Published in final edited form as:
Addiction. 2000 August ; 95(Suppl 2): S189–S210.

The measurement of drug craving

Michael A. Sayette[1], **Saul Shiffman**[1], **Stephen T. Tiffany**[2], **Raymond S. Niaura**[3], **Christopher S. Martin**[4], and **William G. Shadel**[5]

1*University of Pittsburgh, Pittsburgh, USA*

2*Purdue University, Pittsburgh, USA*

3*The Miriam Hospital and Brown University School of Medicine, Pittsburgh, USA*

4*University of Pittsburgh School of Medicine, Pittsburgh, USA*

5*The Miriam Hospital and Brown University School of Medicine, Pittsburgh, USA*

Abstract

This paper reviews theoretical and methodological issues in the measurement of drug craving, with an emphasis on self-report assessment. Despite the important role that craving plays in many research and clinical settings, the way in which the construct is conceptualized and measured rarely receives sufficient attention. Issues pertinent to conceptualizing craving are identified. Because there is no single perfect measure of craving, it is essential that researchers understand the limitations of each measure. Measurement performance concerns that affect the validity of different measures are reviewed. Non-verbal assessment methods are also reviewed. Research factors that may help determine the optimal measures for a given study are highlighted. It is concluded that advances in assessment will proceed only when combined with additional research and a better theoretical understanding of craving.

Introduction

Researchers have long posited a relationship between craving and addiction (Wikler, 1948; World Health Organization, 1955). Yet the importance of this relationship, and the very nature and validity of craving as a construct, have been vigorously debated (Mello, 1978; Tiffany, 1990; Kassel & Shiffman, 1992). Nevertheless, there has been a recent increase in the study of craving from a variety of biological and psychological perspectives (e.g. Pickens & Johanson, 1992; Lowman *et al.*, 2000). Unfortunately, the manner in which craving is measured rarely receives sufficient attention. Further advancement in craving research is hampered by divergent conceptualizations and inconsistent measures (Kozlowski & Wilkinson, 1987).

Because there is no single accepted measure of craving, the challenge is to select the optimal measure(s) for a particular research or clinical application. The primary goal of this paper is to offer some guidelines for selecting appropriate measures of drug craving. We begin with a brief summary of the principal definitions of craving and related conceptual issues. Next, we review and evaluate self-report approaches to the assessment of craving. Included in this section is a review of basic measurement performance issues. Non-verbal measures are then discussed. Finally, factors that influence choice of craving measure are addressed. Unless otherwise noted, our recommendations are meant to apply to a range of addictive substances. We will not be covering the numerous models of craving that have appeared in the literature, as they are

Correspondence to: Michael Sayette PhD, Department of Psychology, 604 OEH, University of Pittsburgh, Pittsburgh, PA 15260 USA. E-mail: E-mail: Sayette+@Pitt.edu.

beyond the scope of this paper and have been addressed in detail elsewhere (e.g. Baker, Morse & Sherman, 1987; Niaura *et al.*, 1988; Wise, 1988; Tiffany, 1990; Kassel & Shiffman, 1992; Robinson & Berridge, 1993).

Craving: definitional issues

Although drug craving has been defined in numerous ways, it has generally been regarded as a desire to use a drug. Craving is usually considered to be a subjective experience, in the sense that one must be aware of a desire in order to crave (Niaura *et al.*, 1988; Kassel & Shiffman, 1992). With few exceptions (e.g. Tiffany, 1990), craving has been conceptualized as reflecting a drug-acquisitive state which motivates drug use. Despite the popular appeal of craving as a construct, there are a number of definitional issues that remain unresolved.

First, researchers have differed regarding the scope of their craving definition. While some restrict the definition to a desire for drug use, other definitions have different foci. For example, Buydens-Branchey *et al.* (1997) target the behavioral intention to use a drug. Specifically, these authors operationalize alcohol craving as "the chance they [patients] would use alcohol if they were outside the hospital and had unlimited financial resources" (p. 221). Marlatt (1985) has suggested restricting the term "craving" to refer to desire for the effects of a drug while using the term "urge" to indicate a behavioral intention to use. Still other investigators suggest the term "craving" to encompass a broad range of phenomena including anticipation of a drug's reinforcing effects, intention to engage in drug use and desire for the drug (Tiffany & Drobes, 1991). Views of craving that include concepts other than desire have been criticized for including processes that are correlated with, but conceptually distinct from, craving (Kozlowski *et al.*, 1996).

Secondly, debate continues regarding the time frame of a craving experience. Some investigators have assessed craving as if it were relatively stable. Accordingly, individuals are asked to rate their craving for a particular day or week (West, Hajek & Belcher, 1989; Anton, Moak & Latham, 1995). Alternatively, craving can be viewed as a momentary "pulsitile" state (Gawin, 1991; Shiffman *et al.*, 1996). From this perspective, measures that encompass long time intervals may lack the precision to capture changes in craving. Measures that can be completed during momentary craving experiences are typically administered in controlled laboratory settings, although recent advances in assessment methodology have yielded near-real time assessment in the natural environment (Shiffman *et al.*, 1996). These real-time assessments support the view that craving vacillates substantially even within the course of a day, and that craving reports obtained at different times have different meanings and predictive power (Shiffman *et al.*, 1996, 1997a).

Thirdly, researchers have questioned whether craving should be characterized as a dimensional construct that can be assessed along a broad continuum of desire, or instead be restricted to an experience characterized by extreme desire (e.g. West & Schneider, 1987). Kozlowski & Wilkinson (1987) prefer the latter definition of craving and reserve the term "urge" to indicate a less extreme state of desire. This position begs the unresolved question of determining at what point an urge becomes craving (Shiffman, 1987; West, 1987). Nor is it clear from conventional dictionary definitions that urge is a broader, more dimensional term than craving. In terms of assessment, self-report responses from the same individuals to the items "urge" and "craving" have been nearly identical (e.g. Shiffman, 1987; Tiffany & Drobes, 1991; Tiffany *et al.*, 1993).

Fourthly, some investigators have suggested that craving might exist even in the absence of awareness (Miller & Gold, 1994; Berridge & Robinson, 1995). According to this perspective, craving is a motivational construct that can operate independently of conscious awareness. Berridge & Robinson (1995) suggest that it is possible to crave a drug without awareness. If

one can crave without awareness (i.e. if there can exist dissociations between the underlying process of craving and awareness of that craving), then non-verbal measures such as drug-seeking and drug-use behavior ought to feature at least as prominently as self-report measures in craving assessment. When craving is only measured as drug use behavior, however, one can question the utility of invoking any craving construct, conscious or otherwise (Mello, 1978).

A final definitional issue to be addressed is the degree to which craving and drug use should be associated. For example, some theorists have suggested particular situations in which measures drawn from different response systems (e.g. psychophysiology, self-report, behavior) should converge (Baker *et al.*, 1987). Of particular interest has been the putative association between craving and relapse. Many investigators have linked craving to relapse, claiming that craving is necessary for relapse (e.g. Ludwig & Wikler, 1974). Such a view has been criticized on both conceptual and empirical grounds (Mello, 1978; Tiffany, 1990), and several recent accounts of addiction do not consider craving to be an inevitable concomitant of relapse (e.g. Baker *et al.*, 1987; Tiffany, 1990; Kassel & Shiffman, 1992). Even if craving need not precede relapse, presumably there should at least be a relationship between reported craving and subsequent use among treated addicts. Such an association was found in three recent smoking studies reported by Killen & Fortmann (1997) and by Shiffman *et al.* (1997a). Nevertheless, previous reviews have concluded that the relationship among different measures of craving, and between measures of craving and relapse, while they exist, are not strong (Rohsenow *et al.*, 1990–91; Tiffany, 1990). The lack of robust associations between craving and relapse may be attributed to (a) the presence of competing motives in addition to the motivation for drug use that can influence the decision to consume a substance, (b) the possibility that drug use might operate independently of the processes subserving craving (Tiffany, 1990) and, of most relevance to this paper, (c) limitations in how craving and drug use are measured.

In summary, there remains much that is still debated about the definition of craving. Difficulty in selecting appropriate measures has been due, in part, to a failure to articulate what is meant by craving. The literature is replete with attempts to settle this issue once and for all by providing a universally held definition for craving (e.g. Kozlowski & Wilkinson, 1987; Marlatt, 1987), and we will not add to this list. Regardless of the definition chosen, however, it is essential for researchers to make an explicit definition of this construct, to select measures appropriate for this definition, and to draw conclusions that are appropriate to the measure. The remainder of this paper addresses a range of methodological issues associated with the assessment of drug craving.

Self-report measures

Nearly all conceptualizations of craving assume that drug-motivational processes can be indexed, at least in part, through self-report measures of subjective experience. Not surprisingly, such measures are popular because they display a high degree of face validity, can be constructed with relatively little difficulty, and are fairly easy to collect. It is problematic, however, to view self-reports of craving as providing a direct readout of an individual's craving state. This interpretation of the meaning of a self-report measure, which assumes a one-to-one mapping of verbal reports to hypothetical internal states, has been described as the correspondence view of test meaning (Buchwald, 1961; Wiggens, 1973). Although intuitively appealing, the correspondence view rests on several implausible assumptions.

As noted by Wiggens (1973), in order for "accurate" measurement to occur from the perspective of the correspondence view, the following conditions must hold: (1) Craving item (s) have common meaning among participants and between the participant and the researcher. (2) The participant must be able to accurately assess his/her own internal states. Distortions due to defensiveness or insensitive observations must be minimized. (3) The participant must

report those internal states honestly to the tester. (4) The craving item or items are, in fact, related to the concept of "craving" as used by the researcher. In light of these requirements, Wiggens (1973, p. 384) concluded, "Clearly the conditions which must be fulfilled to meet the correspondence assumption are unrealistic even under the most optimistic view of the honesty of the subjects and the semantics of verbal behavior."

The correspondence view can be contrasted with the position that, along with the operation of hypothesized craving processes, responses on a craving scale will also reflect the confluence of a variety of other psychological influences. From this perspective, the meaning of a measure of craving report must be construed in terms of particular theories of craving as well as an appreciation of the myriad forces that may shape the verbal behavior of participants. Craving researchers need to be sensitive to memory biases, misattributions, response styles and other types of craving-irrelevant influences that can affect craving ratings (Kozlowski & Wilkinson, 1987; Drummond, Cooper & Glautier, 1990; Tiffany & Drobes, 1991; Bohn, Krahn & Staehler, 1995). For instance, individual differences on variables such as social desirability or self-consciousness may have an impact on a range of self-report measures associated with addiction, including craving (see Baker & Brandon, 1990). Self-reports involving retrospective recall or global summaries of prior experiences (e.g. how much you *generally* experience craving) can be particularly subject to bias. The process of accessing autobiographical memory has inherent potential for inaccuracy and bias, which can occur without respondent awareness or intention to distort (Hammersley, 1994; Shiffman *et al*., 1997b). Of course, these issues are no different for craving ratings than for any other self-report measure.

