How to measure mood in nutrition research

Richard Hammersley1*, Marie Reid1 and Stephen L. Atkin2
1Department of Psychology, University of Hull, Hull HU6 7RX, UK
2Department of Medicine, Weill Cornell Medical College in Qatar, Doha, Qatar

Abstract
Mood is widely assessed in nutrition research, usually with rating scales. A core assumption is that positive mood reinforces ingestion, so it is important to measure mood well. Four relevant theoretical issues are reviewed: (i) the distinction between protracted and transient mood; (ii) the distinction between mood and emotion; (iii) the phenomenology of mood as an unstable tint to consciousness rather than a distinct state of consciousness; (iv) moods can be caused by social and cognitive processes as well as physiological ones. Consequently, mood is difficult to measure and mood rating is easily influenced by non-nutritive aspects of feeding, the psychological, social and physical environment where feeding occurs, and the nature of the rating system employed. Some of the difficulties are illustrated by reviewing experiments looking at the impact of food on mood. The mood-rating systems in common use in nutrition research are then reviewed, the requirements of a better mood-rating system are described, and guidelines are provided for a considered choice of mood-rating system including that assessment should: have two main dimensions; be brief; balance simplicity and comprehensiveness; be easy to use repeatedly. Also mood should be assessed only under conditions where cognitive biases have been considered and controlled.

Key words: Affect: Mood assessment: Mood rating

Introduction
Measuring mood is important in nutrition research because changes in mood motivate human ingestion choices, including what is eaten and drunk, when, and how often, which can have major consequences for health. Mood is usually measured with some type of questionnaire rating system. Common dimensions of mood include subjective energy or level of arousal, positive v. negative mood or good v. bad mood, and tension v. calmness.

The present paper reviews theory underlying both mood and mood assessment, drawing three important distinctions: between mood and emotion; between protracted and transient mood; that transient mood is probably not a state of consciousness, but rather a tint to consciousness. The complex causes of mood also will be discussed, which raise difficulties for assessing mood changes that might occur after eating or drinking and for considering how observed changes in mood ratings should be interpreted. The paper goes on to briefly review experiments that have focused on the effects of food on mood, to review both the common and some alternative methods of rating mood, and to suggest good practice in choosing and interpreting mood-rating systems in nutrition research.

It will be proposed that: transient mood assessment needs to be in accord with contemporary theories of mood, which prefer mood to have two main dimensions; the questionnaire or other assessment needs to be brief and administrable quickly before transient mood is altered by time and the processes of questionnaire completion; assessment should strike a balance between simplicity and comprehensiveness; the assessment method should be easy to use in order to minimise bias and error when it is being completed repeatedly; mood should be assessed only under conditions where theoretically irrelevant cognitive factors that can influence mood rating have been considered and controlled.

It is difficult to measure mood, for it is inherently phenomenological, so methods of mood assessment are problematic(1,2). Nonetheless, ingestion is a major method that individuals use to try to manage their conscious mood state, in a complexity of ordinary foods, nutritional foods, functional foods, medicines, psychoactive drugs, vitamins and dietary supplements. The everyday experience of influencing mood by ingesting something often fails to appear as a significant difference in a controlled experiment. As will be seen, mood assessment techniques can fail to capture transient phenomenological experience.

Abbreviations: ADACL, Activation–Deactivation Adjective Check-List; POMS, Profile of Mood States; VAS, Visual Analogue Scales.

*Corresponding author: Professor Richard Hammersley, email r.hammersley@hull.ac.uk
Here, ‘phenomenological’ is used after Merleau-Ponty\(^{(5)}\) to mean the content and form of human experience, without making a priori assumptions about the causes or structure of that experience. This term is used rather than ‘subjective’ because the latter has come to be the antonym of ‘objective’, which is misleading here because the ultimate aim of mood measurement systems is objectively to measure phenomenological experience. Additionally, as will become clear, there is a need for a term to describe subjective experience without implying the assumption that such phenomenology involves full conscious awareness, which can easily be reported.

As shall be reviewed, reward, or reinforcement, can occur without mood, and mood reports are influenced by many things other than the current or immediately prior phenomenological experience. It is widely recognised that experienced mood is not caused only by underlying changes in physiological state but also by a range of other factors, and the connections between experienced mood and rated mood are cognitively mediated. However, some nutrition research over-simplifies these issues and treats rated mood as if it were an uncomplicated approximation of physiological state\(^{(4)}\).

