine consumed in an average week.

Procedures

Participants were enrolled in introductory psychology classes and they received credit for their participation in the study. The university Institutional Review Board approved the study. All participants reported to a computer laboratory to complete the measures (see Fig. 1). After providing written informed consent, participants were asked to read directions and complete the demographic survey and the CCQ-R (self-report of their own personal caffeine use to be used for descriptive purposes: administration 1) using a web-based survey (Qualtrics). Upon completing the demographic survey and the CCQ-R, each participant was presented with a box of products containing both caffeinated items (e.g., Coca-Cola, chocolate bar, Starbucks Coffee) and noncaffeinated foil items (e.g., Sprite, crackers). Each box of products contained 13 or 14 items; up to 12 items in each box contained caffeine and the others were noncaffeinated foils (not specifically labeled as caffeine-free). Participants were instructed to imagine that the items inside the provided box were what they personally consumed in an average week. Participants were then asked to complete the CCQ-R accordingly, as if the content of the provided box was what they consumed (box content; administration 2). The minimum/mean/maximum box totals (total amount of caffeine present in the box based on scoring values) were calculated to compare with the box scores reported by participants. Items that were reported as other were hand scored and manually added to the box totals (administration 2).
Data analyses

A total of 96 participants completed the measures; of those, 15 (females = 9) cases were excluded from analyses. Data exclusion occurred if (1) <75% of the total number of caffeinated items in the box were reported, (2) more items than were present in the box were reported, or (3) multiple items were entered that were not present in any of the boxes. Most cases of exclusion occurred because participants entered serving sizes in lieu of number of servings. Analyses were conducted with remaining participants’ data (N = 81). Those excluded from analyses were not statistically different (p > 0.05) from the remainder of the sample with respect to any demographic variables. All analyses were conducted using the mean value estimate score (see Supplementary Data: CCQ-R scoring; Appendix A for scoring details, Appendix B for SPSS scoring syntax, and Appendix C for scoring codes; Supplementary Data are available online at www.liebertpub.com/jcr). For an example of box contents and administration 2 responses, see Supplementary Data: Appendix D. Percent correct scores were determined by using the following formula:

$$100 - \frac{\text{the absolute value of the difference between reported mg of caffeine and actual calculated mg of caffeine in the box}}{\text{the actual calculated mg of caffeine in each box}}$$

As such, higher values (i.e., percent correct scores) indicated a greater agreement between participant-reported values and actual calculated values of caffeine. The absolute value was employed for tests of inferential statistics to ease interpretation regarding the magnitude of the estimate difference. Roughly 40% of participants achieved a percent correct score above 90%; 59% of participants achieved a percent correct score above 85%. Approximately 43% of the sample under-reported caffeine content of items in boxes.

Results

Average caffeine consumption among the sample (as reported by the CCQ-R) was 1327.15 mg (SD = 1071.82; range = 0–4681 mg) per week or 189.59 mg per day (items reported as other were not included in CCQ-R personal use scoring). Average number of days per week consuming caffeine was 4.43 (SD = 2.09) and five participants reported no caffeine use. A one-way analysis of variance revealed no differences in participant accuracy with respect to the various boxes assigned (p > 0.05)—regardless of the box provided, accuracy was similar. Similarly, there were no differences across races or genders with respect to accuracy (p > 0.05). Furthermore, personal caffeine use (measured by administration 1; see Procedures section) was not correlated with percent correct scores (r = -0.03, p = 0.811). Participants’ estimates of their own consumption (in mg/day) were not correlated with CCQ-R scores of their personal use (administration 1; r = 0.13, p = 0.264) or percent correct scores (r = 0.20, p = 0.071).

An independent t-test revealed a significant difference between participants who reported having been advised to avoid caffeine (n = 19; M = 89.25, SD = 10.08) and those who had not (M = 81.94, SD = 13.26; t(79) = -2.11, p = 0.04) with respect to CCQ-R accuracy. The average percent correct score was 83.47% (SD = 12.95). The range of percent correct scores was 57.03 (42.96–99.99%). A dependent t-test revealed no significant difference between participants’ CCQ-R scores for the boxed items and the actual amount of caffeine in the boxes, t(79) = 0.96, p = 0.340.

Discussion

We sought to examine the practical utility and potential validity of the CCQ-R for estimating caffeine use. The CCQ-R is a more flexible and comprehensive self-report measure of caffeine use than its predecessors that may be administered and scored electronically. Although participants were unable to accurately estimate
their own caffeine use (compared with their personal CCQ-R scores; administration 1), participants’ CCQ-R scores for the items included in the provided boxes were, on average, ~84% correct. We suggest that the CCQ-R may provide a valid estimation of caffeine consumption.

CCQ-R accuracy was not impacted by any demographics except whether participants had been advised to avoid caffeine. We believe these data provide evidence for construct validity, in that we would expect no influence of most demographics on the ability to complete the measure accurately; however, we might expect that those instructed to avoid caffeine would have better knowledge of what products may contain caffeine (and thus make fewer reporting errors) than those who have not been instructed to avoid caffeine. Percent correct scores yielded nearly 84% average accuracy on the CCQ-R and ~60% of the sample scored 85% correct or higher; 85% agreement is an accepted standard of inter-rater agreement. The CCQ-R demonstrated criterion validity, in that a majority of percent correct scores reached and exceeded 85% correct inter-rater agreement (i.e., researcher-calculated scores compared with participant-reported scores).

In addition to evidence for validity, the CCQ-R has a number of strengths. The measure is easily administered using a web-based survey or it can be printed for paper/pencil completion. The CCQ-R is readily scored using Statistical Package for the Social Sciences (SPSS) scoring syntax that includes caffeine concentration values from the US Food and Drug Administration (FDA), Mayo Clinic, manufacturers’ company websites, and product packaging; these values can be readily updated in the scoring syntax should new concentration data become available for any items. Additionally, three different scores may be garnered from our scoring syntax: a conservative estimate using minimum caffeine concentration values available in calculations, a mean estimate using average caffeine concentration values, and a liberal estimate using maximum caffeine concentration values. Furthermore, researchers and clinicians may choose to be more or less inclusive with respect to item categories for which data are gathered and/or scored. For example, if only coffee consumption is of interest, respondents may be easily avoided by complete knowledge of which items contain caffeine versus those that do not. In the current study, participants who had been advised to avoid caffeine yielded significantly higher percent correct scores than those who had not, presumably because avoiding caffeine requires a relatively complete knowledge of which items contain caffeine. Providing a comprehensive list of specific items containing caffeine may be helpful for improving accuracy of reporting. Finally, in some cases, it may be appropriate to ask respondents to keep a daily caffeine consumption log to aid in completion of the CCQ-R.

Conclusions

In summary, the CCQ-R is the first self-report measure for caffeine use for which there are data to provide evidence for its validity. The CCQ-R is also user-friendly for both researchers and clinicians, as well as for their respective participants and patients. Specifically, it can be administered using paper and pencil (see Supplementary Data: Appendix A) or as a web-based questionnaire (www.caffeineconsumptionmeasure.com). Scoring is automated and total consumption along with consumption in item-specific categories can be automatically calculated with syntax in SPSS (see Supplementary Data: Appendix B). Use of the CCQ-R outside of a research and clinical context (i.e., self-monitoring purposes) is an additional possible indication in need of further study.

Author Disclosure Statement

No competing financial interests exist.
References


Address correspondence to:
Jessica G. Irons, PhD
Department of Psychology
James Madison University
Harrisonburg, VA 22807

E-mail: ironsjg@gmail.com