A number of procedures have been adopted to combat potential limitations of self-report assessments, including: (a) approaches to conceal the purpose of a study or implications of responses; (b) techniques to either facilitate reporting of undesirable responses or minimize reporting of falsely desirable answers; and (c) methods to protect against response sets to agree or disagree with items, or to provide extreme responses (Cook & Selltiz, 1964). Moreover, it is clear that self-report can often provide important information about a person's craving, and should continue to play a major role in the measurement of craving (Tiffany, Carter & Singleton, 2000). There are numerous types of self-report craving measures that differ along several parameters addressed below.

Measurement performance issues

Validity—The most critical concern when considering measures of craving is that of validity. In a general sense, validity refers to the appropriateness of inferences made about scientific data (Cook & Campbell, 1979; Geisinger, 1992). The validity of a measure has been defined as the "appropriateness, meaningfulness, and usefulness of inferences" made from test scores (American Educational, 1985). Almost all empirical studies using scores on a measure provide some evidence regarding the validity of that measure (Anastasi, 1988). Importantly, a measure that was validated on one sample (e.g. severely dependent smokers) may not be valid for another group (e.g. light smokers). It therefore is the responsibility of each investigator to check the appropriateness of its use (Geisinger, 1992).

There are several different types of validity which apply to different aspects of measurement performance. *Face validity* refers to the apparent plausibility or directness of a measure. Many self-report measures of craving have face validity; they seem to assess craving directly. Appearing to assess the real world concept, however, may not reflect a measure's ability to capture differences in the underlying conceptual variable. *Content-related validity* refers to the extent to which a measure adequately represents the domain it is supposed to be measuring. Content validity requires an evaluation of (a) the appropriateness of the type of items for assessing the domain of interest and (b), the completeness of the item sample. *Criterion-related validity* refers to correspondence between scores on a dependent measure and some

independent, objective measure of the variable under study that is of known validity. Criterion-related validity is less relevant to craving research because there is no single "gold-standard" criterion for craving. Because the relationship between craving and drug use is itself still open, the lack of a correlation between a craving measure and drug use does not necessarily invalidate the measure. Of course, if relapse is the criterion of interest, rather than craving, then this form of validity is relevant.

Unlike criterion-related validity, *convergent validity* and *discriminant validity* are highly relevant to the assessment of craving. Convergent validity refers to associations of a measure with other measures that also reflect the construct or closely related constructs. Discriminant validity is indicated when a craving measure is not highly associated with measures that reflect constructs that are distinct from craving. Convergent and discriminant validity are best evaluated together, using multiple measures thought to reflect craving and non-craving constructs. One approach to these two validation processes involves a matrix of intercorrelations among tests representing multiple traits, each measured by multiple methods (Campbell & Fiske, 1959). When convergent and discriminant validity are present, measures of the same trait using different methods are more associated with each other than are measures of different traits using the same method. While this approach is widely accepted, it is rarely implemented (Fiske & Campbell, 1992).

Recently, consensus has been building among psychometricians that all forms of validity can be subsumed under the concept of construct validity (Geisinger, 1992). *Construct validity* refers to the adequacy and explanatory power of scientific concepts, such as craving. Construct validation involves associating multiple measures with other theoretically relevant behaviors under conditions that are believed to affect the construct (e.g. during drug abstinence or drug cue exposure). This process typically requires the amassing of data from a variety of sources and experimental contexts (Cronbach & Meehl, 1955; Anastasi, 1988). Construct validity is not derived from a single study; rather it is inferred from the accumulation of data through ongoing research.

The evaluation of construct validity is meaningless in the absence of a developed theory. That is, construct articulation must precede construct validation. Simply correlating a self-report measure with other presumptive indicants of the same construct is not sufficient for construct validation. One must state, a priori, specific theory-driven predictions. For example, a withdrawal-based theory of craving assumes a strong association between craving report and withdrawal symptoms. In contrast, an incentive–motivational model would not make this prediction, instead linking craving reports to a drug's appetitive effects. In general, there is no single craving construct; there are as many craving constructs as there are craving theories. Evidence for construct validity is more likely to develop as the adequacy of craving theories increase.

The validity of craving measures is affected by a number of criteria. These include reliability, reactivity (obtrusiveness), specificity, sensitivity and floor and ceiling effects.

Reliability—The reliability of a self-report measure typically refers to its ability to provide scores that are consistent and stable. Several different kinds of reliability can be measured, reflecting different sources of variance in observed scores. One type of reliability important for multi-item self-report measures of craving is *internal consistency*, which measures variation in scores among different items on a scale. Relatively high levels of internal consistency reflect strong associations of responses across a set of items, and are an important attribute of multi-item self report craving measures. Low internal consistency indicates that variables other than the construct of interest produce substantial variation in responses across scale items. High internal consistency usually indicates that a scale is uni-dimensional, because it implies a high

average inter-item correlation, whereas multidimensionality implies two or more clusters of items at best weakly correlated with each other. Since the internal consistency of a scale is a function of the average inter-item correlation, for any given number of items, uni-dimensional scales will have higher internal consistency. Multidimensional scales can have high internal consistency only if they are very long (so that the large number of items makes up for low inter-correlation) or if they have a partially uni-dimensional structure, as when subscales are part of an underlying common or higher-order factor (so that there are substantial intercorrelations among items across scales).

Test–retest reliability measures the consistency of obtained scores over repeated testings in the same individuals. High test–retest reliability for a measure can be expected only when the construct of interest is highly stable over time. Craving, however, is not considered to be a stable trait-like construct in most models. Craving ratings are known to fluctuate over time, and are sensitive to a wide range of environmental influences (Gawin, 1991; Shiffman *et al.*, 1996) and experimental manipulations (Rohsenow *et al.*, 1990–91). Although test–retest correlations have been used to support the reliability of a craving measure (e.g. Voris, Elder & Sebastian, 1991), this is better conceived as an index of *stability* (Carlsmith, Ellesworth & Aronson, 1976; Tiffany, 1992). Lower stability can reflect changes in craving over time and does not necessarily indicate poor performance of a measure. Indeed, the assessment of changes over time in craving levels has been the focus of observational studies (Shiffman *et al.*, 1997a).

A specific type of test-retest reliability especially relevant to laboratory research is *repeatability*. Repeatability refers to the consistency of scores across repeated testings under similar assessment conditions (e.g. in smokers after 12 hours of nicotine abstinence). While repeatability is a desirable test attribute, when assessing craving establishing similar assessment conditions across repeated testings is difficult. It is likely that levels of craving change over time even when assessed under similar experimental conditions, due to a wide variety of other influences.

Perhaps the most important advantage of high reliability is the relationship between reliability and validity. The reliability coefficient expresses the proportion of variance in observed scores attributed to the “true” measure of craving, as opposed to other sources of variation. Therefore, reliability sets an upper bound on the validity of a measure. So, for example, in a correlational study, the maximum correlation between a craving measure and any other variable cannot exceed the reliability of the craving assessment. [More technically, a correlational index of validity cannot exceed the square root of the reliability coefficient (Maloney & Ward, 1976)]. Consequently, investigations of the correlates of craving report must be interpreted in light of the reliability of the measures.

The critical relationship between reliability and validity is not limited to correlational studies. Data derived from experimental manipulations of craving are equally sensitive to the reliability of the dependent measures. That is, the power of an experiment to reveal the impact of any pharmacological or behavioral manipulation on craving will depend, in part, on the extent to which the measure of craving is reliable. Because highly reliable measures of craving enhance the ability to detect effects in experimental research, other things being equal, they also will increase statistical power. Thus, there can be economic as well as scientific benefits to the systematic development and utilization of reliable craving measures.

When investigators choose a reliable craving scale, it is important to recognize that reliability may change under different testing conditions. Internal consistency will probably decline, for instance, when a sample presents a restricted range of scores on the scale. Accordingly, it is

ideal to check the reliability of a multi-item scale for each experimental or clinical context in which it is being used (Tiffany *et al.*, 2000).

Reactivity—Measuring craving using a self-report assessment necessarily requires participants to reflect on their experience which, under certain conditions, may influence craving itself. It is well established that self-monitoring can change the behavior being monitored (Perlmuter, Noblin & Hakami, 1983). Self-report assessment is particularly reactive, compared to other response domains, in that it brings craving into awareness by eliciting controlled processing that demands attentional resources. Because many aspects of a drug use routine may be automatized (Tiffany, 1990), such a shift in processing resulting from assessment may create or modify a conscious craving experience (Baker & Brandon, 1990). Increasing the number of times that a measure is administered may exacerbate this problem. Generally, studies using pre- and post-manipulation designs include non-craving "control" manipulations to account for a main effect of measurement. This approach does not control for *interactions* between the repeated use of a measure and a craving manipulation. That is, the combination of a pre-manipulation baseline assessment and a subsequent craving manipulation may alter the post-craving experience (Carlsmith *et al.*, 1976).

Specificity—Specificity refers to the ability of a measure to capture the construct of interest, without being affected by other variables. Simply put, someone who is not craving should not score highly on a measure of craving that has adequate specificity. A drug craving measure with adequate specificity should be relatively insensitive to other factors, such as hunger or fatigue. In part due to the myriad conceptualizations of craving that exist, there has been little research aimed at determining the specificity of craving measures.

Sensitivity—Sensitivity refers to the ability of a measure to detect variation in the construct of interest. Different measures, or different items on a scale, may have different thresholds for detecting craving (see Cook & Selltiz, 1964). Thus the item, "I would do almost anything for a cigarette now" may have less sensitivity than "It would be nice to smoke right now", as presumably a stronger degree of craving would be required before change would result on the former item, relative to the latter one.

Ceiling and floor effects—Ceiling and floor effects are especially relevant to craving assessment. A ceiling effect occurs when individuals tend to report similarly high craving levels at or near the endpoint of the measurement scale, despite actual differences in experienced craving. Most published studies of craving have not found ceiling effects (Carter & Tiffany, 1999). Nevertheless, it appears that under certain conditions (e.g. heavy smokers who are nicotine-deprived) ceiling effects are possible regardless of the end-points of the rating scale, number of items or labels (Zinser *et al.*, 1992; Perkins *et al.*, 1994; Sayette & Hufford, 1994; Hutchison, Niaura & Swift, 1997). Many studies include a baseline assessment of craving followed by a drug cue exposure manipulation designed to increase craving. In some cases, participants report their pre-exposure craving state as being extremely high, and thus do not leave themselves sufficient room on a scale to accurately characterize any potential increase. Of course, uniformly high craving scores may be accurately reflecting "true" high levels of craving.

A floor effect occurs when a scale is insensitive to craving, leading to scores that are uniformly low. Low levels of craving report (e.g. Powell, 1995) may be an accurate depiction of a minimal craving experience (perhaps due to an ineffective manipulation) or may be due to floor effects resulting from a scale with a high sensitivity threshold.