This applies across different research topics including: the extent to which the obesity epidemic has been caused by abundant sweet, high-fat, and otherwise highly reinforcing food, and which reinforcement mechanisms apply; whether macronutrients have specific impact on mood; whether functional foods can be designed to enhance mood or performance; whether abnormalities of food choice and eating behaviour are caused or worsened by the foods chosen. Mood has been measured in nutrition research for over 30 years, yet there has been little progress in agreeing standards regarding the choice and use of appropriate measurement instruments\(^{(1)}\).

### Defining mood

Affect science distinguishes ‘mood’ from ‘emotion’. Emotions are strong affective responses that usually have visible behavioural effects, such as changes in facial expression\(^{(5)}\) and are fundamentally communicative acts\(^{(4)}\). One way of defining emotion is according to facial expression – happiness, sadness, surprise, fear, disgust, anger\(^{(6)}\) – although this may not cover all emotions. Moods are weaker phenomenological experiences that may not have behavioural effects\(^{(2)}\).

It is also important to distinguish ‘protracted mood’ over a period of hours or days, which more readily can be assessed by questionnaire\(^{(1,2)}\), from ‘transient mood’, which fluctuates\(^{(4)}\). Most research with ‘mood’ in the title is about protracted, usually depressed, mood. Reports and ratings of protracted mood draw upon and somehow average information from episodic memory to generate the phenomenon of a relatively stable state from the underlying momentary flux of feelings. For example, depressed individuals exhibit specific memory and attention biases towards reporting negative events and thoughts\(^{(7)}\), although in reality they also experience positive events and thoughts. Thus, although depression is probably caused by a complex of metabolic, neurological and cognitive factors\(^{(8,9)}\) and has behavioural effects as well as phenomenological ones, depression as a protracted mood is cognitively formed and involves selective attention to different aspects of experience, as for instance in rumination\(^{(10)}\).

Transient mood, being much shorter, is inherently variable. Additionally, the determinants of transient mood may be unconscious, or only briefly accessible to conscious and rapidly forgotten\(^{(11)}\). A transient mood does not dominate consciousness, for if it does then either it becomes strong enough to manifest in behaviour as an emotion, or it becomes a protracted mood, which permeates consciousness and/or behaviour for some time.

It may be inaccurate to conceptualise transient mood as a ‘state’\(^{(4)}\). By definition, transient mood is relatively brief and prone to change; also ‘state’ implies something that dominates consciousness and has content that can be introspected. Some affect research uses ‘valence’ to mean a tendency to orient towards some types of stimuli rather than others, or respond in certain ways, as in the attentional biases of depression\(^{(7)}\). In animal models ‘value’ is sometimes used in a similar way\(^{(12)}\).

As will be reviewed, transient mood effects may not always involve an overt orientation of this kind. Hammersley & Reid\(^{(4)}\) suggest the use of ‘tint’. A transient mood can tint consciousness, but awareness of the tint itself varies, and the tint subtly emphasises some aspects of phenomenological experience and de-emphasises others. Valence and value are two examples of the effects of a mood tint, but there may be others, such as cognitive biases of various kinds.

Research on the effects of single acts of ingestion usually hypothesises changes in tint, rather than effects on protracted mood, or emotion. For example, high-carbohydrate breakfasts might tint consciousness with happiness\(^{(13)}\), without causing the happy facial expressions commonly shown in advertising.

Transient mood is caused by cognitive processes, along with physiological ones\(^{(4)}\). Asking individuals to evaluate, rate or report their mood tends to orient individuals to the affective tinting of their conscious state and also to change the contents of consciousness, filling it with the relevant mood assessment task. At the extreme, individuals may be unaware of their transient mood unless they are asked about it, because transient mood can exist without full awareness\(^{(4)}\). Affective priming involves both fully conscious and less conscious processes\(^{(14)}\). One example is irritable mood, when an individual may feel normal and show no overt signs of irritability, but nonetheless have more propensity to exhibit irritation to stimuli that might not irritate them normally.
Affective priming is one cause of transient mood, where prior events and thoughts cause a tendency to be oriented towards cognitions of specific emotional significance. For example, mood-salient words embedded in a task can influence subsequent mood ratings without the subject noticing their systematic presentation. Similarly, foods or drinks with learned associations to specific moods might trigger the mood. Sweet taste consistently causes a transient facial expression of happiness, whereas bitterness or bitter–sweet tastes have less consistent effects. Even strong and immediate affect experiences are cognitively mediated, so it is unsafe to assume that rated mood represents the unmediated physiological effects of ingestion on subjective state. Research participants’ feelings can be influenced by their prior outcome expectancies. It is also necessary to consider the demand characteristics of the experiment, including the nature and demeanour of the experimenter. To minimise experimenter effects, social priming effects, and expectancy effects, ideally all research with individuals should be double-blind whenever feasible.