Assessment parameters

Several aspects of the format of self-report craving measures have varied across studies. In most cases, selection of one format over another has been idiosyncratic and little or no data are available to evaluate them. This section will review some of these different measurement formats with the aim of identifying formal comparisons in future research.

Modality—The most common approaches to self-report craving assessment have been to ask participants to write down their ratings, or to manipulate a dial or joystick linked to a measurement device. Recently cravings have been assessed outside the laboratory by asking people to input their ratings directly into hand-held computers (e.g. Shiffman *et al.*, 1996). Ratings also have been provided vocally by respondents (e.g. Droungas *et al.*, 1995). One advantage of vocal ratings is that participants whose hands are otherwise preoccupied during experimental tasks [e.g. holding a cigarette, engaging in button pressing tasks (Sayette & Hufford, 1994)] can still respond. A disadvantage may be that participants feel uncomfortable proclaiming their private experiences accurately, especially in cases where they may feel ashamed or frustrated by their cravings. [A study comparing traditional questionnaires and computer-assisted self-interviewing procedures indicates that even paper and pencil questionnaires may inhibit respondents from revealing sensitive self-relevant information (Turner *et al.*, 1998)]. In addition, participants may have difficulty recalling the correct response format using a vocal administration. Vocal measures may be especially difficult to administer if there are multiple items included in the measure.

Semantic dimensions—An important issue to consider when selecting a self-report craving measure is whether the items on a scale adequately represent the semantic dimensions that people use to characterize their craving experience. For example, when substance abusers were asked to choose the best description for the term “craving” there was substantial disagreement, with some equating craving with any level of desire, and others viewing it only as a strong desire (Kozlowski *et al.*, 1989). More generally, drug users differ in the response domain that is most affected, with some emphasizing physiological sensations, some emphasizing behavior and some the subjective emotional experience (Shadel *et al.*, 2000). Given the diversity of concepts associated with craving among researchers, there are probably differences in how craving is regarded and described verbally by drug users.

Scale characteristics—The most common types of self-report craving measures are rating scales. Originally rating scales, such as the Likert scale, were developed to assess attitudes, using a series of numerical ratings ranging from “strongly disapprove” to “strongly approve”. Rating scales of craving have required participants to judge their responses to items such as “craving”, “urge” and “desire”. Typically, the end-points of the scale are labeled “not at all” and “extremely” or words to that effect (e.g. Kaplan *et al.*, 1985; Childress *et al.*, 1986). In another approach, end-points are labeled “strongly disagree” and “strongly agree” (e.g. Tiffany & Drobes, 1991) or “yes” and “no” (e.g. Shiffman *et al.*, 1995). Conceptually there is a distinction between these different types of end-points. For instance, someone who is absolutely certain that s/he has a mild urge would report a different score on an “urge” item depending on the type of end-point label. We are unaware, however, of any empirical data examining whether the different types of end-points lead to different results. Nor are we aware of data contrasting rating scales that use *absolute* (e.g. “extremely”) versus *relative* (e.g. “the strongest urge I have ever experienced”) end-point labels. Use of either type of scale may lead to individual differences in reporting of craving as people vary in their interpretation of the labels (McCusker & Brown, 1991). This concern has led some investigators to provide descriptive anchors throughout their scale (e.g. Ludwig, Wikler & Stark, 1974).

A measurement question that rarely is addressed is whether a scale should be unipolar, with the starting point being neutral (i.e. no craving for the substance), or instead be bipolar, with the starting point reflecting an aversion to the substance, and neutral appearing at the midpoint of the scale (Kozlowski & Wilkinson, 1987). In the case of the agree/disagree format, there is ambiguity about whether people "strongly disagreeing" with the statement that they have an urge, are experiencing an aversion to the substance, or just no urge. Because craving may reflect elements of potentially conflicting experiences [e.g. one may simultaneously feel both aversion and attraction while craving a drug (Avants *et al.*, 1995)], and because aversion and desire may not exist on the same dimension, we suggest using unipolar scales.

Ideally, responses to craving items would not be influenced by the manner in which items were presented. That is, a craving questionnaire should be constructed so as to maximize content variance and minimize sources of method variance. One potentially intrusive source of method variance may be the tendency of some participants to endorse items regardless of the items' content. There are a number of ways to reduce this acquiescence bias (Jackson, 1967); perhaps the most direct solution is to balance a craving scale by keying a substantial proportion of the items negatively and the remaining proportion of the items positively (e.g. Tiffany & Drobes, 1991; Tiffany *et al.*, 1993).

Negatively keyed items can provide some degree of internal control over acquiescence bias. When constructing such items, however, it is important to consider that negative keying can be achieved through both negative and positive wording. For example, "I do not want cocaine now" uses negative wording to create a negatively keyed item. In contrast, a positively worded yet negatively keyed item might state, "I want to avoid cocaine now". There are no direct data from the craving literature regarding the impact of negative or positive wording on the psychometric properties of negatively keyed items. Nevertheless, participants sometimes report difficulty interpreting certain negatively worded items. Moreover, there is evidence from personality questionnaires that negatively worded items tend to have lower validity than positively worded items (Schreisman & Hill, 1981; Holden, Fekken & Jackson, 1985). Thus, it may be preferable to word items positively when constructing craving scales with any form of negative keying.

In addition to labels, the range of values on scales of craving has varied, with maximum values as low as 4 (e.g. Shiffman *et al.*, 1995) and as high as 256 (e.g. Zinser *et al.*, 1992). It is unclear whether the size of the range influences assessment, as different scales have rarely been compared in the same experiments. Recent data collected in the first author's laboratory using nicotine-deprived heavy smokers during a cigarette exposure manipulation (see Sayette & Hufford, 1994 for methodological details) found maximum values were reported more frequently on 10-point than 100-point scales, suggesting that 100-point scales may be less vulnerable than 10-point scales to ceiling effects.

Magnitude estimation—Ceiling effects can sometimes prove troubling in craving studies when pre-manipulation baselines approach the maximum values of the scale, leaving little or no room for any additional increases to be reported. One statistical attempt to address this problem is to transform the data (e.g. a Tobit transformation). Nevertheless, such approaches cannot address problems occurring when many people choose the maximum value on the scale. Another approach to reducing problems associated with ceiling effects is to use a measure that does not include a maximum value.

One way to assess craving following a manipulation is to ask participants to rate their current craving relative to their baseline craving. This approach also can be used by having participants compare their craving during one manipulation with the craving produced by a prior manipulation. That is, they might be asked if their craving is higher or lower (a dichotomous

measure), or the extent to which their craving is higher or lower following Condition X compared to the prior condition. This latter approach to dealing with ceiling effects involves magnitude estimation.

Magnitude estimation has been used to scale the intensity of psychological responses, including ratings of sensory stimuli (e.g. Stevens, 1955), emotional states (e.g. Grossberg & Grant, 1978; Acri & Grunberg, 1992) and perceived level of intoxication following alcohol consumption (e.g. Ekman *et al.*, 1964; Sayette & Wilson, 1991). The critical feature of magnitude estimation is that the scale has no maximum endpoint. Participants first imagine a standard stimulus intensity (e.g. their craving level after not smoking for 6 hours or at the start of an experiment). This value is assigned a number, e.g. 10, which then serves as a comparison for all subsequent estimates. Following this initial assessment, participants rate their perceived magnitude of craving in relation to the initial value. If craving has doubled, due perhaps to exposure to a cigarette, then the assigned number would be 20. If craving is now only half what it had been originally, the value would be 5.

The appeal of a magnitude estimation stems from situations in which individuals might begin a study with a strong craving, such as in studies that require a period of abstinence prior to entering the laboratory. In some of these cases, ceiling effects have occurred at the time of baseline craving assessment (e.g. Zinser *et al.*, 1992; Perkins *et al.*, 1994; Sayette & Hufford, 1994; Hutchison *et al.*, 1997).

Although there are apparent advantages to using magnitude estimation in studies where ceiling effects are anticipated, there are also potential limitations. Depending on the variability of the sample the standard intensity, as interpreted by each person, may reflect very different experiences. For example, a heavy smoker deprived of nicotine for 12 hours may be comparing his/her post-manipulation craving levels to a very different baseline level than an occasional smoker who feels little effects of nicotine deprivation. This is in marked distinction to traditional uses of magnitude estimation in some areas of psychophysics in which an individual's report can be mapped onto a known dimensional manipulation (e.g. weight, wavelength, pitch). Of course, there are no objective dimensional anchors for any self-report craving scales, so this limitation of magnitude estimation is a concern for all types of self-report measures of craving. Moreover, use of a standard number when using a magnitude estimation can be confusing, and requires a basic level of mathematical skills. This may be especially confusing to some participants when they are using a scale that involves more than one item. It is essential when using magnitude estimation to include training prior to craving assessment. Finally, magnitude estimation scores are unlikely to be normally distributed and care should be taken to select appropriate forms of analyses (e.g. data transformations, non-parametric analyses).

In summary, empirical data are needed to test the utility of magnitude estimation measures to assess craving. At this point we would not advise relying solely on these scales. Nevertheless, in situations in which many people may experience ceiling effects including a magnitude estimation approach may prove useful.

Number of test items—One of the most pressing questions regarding the validity of self-report measures of craving is the number of items contained in the assessment. There are several criteria to consider when selecting the number of items in a self-report measure of craving.

Single-item ratings may lack the breadth required to capture the various semantic dimensions used by people to describe their craving (Tiffany & Drobes, 1991). Some investigators have responded to this concern by providing participants with an operational definition of craving (e.g. Powell *et al.*, 1990), although it has been suggested that training in the use of a scale can

bias responses (Pickens & Johanson, 1992). Alternatively, researchers have created multi-item scales that aim to capture a broad range of conceptualizations of craving. Although the items on multi-item scales of craving may be semantically distinct (e.g. “craving a cigarette” and “difficulty not smoking”), high inter-item correlations have led some to contend that they are capturing a similar underlying construct while also improving reliability over scales containing a less diverse set of items (e.g. West, 1987). Such high levels of association across items on a craving scale may be due in part to a response bias to present consistent responses, or to the tendency for a participant’s response to the initial item to “anchor” responses to subsequent items.

It is possible that increased numbers of test items would improve sensitivity. For example, during opiate withdrawal, Schuster *et al.* (1995) found a multi-item craving questionnaire to detect changes while a single-item “I feel like shooting up now” rating did not. Inclusion of items such as “All I want is a cigarette” and “I have an urge to smoke” may increase sensitivity throughout all phases of a study, by reducing the chances of ceiling or floor effects compared to a single item measure. When many items are used to create the scale, however, numerous combinations of responses can lead to the same score (Carlsmith *et al.*, 1976). Consideration of this issue is affected by one’s articulation of the craving construct. If we take the position that we do not know which types of items are the purest indicators of craving, then the addition of a broader range of items is advantageous (West, 1987; Tiffany & Drobes, 1991). If instead we posit that certain items are more accurate reflections of craving than others, than the inclusion of a broad range of items may be inappropriate (Kozlowski *et al.*, 1996).

Clearly reliability, and thus power, can be increased with the use of multi-item, relative to single-item, scales. Yet it also is likely that increasing the number of items on the self-report measure would increase the reactivity of the measure. The reactivity caused by a longer measure is a function of the number of times that a participant is required to complete it, the amount of time elapsing between assessments and the nature of the craving manipulation (if any) that is being used. During certain craving induction procedures, immediate “real time” responses to craving manipulations are desired (e.g. Sayette & Hufford, 1995), and especially brief ratings may reduce the chance that the time required to complete the measure will interfere with (either by increasing or decreasing) craving levels. Questionnaires with as many as 30 items are impractical for studies requiring multiple craving assessments.

Very little is known about the reactivity of self-report craving scales. Research on other psychological states suggests that using relatively long self-report craving measures may enhance, or even create, an underlying craving experience. Mark, Sinclair & Wellens (1991) found that simply completing the Beck Depression Inventory affected subsequent scores on a mood measure. Clark *et al.* (1988) induced panic attacks in 83% of panic patients, merely by asking them to read a set of word-pairs (e.g. palpitation–dying) related to anxiety. Accordingly, it is not unreasonable to question whether the administration of multi-item craving measures may affect craving, especially when participants are drug deprived; nor is it clear that the use of even a single-item measure would fully allay this concern, although it might at least minimize it. The potential reactivity of longer craving scales could be studied by measuring craving levels as a function of item position. This would require generating multiple versions of the test with items placed in different positions. If completing the scale was affecting craving, then craving should rise over the course of the questionnaire. Other potential concerns with multi-item scales are that responses to multiple craving items may produce a response bias that could serve to distort responding. It is also possible that responses to initial items may serve to anchor participants as they complete the remaining items.

In a single-factor scale, all the items are posited to measure (albeit imperfectly) a single underlying construct. In this context, the unique meanings of individual items are regarded as

measurement error (in relation to the single construct assumed to underlie all the items). This is precisely why adding items to a scale boosts its reliability; each item imperfectly measures craving, but with unique errors. Across a pool of items the unique errors cancel each other out, thus allowing the "signal" of the craving assessment to rise better above the "noise" inherent in individual items.

The degree to which the items in a scale jointly tap a central construct is indicated by the average correlation between pairs of items (because the correlations are attributed to the items' convergence on the underlying construct). Indeed, the internal consistency of a scale is entirely a function of the average inter-item correlation and the number of items. Figure 1 shows a plot of this function. It can be seen that the number of items required to reach a given degree of reliability (0.80 is a commonly-used criterion) depends on how highly correlated the items are. It should also be evident that, beyond a certain point, adding more items has relatively little impact on reliability. Another implication of this function is that additional items boost reliability only so long as they are correlated with the other items and do not lower the average inter-item correlation. At the same time, it is wise to avoid items that use nearly identical wording, because they may lead to correlated errors of measurement: e.g. if all the items use the word "craving", a participant who misinterprets the word will do so for all the items. This may increase reliability, but at the expense of validity. In constructing craving scales, one also needs to be sure that the items in the scale fully capture the construct; selecting highly-correlated items does not guarantee this. Within these bounds, however, selection of items that are strongly related can offset the need to include many items on a scale.

We should distinguish between the introduction of additional items as a way to boost the reliability of a uni-factorial scale (i.e. one that measures a single unitary construct) vs. the introduction of a more diverse item content, with the intention to capture multiple dimensions or factors thought to comprise or be related to craving. The above holds for uni-factorial scales that are posited to tap a single construct. On the other hand, a multi-factor concept of craving (the basis for multi-factor scales) posits that the construct encompasses several domains that are moderately to weakly correlated with each other. In this case, the differences among items are not necessarily error—they may represent truly different constructs or scales. (Note that, within each scale, all the considerations for uni-factorial scales apply.) Thus, diversity of item content is helpful only if one believes that the construct of craving is heterogeneous and should encompass such diversity.

A recent example of a multi-item craving scale is the Questionnaire of Smoking Urges (QSU: Tiffany & Drobes, 1991). The QSU includes items assessing desire to smoke, expectancies of both positive and negative reinforcement from smoking and intention to smoke. During nicotine-deprived states, a negative reinforcement factor derived from the QSU correlated with a behavioral measure of craving-related processes better than a positive reinforcement factor (Willner, Hardman & Eaton, 1995). In contrast, during a non-deprived state, the QSU–behavior relationship was better for the positive-reinforcement than the negative-reinforcement factor. These data are consistent with a multidimensional view of craving.

The fundamental issue regarding the use of scales such as the QSU is whether items referring to intention to use a drug, and expectancies about the effects of drug use, can be considered components of a broad construct of craving. On the one hand, such divergent concepts have been included in various conceptualizations of craving, and typically factor-analytical studies find items associated with desire and intention to end up on the same factor as more traditional items. On the other hand, while expectancies and intentions may be potentially related, they nevertheless may be conceptually distinct from cravings. Kozlowski *et al.* (1996, p. 257) assert: "We think that a measure of smoking urges should be designed to assess not why smokers want to smoke but the degree to which they desire to smoke a cigarette". These investigators

questioned the utility of craving measures that include multiple content domains. They concluded that, unlike certain personality constructs that require multiple items, the assessment of craving can be reliably and validly measured with a very small set of items (e.g. urge, need, desire). Further research is needed that attempts to show the discriminant validity of items assessing different content areas under theoretically-derived conditions (e.g. smokers trying vs. not trying to quit).

In summary, the number of items and content areas to include in a self-report measure of craving remains among the most important issues facing researchers. When diverse groups of participants are included in the study (e.g. heavy and light drug users), then items with different thresholds of sensitivity may be in order. In situations where craving levels are likely to be high (e.g. during a cue exposure manipulation using drug-abstinent participants), where there is a large number of repeated assessments over a fairly short time period or when craving ratings are assessed vocally, then a relatively small number of items, perhaps even a single item, may be useful in addressing concerns about reactivity. With regard to selecting content areas, some differences in measurement stem from continuing controversy over the nature of the craving construct, and measurement will improve as we improve our theoretical understanding of craving.

Factors affecting choice of self-report measure

Between- vs. within-subject comparisons

One important issue that needs to be considered in designing measures of craving is whether one is primarily interested in between-person differences or within-person changes in craving. Within-subject designs provide inherent controls for many variables that could otherwise influence craving ratings. For example, such designs are somewhat less sensitive to the wording of items and to idiosyncratic interpretation of them by participants, as these effects are likely to hold constant across multiple administrations of the assessment. Conversely, within-subject designs are vulnerable to effects arising from repeated measurements, including the possibility that an earlier assessment actually provokes craving that is then reflected in a later assessment.

When an investigator is trying to characterize the craving response of different groups in a between-subject analysis, particular attention needs to be given to whether the assessment is subject to floor or ceiling effects for one group but not the other. For example, if one administers an assessment to highly dependent and non-dependent drug users, the former may concentrate their ratings at the high end of the scale, such that the scale is insensitive to that groups' experience. Parallel issues arise in within-subject designs if there is a manipulation that raises or lowers participants' craving between assessments; an assessment that was sensitive at one time may not be sensitive at the other.

Craving vs. craving change

Many studies involve manipulating participants' craving from an initial baseline level. Investigators must determine the most meaningful index of craving following the manipulation. One approach is to adjust for baseline levels by calculating a change score, a residual score, or by using baseline scores as covariates in analyses. One rationale for obtaining (and correcting for) baseline measures is that individuals differ in the way they use and respond to the measurement scale. That is, some of the between-person variance in craving scores may be due to consistent measurement error (as distinct from either random error on each measurement occasion or true between-person variance). In this case, including a baseline score allows one to "norm" the relevant measures to each individual's standard. It is beyond the scope of the present paper to contrast change scores and other approaches that adjust for baseline levels and the reader is referred to recent examinations of the topic (e.g. Collins & Horn, 1991). An

assumption of these procedures is that it is possible to establish a stable baseline for craving. Depending on the drug (e.g. nicotine), it can be difficult to identify a time period during which research participants are affected by neither the direct effects of the drug, nor the acute effects resulting from initiation of drug withdrawal (see Hughes, 1991).

The distinction between changes in craving and absolute scores has important implications for interpreting experimental data. It is important to recognize a priori whether the key measure in a particular study is one of craving increase or absolute level of craving. Depending on the purpose of the study or one's conceptualization of craving, the variable of interest may differ. For instance, if one's view of craving is restricted to an intense level of desire, in which a certain level is required before reaching threshold, then the absolute score following exposure may be more important than the increase in craving. Alternatively, if one is measuring the ability of a pharmacological agent to reduce momentary bursts of craving that may occur during a day, then one might make a case for using decreases from peak craving as the key measure.

An illustration of how the choice of measure can fundamentally affect how the data are interpreted is provided by recent data collected with smokers who attended two separate sessions, once while nicotine-deprived, and once in a non-deprived state (Sayette & Hufford, 1994). During each session, participants underwent a smoking cue exposure manipulation that involved holding a lit cigarette. Not surprisingly, smokers rated their pre-cue exposure "baseline" craving higher on the deprived day (about 7 on a 0–10 scale) than the non-deprived day (under 3 on the same scale). In both sessions, following smoking cue exposure, craving reports rose, with self-reported craving greater on the deprived day (above 8) than on the non-deprived day (about 6). Depending on how the data are analyzed, different conclusions can be drawn. Comparisons of the change in craving score, or an analysis of covariance adjusting for baseline differences could lead to the conclusion that smokers were more sensitive to the cue exposure on the *non-deprived* day. Analysis of absolute scores during smoking cue exposure indicates, however, that participants reported stronger absolute craving scores following exposure on the *deprived day*. This latter analysis would be appropriate if one were interested in determining which combination of manipulations had the greatest impact on craving.

Craving inductions

There are several approaches to eliciting craving in a laboratory. These include drug use imagery, administration of stressors and manipulations that expose participants to cues associated with a particular drug including, in some cases, administering small doses of the drug. These different manipulations share the aim of providing internal or external cues that may elicit a craving for the drug.

Issues to keep in mind when selecting a measure in all of these approaches include the expected impact, duration and number of trials of the manipulation. One must balance the relative advantages of a longer scale (e.g. increased sensitivity and reliability), with its potential drawbacks (e.g. increased reactivity). If there are multiple trials, one may need brief forms to permit real time assessment for two reasons. First, it is unfeasible to administer a long multi-item scale to participants numerous times over a short time period. Secondly, when studies use multiple control and craving manipulations in rapid succession (e.g. scripted imagery manipulations) each trial is usually brief. As the time needed to complete the form increases, so too does the danger that assessment might obscure the effects of the manipulation. In some experiments this concern has led to abandoning real-time craving assessment, and relying instead on retrospective craving reports completed following a group of craving and control trials or manipulations (e.g. Elash *et al.*, 1994). In studies in which the only craving manipulation is drug abstinence prior to entering the laboratory or clinic, the craving measure may by itself serve as a cue that augments craving. One option that might be investigated is to imbed the "craving" item(s) within a larger set of questions unrelated to craving.