Most procedures reporting on mood involve retrospective reconstructing of recent feelings, so by the time individuals rate or otherwise report on a small mood change, they may have forgotten or distorted its initial nature. Modern conceptualisations of consciousness, such as the working memory framework and fast \( t \) slow processes, regard it as having multiple components with different functions and properties, as well as indistinct boundaries from unconscious processes, rather than it being a simple state that can be examined and reported upon without problem. Many studies of the effects of ingestion on mood have assumed that rated mood represents conscious state, caused by ingestion. Then, research sometimes ‘back-calculates’ the physiological effects of ingestion from mood ratings, which is unsafe because of the above complexity and the relationships between physiological arousal and conscious state are complex. For example, below will be reviewed the challenge of demonstrating that carbohydrate can improve mood due to elevating tryptophan levels in the brain and enhancing serotonin metabolism. Carbohydrate often apparently affects transient mood, but this mechanism applies only amongst individuals who crave carbohydrates and habitually ingest pure carbohydrate deliberately to improve mood.

However, sweetness produces a strong positive affect experience and a happy facial expression, so it may have direct effects on subjective state. As well as in adult humans, this occurs also in young infants, primates and other animals. Depending on how one defines and distinguishes ‘reward’ from ‘reinforcement’ this may suggest that sweetness is a primary positive reinforcer with the potential to cause a positive mood without cognitive processing. Some tastes also trigger disgust reactions. However, there is no convincing evidence that any of the other properties of food are primary reinforcers. Indeed, some orosensory properties that tend to be associated with energy-dense foods, such as viscosity, are represented independently of their reinforcement value. Research on conditioned taste aversions suggests that near-instantaneous responses to foods are often learned.

Research on the effects of carbohydrates on transient mood

This section reviews experiments that examine the effects of food and drink on transient mood. Excluded are experiments using caffeine and other drugs, studies that assess the protracted mood effects of various diets across several days or longer, and studies where mood is assessed but is not reported as a main focus of the study. Of remaining experiments there have been most on the effects of carbohydrates, including studies of sugar, breakfast cereal and breakfast bars. There are also a few studies of protein without comparing it with carbohydrate, a few on the effects of other meals, and some on the effects of miscellaneous foodstuffs that possibly have mood benefits. This section will focus on carbohydrates, because there are too few of the other studies to be able to judge the consistency of their findings, even before considering the challenges of mood rating.

In studies on simple carbohydrates of under 24 h, effects on mood have been found inconsistently. Not all studies have blinded participants to what they ingest, which may help explain variations in findings. Another difficulty is that many studies use multiple measures of mood, and sometimes multiple questionnaires, making it difficult to interpret changes on a small number of mood items. The questionnaires themselves are not optimised for this purpose (see below).

Other difficulties in this literature include that eating has generic mood effects, releasing endorphins, alleviating hunger and the cognitive effects of fasting, and reducing dehydration. Consequently, showing that a meal of specific composition changes mood compared with fasting cannot be attributed to the nutrient content of the meal. To demonstrate this requires comparison with a control meal of systematically varied composition.

Breakfast compared with fasting improves cognitive performance, often without affecting rated mood. When mood is affected by breakfast, then alertness tends to increase and fatigue to decrease, but, as shall be reviewed, there are difficulties with inconsistent mood measurement techniques. It is also conceivable that breakfast improves cognitive function after an overnight fast, then consequently better cognitive function improves rated mood.