Recall vs. real-time

Laboratory studies typically, although not always, ask participants to report their craving in real time ("How are you feeling *now*?"). Shiffman *et al.* (1997b) observed that recall of cravings prior to a lapse was inaccurate and biased, compared to recordings made at the time. When participants are asked to *recall* their prior craving, such reports are subject to a variety of biases that can influence autobiographical recall (Hammersley, 1994). Problems with recall are not typically due to the simple decay of memory stores, but to cognitive processes that can result in distortions (Hammersley, 1994); these biases can affect recall even over very short intervals. Recall can be particularly biased if other craving-relevant events (drug use, other craving experiences) are interposed between the target event and the time of recall. This is not to say that real time measures are necessarily "accurate", as any self-report measure can be subject to distortion. Even if retrospective reports are less accurate than real-time recordings, however, they still may be useful. For example, an addict's recall of past craving—whether accurate or not—may influence the expectancies of future craving experience, the selection among treatments, etc. In other words, *beliefs* about craving may influence future behavior. While there have been recent methodological advances using miniature computers that permit real time assessment in field studies (e.g. Shiffman *et al.*, 1997a), the expenses associated with such assessment suggest that there will continue to be retrospective craving assessment in many studies. Clearly, research that takes into account what is known about memory distortions (Hammersley, 1994) is needed to determine the best approaches for retrospective assessment of craving.

Individual differences

There are a host of individual difference factors that may impact on the responses to various types of craving measures. Individual differences in the severity of drug dependence can affect the choice of self-report measure. The experience of craving may be qualitatively different for an occasional user than for an habitual one (Payne *et al.*, 1996). Although there are some data to suggest that women and men may show different patterns of responding to various craving manipulations (Niaura *et al.*, 1998), the impact of gender on the measurement of craving is unclear. Nor is the influence of ethnicity on the measurement of craving understood. Little is known about similarities and differences in cravings across different abusive substances, although it remains a research priority (Institute of Medicine, 1989). Whether cravings differ across routes of drug administration is also unknown.

Non-verbal measures of craving

This section addresses non-verbal approaches to craving measurement. Non-verbal measures of craving used in research have included (a) reinforcement "proxies", (b) drug self-administration, (c) psychophysiological responding, (d) neurobiological responding, (e) cognitive processing and (f) expressive behavior. Interpretation of these responses depends on one's theory of craving. If self-reports of craving are seen as the gold standard for assessment, then the following responses are probably viewed merely as behaviors or reactions that are associated with (self-reported) craving. Accordingly, they are less central to the measurement of craving. In contrast, if one believes that craving is a construct that is only imperfectly indexed by a variety of self-report and non-verbal measures, then non-verbal measures play a more critical role in the assessment of craving. In this case, however, it is important that the selected measures are consistent with a particular theory of craving (e.g. Baker *et al.*, 1987; Tiffany, 1990) and not simply responses that are correlated with reports of craving. It is a challenge to investigators to provide evidence to indicate that changes on non-verbal measures reflect craving, as opposed to a host of alternative functions and experiences.

Drug reinforcement proxies

If craving is considered to reflect a motivation to engage in drug use, then the degree to which anticipated drug use is perceived to be reinforcing can be used as a measure of craving. The level of anticipated drug reinforcement has been assessed using several different measures; often they involve assessment of choice behavior leading to drug administration. Historically, craving has been inferred by determining the cost (e.g. amount of work or pain) that an animal or human will assume in order to obtain a desired drug (Gardner & Lowinson, 1993). Recently, applications of behavioral economics theory to addiction (e.g. Becker & Murphy, 1988; Loewenstein, 1999), have contributed to attempts to quantify the perceived reinforcement value of drugs by asking participants to choose between drug use and varying amounts of money (e.g. Griffiths *et al.*, 1993). Presumably, the greater the value attributed to drug use, the greater the drug craving. Limitations of these measures include individual differences in the perceived value of the alternative reinforcers (10 dollars may be perceived differently by poor compared to wealthy participants). In addition, there is an assumption that the motivation for drug use and the motivation to use or avoid the alternative reinforcer are the only relevant motivations. Yet there can be additional motivations (e.g. motivation to quit) that may influence the behavioral outcome (Ludwig & Wikler, 1974). Although more data are needed to determine the most effective approach to this type of assessment, generally it holds promise and should be considered in craving research.

Drug self-administration

For researchers who consider craving primarily to reflect a behavioral intention to use a drug, direct assessment of drug use may be preferable to self-report measures of craving (Hughes, 1987). Reliance on drug-use behaviors is also preferable among those who oppose use of any craving construct (e.g. Mello, 1978) or those testing animal models of craving. Researchers who suggest that craving need not be a conscious experience (e.g. Berridge & Robinson, 1995) also emphasize measures of drug self-administration in their studies. Drug use can be influenced, however, by a number of forces other than drug craving. One can experience a strong craving but if the drug is unavailable, or if one is determined to quit, drug use may not ensue. Conversely, one might use a drug in the absence of a craving. Even if drug use is simply viewed as a behavior that is related to craving, however, measures of self-administration still may be included in a craving battery as part of the construct validation process (Rankin *et al.*, 1979).

Several measures of drug use have been selected to infer craving in laboratory studies. One approach is to measure the time that elapses from the moment that a participant is permitted to engage in drug use until initiation of drug use (latency). Other measures of drug use topography also have been used, such as amount of alcohol consumed, cigarette puff frequency, puff duration, strength and volume and interpuff interval (Marlatt, Demming & Reid, 1973; Kashinsky, Collins & Brandon, 1995). Latency to first puff or drink may be the component of drug use behavior that is most closely associated with craving, because subsequent drug use behavior may confound the effects of craving with pharmacological effects of the drug, individual differences in drug use history or the automaticity of drug use behavior (Henningfield & Griffiths, 1979; West & Schneider, 1987; Tiffany, 1990).

Psychophysiological responding

According to a number of models, psychophysiological responses should index craving. Changes in heart rate, skin temperature, blood pressure, skin conductance and salivation have been included in craving studies, although the pattern of physioloical responding has often differed (Niaura et al., 1988; Glautier, Drummond & Remington, 1992; Carter & Tiffany, 1999). Presumably these measures are less vulnerable to conscious control and thus may be more sensitive than self-report measures to detecting craving (Baker & Brandon, 1990).

Psychophysiological measures have been criticized for a number of reasons. Psychophysiological systems serve functions that are independent of a motivation for drug use. In addition, certain physiological responses to drug use, or even to situations associated with craving, can be unrelated to a motivation to use the drug. That is, an increase in cardiac activity during a craving manipulation may relate to effort associated with coping with the situation, a state of physical or cognitive preparedness to acquire and ingest the drug, among other functions (Baker & Brandon, 1990; Tiffany, 1990). Further, it is unclear what patterns of psychophysiological response ought to be related to drug craving (Niaura *et al.*, 1988; Tiffany, 1990). For example, both increases and decreases in heart rate have been considered to be indicative of craving, depending on the model invoked. Indeed, some investigators have argued that physiological changes have not been linked in theoretically meaningful ways to craving (Kassel & Shiffman, 1992; Pickens & Johanson, 1992) and even have wondered whether selection of psychophysiological measures is more a function of convenience than theory (Tiffany, 1991).

Startle reflex

Emotion researchers (e.g. Lang, Bradley & Cuthbert, 1990) have found that the magnitude of the reflexive eye blink to startling stimuli captures the processing of both positive (decreased startle) and negative (increased startle) emotional states. Recently this measure has been used in laboratory investigations to assess craving. Elash *et al.* (1995) found smoking-related imagery to increase startle responding, suggesting an increase in the processing of negative affect. Others have found smoking cue exposure to attenuate prepulse inhibition of the startle reflex, a measure thought to be associated with changes in information processing (Hutchison *et al.*, 1997). Advantages of measuring startle response in studies of craving include its (a) relative ease of measurement, (b) noninvasiveness, (c) reflexive nature, which reduces the likelihood of response biases and (d) good temporal stability. A disadvantage of startle probe methods is a lack of specificity. As more data are collected, the role of startle reflex in craving assessment will be better understood.

Neurobiological responding

Recently, studies using both animals and human participants have isolated brain structures and processes that may underlie craving (Wise, 1988; Robinson & Berridge, 1993). Measures of glucose metabolism [e.g. using positron emission tomography (PET) scans] with human participants reveal metabolic increases during craving manipulations in particular brain structures associated with both emotional (e.g. amygdala) and cognitive (e.g. hippocampus) aspects of memory (Everitt, 1997). Nevertheless, it is important to recognize that imaging studies can only show that a specific craving manipulation produces increases in indices of brain activation. Neurobiological changes may be correlated with craving, but like other psychophysiological measures these changes do not necessarily produce craving or cause the effects that are observed on other dependent variables (Van Orden & Paap, 1997). The distinction rests in the differentiation between a measure of craving and the phenomenon of craving itself. At this point, research using neuroimaging methodology to study craving is in its infancy and more research is indicated to understand the neural circuitry of craving.

Cognitive processing

The view that craving is associated with changes in cognitive processing has a long history (Sorokin, 1942; Keys *et al.*, 1950). Some posit that craving requires, or at least is associated with, a shift in non-automatic processing resources (Tiffany, 1990; Sayette & Hufford, 1994). One approach to assessing redistribution of cognitive resources during craving involves divided attention tasks. An example of this approach is the secondary response time (RT) paradigm. RT paradigms have been used to identify the degree to which a primary task draws on limited-

capacity cognitive resources, by recording performance decrements on a secondary RT task. When more cognitive resources are required in a primary task, less capacity remains available for a secondary RT task and RT latency increases (Kerr, 1973; Wickens, 1984). Studies using both smokers and alcoholic participants have revealed increases in secondary RTs during peak craving periods, relative to non-craving baseline periods (Sayette et al., 1994; Sayette & Hufford, 1994; Cepeda-Benito & Tiffany, 1996; Juliano & Brandon, 1998).

Another cognitive performance approach to craving assessment involves explicit memory tasks such as cued recall. Presumably, words that can be recalled most readily from a previously learned list are most salient to an individual. Consistent with this formulation, abstinence appears to increase ability to recall already presented smoking-related words (Zeitlan, Potts & Hodder, 1994). One difficulty with this measure which may reduce its sensitivity is that once participants realize that some of the words are related to a drug then their organizational framework for encoding items is altered, which may facilitate retrieval of subsequently presented drug-related words. That is, increased recall of drug-related words may be due to encoding strategies rather than the salience produced by craving.

Several implicit memory tasks have assessed processes associated with craving. The general purpose of these tasks is to examine the salience of drug-related information for individuals without making them aware of the purpose of assessment. This salience presumably reflects a cognitive preoccupation with drug-related material. Stem completions, perceptual identification, categorization tasks and color-naming variants of the Stroop task have all been used (Litz, Payne & Colletti, 1987; Jarvik *et al.*, 1995; Zeitlin *et al.*, 1994). These studies of implicit memory have provided conflicted findings, with drug deprivation leading to facilitation (e.g. Gross *et al.*, 1993; Jarvik *et al.*, 1995) as well as inhibition (e.g. Zeitlin *et al.*, 1994) in the selective processing of drug-related information. A concern with these tasks is that other sources of priming besides craving can influence performance. As with the explicit memory tasks, participants, often aware that they are involved in a drug study, may recognize a drug-related word in the beginning of a task, which then primes subsequent recognition of other drug-relevant words, independent of an underlying preoccupation with smoking. This effect is known as response chaining. One response to this concern is to focus exclusively on initial responses (Stacy, Leigh & Weingardt, 1994).

In summary, cognitive tasks have proven useful in suggesting memory structures associated with craving as well as cognitive processes that may change during craving. As with other measures, however, it is important for researchers to show that the observed changes are related to craving. This requires articulation of a craving definition, as well as appropriate experimental procedures, in which factors such as drug use history, level of drug abstinence and presence of drug-related stimuli are manipulated or controlled.

Expressive behavior

Analysis of facial expressive behavior is an especially important measure of emotional responses (Darwin, 1872/1965;Barlow, 1988). Because craving may be affective in nature, the use of an established, anatomically based facial coding system may prove useful for detecting craving. Depending on one's model of craving, these facial signals may reflect craving-related affect or a measure of craving itself—if one views craving to be fundamentally affective in nature (e.g. Baker *et al.*, 1987). Advantages of such objective facial coding systems include the capacity to comprehensively measure immediate responses in an unobtrusive manner, to assess positive and negative affects as orthogonal dimensions, and even to discriminate between different emotions (see Ekman & Rosenberg, 1997).

Using the Facial Action Coding System (FACS: Ekman & Friesen, 1978) to code participants' responses to smoking cues under two experimental conditions, Sayette & Hufford (1995) found

that high craving ratings could be linked to both positive affect-related signals (when expecting to smoke) and negative affect-related signals (when not permitted to smoke). Limitations of FACS include reliance on expressions that must be visible to humans, the need for extensive coder training [more than 100 hours (Ekman, 1982)], as well as the extremely time-consuming nature of coding facial expressions. Consequently, use of FACS may be restricted to brief "critical" craving coding intervals. Future research may identify specific facial action patterns that are associated with craving. If so, coding time could be reduced substantially by coding only these pre-selected action units. Recent advances in the use of computer vision-based approaches to facial expression analysis hold promise for the examination of facial movements related to craving. These approaches convert videotape into digital signals that can be fed into a computer. Initial studies have already established high correspondence with FACS for a subset of facial muscle movements (e.g. Cohn *et al.*, 1999).

An alternative measure of facial muscle activity is facial electromyography (EMG), which has been used in studies of craving (e.g. Elash *et al.*, 1995; Drobes & Tiffany, 1997). Facial EMG can detect even those subacute muscle movements that are not visible to the human eye. Evidence suggests that EMG can effectively assess global positive and negative facial expressions (e.g. Caccioppo *et al.*, 1988). A disadvantage of facial EMG is that the ability to capture a range of discrete emotions (e.g. fear, anger, sadness, disgust, happiness) requires use of increasing numbers of electrodes, which may interfere with a participant's natural craving experience. Moreover, studies indicate that emotional specificity is difficult to assess because several different negative emotional responses may activate the same muscle (corrugator supercillii) (Brown & Schwartz, 1980; McHugo *et al.*, 1985; Sirota, Schwartz & Kristeller, 1987; Lang *et al.*, 1993). As is true with facial coding, activation of a particular muscle is not specific to craving. Consequently, use of any of these measures of expressive behavior is best restricted to laboratory environments in which craving is elicited and appropriate experimental controls are employed.

In summary, several types of non-verbal measures have been used to study craving. While generally they may be less vulnerable to the response biases or reactivity created by self-report assessment, they still are susceptible to a range of conceptual and measurement problems. Some may lack the sensitivity to reveal craving. Even more problematic are concerns with specificity, in which changes on these measures often are due to factors other than craving. Most importantly, how these non-verbal measures relate to self-report measures of craving depends entirely on one's theory of craving. Thus, articulation of a craving theory is needed to best interpret these non-verbal responses.

Conclusion

Growing interest in the study of craving requires examination of how craving—and processes associated with craving—are measured. We have argued that it is important for investigators to articulate a definition of craving, and a theoretical framework for which the selected measures are conceptually relevant. Currently, investigators assembling a set of craving measures can choose from among a wide range of self-report and non-verbal approaches, some of which have only recently been introduced into the literature. None of the measures is appropriate across all settings and thus measurement selection poses a formidable challenge. Relevant issues include the anticipated potency of the craving experience, amount of time available for administering the measures and number of times that the measure will be administered. Awareness of the limitations and strengths of different measures should lead to improved craving assessment.

While some guidelines have been offered to help select appropriate measures, it is clear that theory-driven, multi-method research contrasting different assessment approaches and formats

is needed. In addition, a multi-dimensional approach to assessment permits examination of the coherence of different measures across response systems. By systematically manipulating craving levels, putative measures of craving can be evaluated. Such efforts are required in order to provide support for the construct validity of craving. Improved measurement is essential if we are to optimize research efforts directed at better understanding drug craving.

Acknowledgement

Preparation of this manuscript was supported in part by grants from the National Institute on Alcohol Abuse and Alcoholism (AA09918) and the National Institute on Drug Abuse (DA10605) to Michael Sayette.

References

Acri JB, Grunberg NE. A psychophysical task to quantify smoking cessation-induced irritability: the reactive irritability scale (RIS). Addictive Behaviors 1992;17:587–601. [PubMed: 1488939]

American Educational Research Association, American Psychological Association & National Council for Measurement in Education. Standards for educational and psychological testing. Washington, DC: APA; 1985.

Anastasi, A. Psychological testing. Vol. 6th edn. New York: MacMillan; 1988.

Anton R, Moak D, Latham P. The obsessive compulsive drinking scale: a self-rated instrument for the quantification of thoughts about alcohol and drinking behavior. Alcoholism: Clinical and Experimental Research 1995;19:92–99.

Avants SK, Margolin A, Kosten TR, Cooney NL. Differences between responders and nonresponders to cocaine cues in the laboratory. Addictive Behaviors 1995;20:215–224. [PubMed: 7484315]

Baker, TB.; Morse, E.; Sherman. The motivation to use drugs: a psychobiological analysis of urges. In: Rivers, C., editor. The Nebraska Symposium on Motivation: alcohol use and abuse. Lincoln: University of Nebraska Press; 1987. p. 257-323.

Baker, TB.; Brandon, TH. Behavioral Assessment. Vol. 12. 1990. Validity of self-reports in basic research; p. 33-51.

Barlow, DH. Anxiety and its Disorders. Guilford Press; 1988.

Becker G, Murphy K. A theory of rational addiction. Journal of Political Economy 1988;96:675–700.

Berridge KC, Robinson TE. The mind of an addicted brain: neural sensitization of wanting versus liking. Current Directions in Psychological Science 1995;4:71–76.

Bohn MJ, Krahn DD, Staehler BA. Development and initial validation of a measure of drinking urges in abstinent alcoholics. Alcoholism: Clinical and Experimental Research 1995;19:600–606.

Brown S, Schwartz GE. Relationships between facial electromyography and subjective experience during affective imagery. Biological Psychology 1980;11:49–62. [PubMed: 7248403]

Buchwald, AM. Verbal utterances as data. In: Feigl, H.; Maxwell, G., editors. Current ssues in the philosophy of science. New York: Holt; 1961. p. 461-468.

Buydens-Branchey L, Branchey M, Fergeson P, Hudson J, McKernin C. Hormonal, psychological, and alcohol craving changes after m-chlorophenylpiperazine administration in alcoholics. Alcoholism: Clinical and Experimental Research 1997;21:220–226.

Caccioppo JT, Martzke JS, Petty RE, Tassinary LG. Specific forms of facial EMG response index emotions during an interview: from Darwin to the continuous flow hypothesis of affect-laden information processing. Journal of Personality and Social Psychology 1988;54:592–604. [PubMed: 3367281]

Campbell DT, Fiske DW. Convergent and discriminative validation by the multitrait–multimethod matrix. Psychological Bulletin 1959;56:81–105. [PubMed: 13634291]

Carlsmith, JM.; Ellesworth, PC.; Aronson, E. Methods of Research in Social Psychology. Reading, Massachusetts: Addison Wesley; 1976.

Carter BL, Tiffany ST. Meta-analysis of cue reactivity in addiction research. Addiction 1999;94:327–340. [PubMed: 10605857]

Cepeda-Benito A, Tiffany ST. The use of a dual-task procedure for the assessment of cognitive effort associated with cigarette craving. Psychopharmacology 1996;127:155–163. [PubMed: 8888382]

Childress AR, McLellan AT, O'Brien CP. Conditioned responses in a methadone population: a comparison of laboratory, clinic, and nature settings. Journal of Substance Abuse Treatment 1986;3:173–179. [PubMed: 3806730]

Clark, DM.; Salkovskis, PM.; Gelder, M.; Koehler, C.; Martin, M.; Anastasiades, P.; Hackman, A.; Middleton, H.; Jeavons, A. Tests of a cognitive theory of panic. In: Hand, I.; Wittchen, H., editors. Panic and Phobias. New York: Springer-Verlag; 1988. p. 149-158.

Cohn, JF.; Zlochower, AJ.; Lien, J.; Kanade, T. Psychophysiology. Vol. 36. 1999. Automated face analysis by feature point tracking has high concurrent validity with manual FACS coding; p. 35-43.

Collins, LM.; Horn, JL., editors. Best Methods for the Analysis of Change: recent advances, unanswered questions, future directions. Washington, DC: American Psychological Association; 1991.

Cook, TD.; Campbell, DT. Quasi-experimentation. Chicago: Rand McNalley; 1979.

Cook SW, Selltiz C. A multiple indicator approach to attitude measurement. Psychological Bulletin 1964;62:36–55. [PubMed: 14176651]

Cronbach LJ, Meehl PE. Construct validity in psychological tests. Psychological Bulletin 1955;52:281–302. [PubMed: 13245896]

Darwin, C. The Expression of the Emotions in Man and Animals. Chicago: University of Chicago Press; 1965.