Five papers explicitly compare high-carbohydrate breakfasts with high-fat or high-protein breakfasts of otherwise similar content. For this comparison, there may be issues
regarding the different speeds at which different meals are digested. Two have found no distinct mood effects of different meals using the Visual Analogue Scales (VAS) and the Profile of Mood States (POMS)\(^{(42, 43)}\). Experiments finding specific ‘significant’ mood effects of one meal type compared with another are difficult to interpret, because effects found differ. Carbohydrate may reduce fatigue and dysphoria (VAS)\(^{(53)}\), or increase fatigue (VAS)\(^{(41)}\), or increase alertness and happiness whilst reducing nervousness and thirst (VAS)\(^{(13)}\). It is impossible to tell whether reduced ‘fatigue’ and increased ‘alertness’ are the same thing, or how VAS ‘fatigue’ can be translated into POMS scales or those of another instrument. Nor can one tell whether one instrument is more sensitive than another.

There have also been four experiments looking at giving glucose drinks during vigorous exercise, compared with only water. Such drinks had no mood (all POMS) or performance effects on recreational cycling\(^{(72)}\). Playing soccer, compared with placebo they increased ratings of activation and of perceived exertion\(^{(73)}\). They improved vigilance, reduced confusion and increased vigour with another are difficult to interpret, because effects found differ. Carbohydrate may reduce fatigue and dysphoria (VAS)\(^{(53)}\), or increase fatigue (VAS)\(^{(41)}\), or increase alertness and happiness whilst reducing nervousness and thirst (VAS)\(^{(13)}\). It is impossible to tell whether reduced ‘fatigue’ and increased ‘alertness’ are the same thing, or how VAS ‘fatigue’ can be translated into POMS scales or those of another instrument. Nor can one tell whether one instrument is more sensitive than another.

There have also been four experiments looking at giving glucose drinks during vigorous exercise, compared with only water. Such drinks had no mood (all POMS) or performance effects on recreational cycling\(^{(72)}\). Playing soccer, compared with placebo they increased ratings of activation and of perceived exertion\(^{(73)}\). They improved vigilance, reduced confusion and increased vigour with another are difficult to interpret, because effects found differ. Carbohydrate may reduce fatigue and dysphoria (VAS)\(^{(53)}\), or increase fatigue (VAS)\(^{(41)}\), or increase alertness and happiness whilst reducing nervousness and thirst (VAS)\(^{(13)}\). It is impossible to tell whether reduced ‘fatigue’ and increased ‘alertness’ are the same thing, or how VAS ‘fatigue’ can be translated into POMS scales or those of another instrument. Nor can one tell whether one instrument is more sensitive than another.

Another study relevant to the tryptophan hypothesis\(^{(83)}\) fed high- and low-neuroticism-scoring participants high-carbohydrate or high-protein diets across breakfast, a snack and lunch and gave them a stressful task. High-neuroticism participants who had eaten the high-carbohydrate diet did not show the stress-induced rise in depression, decline in vigour (POMS) and elevated cortisol levels that they showed after the high-protein diet. This suggests that carbohydrate may moderate or protect against a stressed mood response, rather than changing normal mood.

In summary, it appears that carbohydrates can affect mood but it remains difficult to specify the relevant conditions. Nutritional status, stress, and the habitual use of carbohydrates to affect mood may be important, but findings are obscured by some designs that do not blind participants to what they are ingesting, and by lack of a consistent and planned approach to mood measurement. It is possible that the POMS and VAS measure somewhat different things, and their sensitivity to nutritional manipulations may also differ.

### Causes of transient mood rating

Mood rating may only take from a few seconds to a few minutes, depending on how many items are used, but this is ample time for complex unconscious and conscious cognitive processes to affect mood rating\(^{(84)}\). Priming effects may be particularly important, where prior stimuli affect subsequent responses although the individual does not consciously recall the stimuli, and the stimuli may be non-salient and incidental to the response.