Drobes DJ, Tiffany ST. Induction of smoking urge through imaginal and *in vivo* procedures: physiological and self-report manifestations. Journal of Abnormal Psychology 1997;106:15–25. [PubMed: 9103714]

Droungas A, Ehrman RN, Childress AR, O'Brien CP. Effect of smoking cues and cigarette availability on craving and smoking behavior. Addictive Behaviors 1995;20:657–673. [PubMed: 8712062]

Drummond CD, Cooper T, Glautier SP. Conditioned learning in alcohol dependence: implications for cue exposure treatment. British Journal of Addiction 1990;85:725–743. [PubMed: 2198966]

Ekman, P.; Friesen, WV. Facial Action Coding System. Palo Alto, CA: Consulting Psychologists Press; 1978.

Ekman, P. Methods for measuring facial action. In: Scherer, KR.; Ekman, P., editors. Handbook of Methods in Nonverbal Behavior Research. Cambridge: Cambridge University Press; 1982. p. 45-90.

Ekman, P.; Rosenberg, EL., editors. What the Face Reveals: basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS). New York: Oxford University Press; 1997.

Ekman G, Frankenhaeuser M, Goldberg L, Hagdahl R, Myrsten A. Subjective and objective effects of alcohol as functions of dosage and time. Psychopharmacologia 1964;6:399–409. [PubMed: 5830962]

Elash CA, Tiffany ST, Vrana SR. Manipulation of smoking urges and affect through a brief-imagery procedure: self-report, psychophysiological, and startle probe responses. Experimental and Clinical Psychopharmacology 1995;3:156–162.

Everitt B. Craving cocaine cues: cognitive neuroscience meets drug addiction research. Trends in Cognitive Sciences 1997;1:1–2.

Fiske DW, Campbell DT. Citations do not solve problems. Psychological Bulletin 1992;12:393–395.

Gardner EL, Lowinson JH. Drug craving and positive/negative hedonic brain substrates activated by addicting drugs. Seminars in the Neurosciences 1993;5:359–368.

Gawin FH. Cocaine addiction: psychology and neurophysiology. Science 1991;251:1580–1586. [PubMed: 2011738]

Geisinger KF. The metamorphosis of test validation. Educational Psychologist 1992;27:197–222.

Glautier S, Drummond DC, Remington B. Different drink cues elicit different physiological responses in nondependent drinkers. Psychopharmacology 1992;106:550–554. [PubMed: 1579627]

Griffiths RR, Troisi JR, Silverman K, Mumford GK. Multiple choice procedure: an efficient approach for investigating drug reinforcement in humans. Behavioural Pharmacology 1993;4:3–13. [PubMed: 11224166]

Gross TM, Jarvik ME, Rosenblatt MR. Nicotine abstinence produces content-specific Stroop interference. Psychopharmacology 1993;110:333–336. [PubMed: 7831427]

Grossberg JM, Grant BF. Clinical psychophysics: applications of ratio scaling and signal detection methods to research on pain, fear, drugs, and medical decision making. Psychological Bulletin 1978;85:1154–1176. [PubMed: 704721]

Hammersley R. A digest of memory phenomena for addiction research. Addiction 1994;89:283–293. [PubMed: 8173494]

Henningfield JE, Griffiths RR. A preparation for the experimental analysis of human cigarette smoking behavior. Behavior Research Methods and Instrumentation 1979;11:538–544.

Holden RR, Fekken GC, Jackson DN. Structured personality test item characteristics and validity. Journal of Research in Personality 1985;19:386–394.

Hughes JR. Craving as a psychological construct. British Journal of Addiction 1987;82:38–39.

Hughes JR. Distinguishing withdrawal relief and direct effects of smoking. Psychopharmacology 1991;104:409–410. [PubMed: 1924648]

Hutchison K, Niaura R, Swift R. The effects of high vs. low nicotine cigarettes on prepulse inhibition. Psychophysiology 1997;34:S45.

Institute of Medicine. Prevention and Treatment of Alcohol Problems: research opportunities. Washington, DC: National Academy Press; 1989.

Jackson, DN. Acquiescence response styles: problem of identification and control. In: Berg, IA., editor. Response Set in Personality Assessment. Chicago: Aldine; 1967. p. 71-114.

Jarvik M, Gross T, Rosenblatt M, Stein R. Enhanced lexical processing of smoking stimuli during smoking abstinence. Psychopharmacology 1995;118:136–141. [PubMed: 7617799]

Juliano LM, Brandon TH. Reactivity to instructed smoking availability and environmental cues: evidence with urge and reaction time. Experimental and Clinical Psychopharmacology 1998;6:45–53. [PubMed: 9526145]

Kaplan RF, Cooney NL, Baker LH, Gillespie RA, Meyer RE, Pomerlau OF. Reactivity to alcohol-related cues: physiological and subjective responses in alcoholics and non-problem drinkers. Journal of Studies on Alcohol 1985;46:267–272. [PubMed: 4033125]

Kashinsky W, Collins BN, Brandon TH. A telemetric device for measuring smoking topography. Behavior Research Methods, Instruments, and Computers 1995;27:375–378.

Kassel JD, Shiffman S. What can hunger teach us about drug craving? A comparative analysis of the two constructs. Advances in Behaviour Research and Therapy 1992;14:141–167.

Kerr B. Processing demands during mental operations. Memory and Cognition 1973;1:401–412.

Keys, A.; Brozek, J.; Henschel, A.; Mickelsen, O.; Taylor, HL. The biology of human starvation. Vol. vol. 2. Minneapolis: University of Minnesota Press; 1950.

Killen JD, Fortmann SP. Craving is associated with smoking relapse: findings from three prospective studies. Experimental and Clinical Psychopharmacology 1997;5:137–142. [PubMed: 9234050]

Kozlowski LT, Mann RE, Wilkinson DA, Poulos CX. “Cravings” are ambiguous: ask about urges or desires. Addictive Behaviors 1989;14:443–445. [PubMed: 2782125]

Kozlowski LT, Pillitteri JL, Sweeney CT, Whitfield KE, Graham JW. Asking questions about urges or cravings for cigarettes. Psychology of Addictive Beahviors 1996;10:248–260.

Kozlowski LT, Wilkinson DA. Use and misuse of the concept of craving by alcohol, tobacco, and drug researchers. British Journal of Addiction 1987;82:31–36. [PubMed: 3470042]

Lang PJ, Bradley MM, Cuthbert BN. Emotion, attention, and the startle reflex. Psychological Review 1990;97:377–395. [PubMed: 2200076]

Lang PJ, Greenwald MK, Bradley MM, Hamm AO. Looking at pictures: affective, facial, visceral, and behavioral reactions. Psychophysiology 1993;30:261–273. [PubMed: 8497555]

Litz B, Payne T, Colletti G. Schematic processing of smoking information by smokers and nonsmokers. Cognitive Therapy and Research 1987;11:301–313.

Loewenstein, G. A visceral account of addiction. In: Elster, J.; Skog, OJ., editors. Getting Hooked: rationality and addiction. Cambridge, UK: Cambridge University Press; 1999. p. 235-264.

Lowman C, Hunt WA, Litten RZ, Drummond DC. Research perspectives on alcohol craving: an overview. Addiction 2000;95:S45–S54. [PubMed: 11002901]

Ludwig AM, Wikler A. “Craving” and relapse to drink. Quarterly Journal of Studies on Alcohol 1974;35:108–130.

Ludwig AM, Wikler A, Stark LH. The first drink: psychobiological aspects of craving. Archives of General Psychiatry 1974;30:539–547. [PubMed: 4131353]

Maloney, MP.; Ward, MP. Psychological Assessment: a conceptual approach. New York: Oxford University Press; 1976. p. 66

Mark MM, Sinclair RC, Wellens TR. The effect of completing the Beck Depression Inventory on self-reported mood state: contrast and assimilation. Personality and Social Psychology Bulletin 1991;17:457–465.

Marlatt, GA. Cognitive factors in the relapse process. In: Marlatt, GA.; Gordon, JR., editors. Relapse Prevention: maintenance strategies in the treatment of addictive behaviors. New York: Guilford Press; 1985. p. 128-200.

Marlatt GA. Craving notes. British Journal of Addiction 1987;82:42–44.

Marlatt GA, Demming B, Reid JB. Loss of control drinking in alcoholics: an experimental analogue. Journal of Abnormal Psychology 1973;81:223–241.

McCusker CG, Brown K. The cue-responsivity phenomenon in dependent drinkers: “personality” vulnerability and anxiety as intervening variables. British Journal of Addiction 1991;86:905–912. [PubMed: 1912743]

McHuco GJ, Lanzetta JT, Sullivan DG, Masters RD, Englis BG. Emotional reactions to a political leader’s expressive displays. Journal of Personality and Social Psychology 1985;49:1513–1529.

Mello, NK. A semantic aspect of alcoholism. In: Cappell, HD.; LeBlanc, AE., editors. Biological and Behavioral Approaches to Drug Dependence. Toronto: Addiction Research Foundation; 1978. p. 73-87.

Miller NS, Gold MS. Dissociation of “conscious desire” (craving) from and relapse in alcohol and cocaine dependence. Annals of Clinical Psychiatry 1994;6:99–106. [PubMed: 7804394]

Niaura RS, Rohsenow DJ, Binkoff JA, Monti PM, Pedraza M, Abrams DB. Relevance of cue reactivity to understanding alcohol and smoking relapse. Journal of Abnormal Psychology 1988;97:133–152. [PubMed: 3290304]

Niaura RS, Shadel WG, Abrams DB, Monti PM, Rohsenow DJ, Sirota A. Individual differences in cue reactivity among smokers trying to quit: effects of gender and cue type. Addictive Behaviors 1998;23:209–224. [PubMed: 9573425]

Payne TJ, Smith PO, Sturges LV, Holleran SA. Reactivity to smoking cues: mediating roles of nicotine dependence and duration of deprivation. Addictive Behaviors 1996;21:139–154. [PubMed: 8730517]

Perkins K, Epstein LE, Grobes J, Fonte C. Tobacco abstinence, smoking cues, and the reinforcing value of smoking. Pharmacology, Biochemistry, and Behavior 1994;47:107–112.

Perlmuter LC, Noblin CD, Hakami M. Reactive effects of tests for depression: theoretical and methodological considerations. Journal of Social and Clinical Psychology 1983;1:128–139.

Pickens RW, Johanson C. Craving: consensus of status and agenda for future research. Drug and Alcohol Dependence 1992;30:127–131. [PubMed: 1633752]

Powell J. Conditioned responses to drug-related stimuli: is context crucial? Addiction 1995;90:1089–1095. [PubMed: 7549777]

Powell J, Gray JA, Bradley BP, Kasvikis Y, Strang J, Barratt L, Marks I. The effects of exposure to drug-related cues in detoxified addicts: a theoretical review and some new data. Addictive Behaviors 1990;15:339–354. [PubMed: 2248108]

Rankin H, Hodgson R, Stockwell T. The concept of craving and its measurement. Behaviour Research and Therapy 1979;17:389–396. [PubMed: 486042]

Robinson TE, Berridge KC. The neural basis of drug craving: an incentive–sensitization theory of addiction. Brain Research Reviews 1993;18:247–291. [PubMed: 8401595]

Rohsenow DJ, Niaura RS, Childress AR, Abrams DB, Monti PM. Cue reactivity in addictive behaviors: theoretical and treatment implications. International Journal of the Addictions 199091;25:957–993. [PubMed: 2131326]

Sayette MA, Hufford MR. Effects of cue exposure and deprivation on cognitive resources in smokers. Journal of Abnormal Psychology 1994;103:812–818. [PubMed: 7822584]

Sayette MA, Hufford MR. Urge and affect: a facial coding analysis of smokers. Experimental and Clinical Psychopharmacology 1995;3:417–423.