#### Semantic priming of mood rating

One cause of transient mood-rating effects may be semantic priming\(^{(85, 86)}\), even in the absence of initial mood changes. Semantic priming is the well-established phenomenon of the priming of meaning by prior exposure to related material associated with it through learning\(^{(31)}\). A well-known case in psycholinguistics is that prior context primes the recognition of one meaning of a polysemous word rather than the other; ‘Time flies like an arrow’ v. ‘Fruit flies like a banana’\(^{(87)}\). In mood research, mood-induction procedures that use verbal stimuli, such as the Velten procedure, may prime the semantics rather than the actual affective experience\(^{(98)}\), leading for example to more choice of sad words because the text contained words priming sadness. Similarly, a product repeatedly advertised as making individuals happy might increase ratings of happiness without affecting happy mood. Semantic priming can affect tasks involving alcohol-relevant words\(^{(85, 89, 90)}\). Cues associated with substance use, including words, can prime both substance-like and substance-opponent responses, although the mechanisms of cue exposure may be different\(^{(91)}\). There is also a literature showing that emotional stimuli can prime
various kinds of task – affective priming – including recognition of related words \(^{(92–94)}\) and drinking beer \(^{(95)}\). Smells and tastes also can prime episodic emotional memories \(^{(96–98)}\), which might cause moods and/or influence mood ratings.

**Mood rating affects subsequent mood rating**

Initial ratings of mood may influence subsequent ratings. First of all, there are the effects of repetition, whereby simply repeating a response without reinforcement makes the response more likely in the future \(^{(99,100)}\). So a rating of ‘tired’ at time 1 may make that rating more likely at time 2. There is also the phenomenon of semantic saturation, whereby repetitive presentation of similar items can eventually inhibit further similar responding \(^{(101)}\). Thus, repetitive use of the same rating scales might lead individuals to shift their ratings, but not in ways that are easily predicted or controlled, as there are at least two opponent processes involved.

**Sensitisation and orientation**

Moreover, it is likely that completing rating scales about current affective state sensitises respondents to affective information \(^{(102)}\), and makes them more likely to experience and/or report small changes in mood. For example, sometimes in everyday life the individual may not feel hunger until asked if they are hungry. Mood rating is unlikely to be a simple ‘readout’ of current state, but is always based on a second-round cognitive appraisal \(^{(4)}\). This appraisal does not necessarily use information from an initial rapid affect experience, or it may contradict that initial experience. Mood-rating systems therefore have the potential for generating ‘mood’ on the basis of little or no information from consciousness. Consequently, ingestion may sometimes systematically influence mood ratings, without affecting prior subjective state.

**Analysis of mood ratings**

Mood ratings repeated over time generate large quantities of data. There is the potential for type 1 error, so data analysis needs to have a plan to minimise the number of statistical comparisons. Such a plan ideally includes having clear hypotheses about how ingestion should influence mood, regarding which components of mood should be affected, at what time, and for how long. Research should still assess mood comprehensively, to avoid item availability bias, but analysis should focus on the components where effects are hypothesised.

It is not always possible to form clear hypotheses about how ingestion should influence mood, particularly as mood is often assessed as an adjunct to other research questions. In such cases it is important that a minimal number of mood components are analysed, and a conservative approach to analysis is taken that considers the problem of multiple comparisons. Without a considered approach to analysis, there is a risk of the ad hoc identification of isolated ‘significant’ but meaningless differences, on one mood rating at one point in time, or a few ratings at different points in time. This may partially explain inconsistent and not readily replicated findings regarding the effects of ingestion on mood.

In analysis and interpretation, it is important to not lose sight of the fact that mood ratings are ordinal ratings of subjective state, not accurate interval measurements of anything. The absolute rating values are relatively meaningless and analysis should focus instead on change and comparison between groups or conditions and over time. It is advisable to consider individual differences in rating scale use, for example by transforming raw data into z-scores.

**Requirements of a transient mood-rating system**

1. It should cover the theoretically considered main dimensions of mood (arousal and affect).
2. It should also assess common physical feelings such as hunger, thirst, pain and illness, so that changes in these do not manifest as changes in arousal or affect.
3. It should use as few questions as possible, given requirements 1 and 2, to: facilitate repeated administration; reduce test fatigue; reduce non-conscious priming effects, which increase with repetition; reduce other biases; minimise the duration of the rating process in order to capture ‘mood’ across a relatively brief moment in time; facilitate data analysis.
4. Individual rating items should be structured as polar opposites.
5. The granularity of each scale should be fit for purpose, usually either an unnumbered line, or a five- or seven-point scale.
6. The instrument should be straightforward and have face validity because mood is inherently subjective. What are the implications of ‘subtle’ mood effects averaged across multiple items? If someone is subjectively experiencing a mood they should be able to report it with an appropriate instrument administered in a timely fashion.
7. Assessing mood too often may radically alter the natural phenomenological condition. For example, fifty or more questions rated repeatedly across a period of hours may dominate consciousness more than any effects of the initial food. Alternative approaches include making more use of between-subjects designs with mood measured but once in each group, and ecological momentary assessment methods \(^{(103)}\) where participants do not know when they will be asked to rate mood.