Sayette MA, Monti PM, Rohsenow DJ, Bird-Gulliver S, Colby S, Sirota A, Niaura RS, Abrams DB. The effects of cue exposure on attention in male alcoholics. Journal of Studies on Alcohol 1994;55:629–634. [PubMed: 7990473]

Sayette MA, Wilson GT. Intoxication and exposure to stress: effects of temporal patterning. Journal of Abnormal Psychology 1991;100:56–62. [PubMed: 2005272]

Schreisman CA, Hill KD. Controlling acquiescence response bias by item reversals: the effect on questionnaire validity. Educational and Psychological Measurement 1981;41:1101–1114.

Schuster CR, Greenwald MK, Johanson CE, Heishman SJ. Measurement of drug craving during naloxone-precipitated withdrawal in methadone-maintained volunteers. Experimental and Clinical Psychopharmacology 1995;3:424–431.

Shadel WG, Niaura R, Brown RA, Hutchison KE, Abrams D. A content analysis of smoking craving. Journal of Clincal Psychology. 2000in press

Shiffman S. The concept of craving: let us not throw out the baby with the bathwater. British Journal of Addiction 1987;82:37–38.

Shiffman S, Engberg J, Paty JA, Perz W, Gnys M, Kassel JD, Hickcox M. A day at a time: predicting smoking lapse from daily urge. Journal of Abnormal Psychology 1997a;106:104–116. [PubMed: 9103722]

Shiffman S, Hufford M, Hickox M, Paty JA, Gnys M, Kassel J. Remember that? A comparison of real-time versus retrospective recall of smoking lapses. Journal of Consulting and Clinical Psychology 1997b;65:292–300. [PubMed: 9086693]

Shiffman S, Paty JA, Gnys M, Kassel JD, Elash C. Nicotine withdrawal in chippers and regular smokers: subjective and cognitive effects. Health Psychology 1995;14:301–309. [PubMed: 7556033]

Shiffman S, Paty JA, Gnys M, Kassel JD, Hickcox M. First lapses to smoking: within-subjects analyses of real-time reports. Journal of Consulting and Clinical Psychology 1996;64:366–379. [PubMed: 8871421]

Sirota AD, Schwartz GE, Kristeller JL. Facial muscle activity during induced mood states: differential growth and carry-over of elated versus depressed patterns. Psychophysiology 1987;24:691–699. [PubMed: 3438434]

Sorokin, PA. Man and Society in Calamity. New York: Dutton; 1942.

Stacy AW, Leigh BC, Weingardt KR. Memory accessibility and association of alcohol use and its positive outcomes. Experimental and Clinical Psychopharmacology 1994;2:269–282.

Stevens SS. The measurement of loudness. Journal of the Accoustical Society of America 1955;27:815–829.

Tiffany ST. A cognitive model of drug urges and drug-use behavior: role of automatic and nonautomatic processes. Psychological Review 1990;97:147–168. [PubMed: 2186423]

Tiffany ST. The application of 1980s psychology to 1990s smoking research. British Journal of Addiction 1991;86:617–620. [PubMed: 1859929]

Tiffany ST. A critique of contemporary urge and craving research: methodological, psychometric, and theoretical issues. Advances in Behaviour Research and Therapy 1992;14:123–139.

Tiffany ST, Carter BL, Singleton EG. Challenges in the manipulation, assessment and interpretation of craving relevant variables. Addiction 2000;95:S177–S187. [PubMed: 11002913]

Tiffany ST, Drobes DJ. The development and initial validation of a questionnaire on smoking urges. British Journal of Addiction 1991;86:1467–1476. [PubMed: 1777741]

Tiffany ST, Singleton E, Haertzen CA, Henningfeld JE. The development of a cocaine craving questionnaire. Drug and Alcohol Dependence 1993;34:19–28. [PubMed: 8174499]

Turner CF, Ku L, Rogers SM, Lindberg LD, Pleck JH, Sonenstein FL. Adolescent sexual behavior, drug use, and violence: increased reporting with computer survey technology. Science 1998;280:867–873. [PubMed: 9572724]

Van Orden GC, Paap KR. Functional neuroimages fail to discover pieces of the mind in parts of the brain. Philosophy of Science 1997;64:S85–S94.

Voris J, Elder I, Sebastian P. A simple test of cocaine craving and related responses. Journal of Clinical Psychology 1991;47:320–323. [PubMed: 2030141]

West RJ. Use and misuse of craving. British Journal of Addiction 1987;82:39–41.

West RJ, Hajek P, Belcher M. Severity of withdrawal symptoms as a predictor of outcome of an attempt to quit smoking. Psychological Medicine 1989;19:981–985. [PubMed: 2594893]

West RJ, Schneider N. Craving for cigarettes. British Journal of Addiction 1987;82:407–415. [PubMed: 3555575]

Wickens, CD. Processing resources in attention. In: Parasuraman, R.; Davis, R.; Beathy, J., editors. Varieties of Attention. New York: Academic Press; 1984. p. 63-102.

Wiggens, JS. Personality and Prediction. Principles of personality assessment. Reading, MA: Addison-Wesley; 1973.

Wikler A. Recent progress in research on the neurophysiological basis of morphine addiction. American Journal of Psychiatry 1948;105:329–338. [PubMed: 18890902]

Willner P, Hardman S, Eaton G. Subjective and behavioural evaluation of cigarette cravings. Psychopharmacology 1995;118:171–177. [PubMed: 7617804]

Wise RA. The neurobiology of craving: implications for the understanding and treatment of addiction. Journal of Abnormal Psychology 1988;97:118–132. [PubMed: 3290303]

World Health Organization. The craving for alcohol: a symposium by members of the WHO Expert Committees on Mental Health and on Alcohol. Quarterly Journal of Studies on Alcohol 1955;16:33–66.

Zeitlan, SB.; Potts, AJ.; Hodder, SL. Implicit and explicit memory biases for smoking-related cues in nicotine abstinent and non-abstinent smokers (Summary). Convention Proceedings for the 28th Annual Meetings of the Association for advancement of Behavior Therapy; 1994. p. 160

Zinser MC, Baker TB, Sherman JE, Cannon DS. Relation between self-reported affect and drug urges and cravings in continuing and withdrawing smokers. Journal of Abnormal Psychology 1992;101:617–629. [PubMed: 1430600]

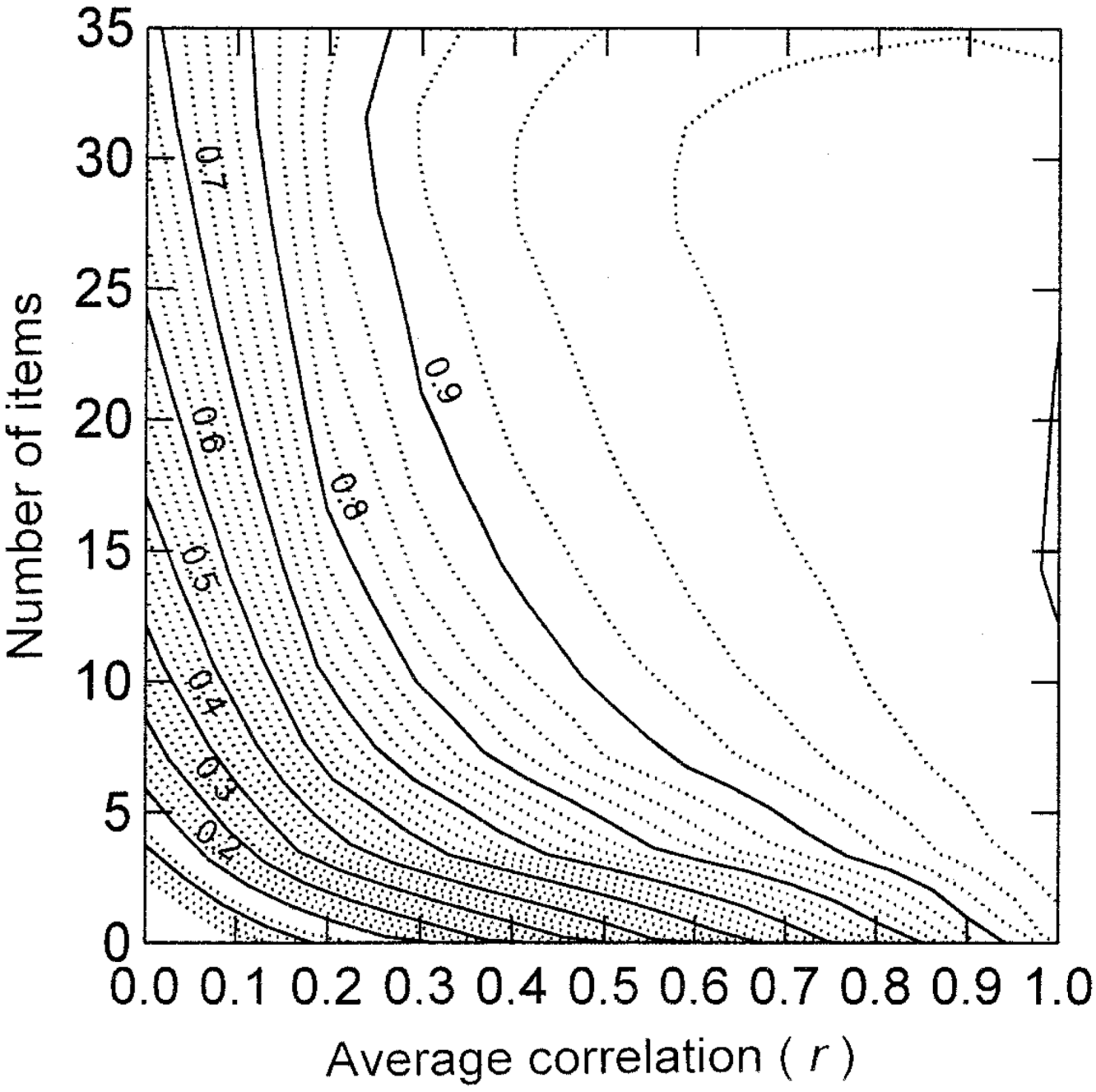

Figure 1.
Reliability (alpha) as a function of items and correlation. The figure shows the computed reliability coefficient (Cronbachs' alpha) as a joint function of the number of items in a scale and the average intercorrelation among them. The lines within the graph represent the computed alpha.