The putative mood effects of substances are usually only one of a range of things to be assessed when measuring...
the short- or long-term effects of what individuals ingest. Many studies have relied on either very basic rating scales, or one or more standard rating scales discussed below. In analysis and interpretation of data, mood-rating data are generally treated as unproblematic measures of internal subjective state. Indeed, abstracts often fail to report how mood was rated, as if for example ‘fatigue’ was the same in every rating system.

Are widely used rating systems fit for purpose?

Table 1 summarises the most commonly used rating systems. The Visual Analogue Mood Scale (VAMS) is too simple, with just one rating item of best to worst mood, and it has not been properly validated for use with transient mood, so it is not recommended.

Most popular in nutrition research is the POMS, which was originally developed to assess how individuals ‘have been feeling recently, including today’, so it was validated to measure protracted mood. It originally asked participants to rate seventy-two mood adjectives on a 1–5 scale of ‘not at all’ to ‘extremely’, generating the six scales shown in Table 1, and it has developed over the years. Now updated to the POMS-2, with sixty-five items, or sixty items in the youth form, or thirty-five items in the short form, giving six scales plus ‘friendliness’ (which is rated separately), labelled somewhat differently to POMS-1.

Widespread use does not of itself validate a measure. The POMS never has been properly validated as a measure of transient mood, and it may be too complex for purpose, particularly for repeated administration. Also, it is entirely empirically derived and it is unclear how the six scales map on to the two primary mood dimensions of arousal and affect: arousal might be measured with ‘vigor’ or ‘fatigue’, while bad mood might be measured with ‘anger’, ‘tension’ or ‘depression’. If POMS results in nutrition research were highly consistent, then one might be more assured of its utility.

The VAS comprise eighteen (or sometimes sixteen) rating items producing two, or three, factors. Although VAS factors are similar to the main dimensions of mood, they are empirically derived and the variation in the number of factors raises questions about the stability of the underlying factor structure and whether those particular questionnaire items are the most appropriate ones.

However, there is some evidence that the VAS are sensitive to changes over periods of less than 24 h, especially to the effects of breakfast.

The evidence that the Activation–Deactivation Adjective Check-List (ADACL) has this sensitivity is more limited because it has not been widely used other than by its author. The ADACL was theoretically derived according to a two-dimensional mood structure as is commonly found in affect research. Dimensions are arousal or energy level and hedonic tone (positive or negative mood) and it is also possible to cross-culturally map emotion words on to these dimensions. The fifty-item version of the ADACL is too long for repeated administration, but the twenty-item short version could be used in this way.

In choosing a method, a caution is that the validated VAS are sometimes cited as justification for using other, non-validated, sets of visual analogue scales, as if it were the scale design rather than the content that had been shown to measure mood. English includes hundreds of mood-relevant words, and the choice of which to use needs to be theoretically or empirically derived. There is a risk of arbitrary items being chosen.

Not commonly used in nutrition research, but worth considering, are the ‘Visual Analogue Mood Scales’, which are widely used in medical research, require minimal cognitive or verbal skills, and measure eight mood dimensions: afraid, confused, sad, angry, energetic, tired, happy and tense. As with the POMS, these dimensions are not theoretically derived and do not appear to have been shown to be sensitive to changes over less than 24 h. Also worth considering are the Positive and

| Table 1. The mood-rating systems most commonly used in ingestion research |
|-------------------|-------------------|-----------------|-------------------|
| **VAMS** | **POMS** | **VAS** | **ADACL** |
| **Advantages** | Simple | Popular | Validated |
| **Disadvantages** | Too simple | Too long, not validated | Are factors appropriate? |
| **Dimensions** | One item: best to worst | Six (POMS-2): anger/hostility; confusion/bewilderment; depression/dejection; fatigue/inertia; tension/anxiety; vigor/activity; plus friendliness | Two or three: arousal; affect; calmness |
| **Number of items** | 1 | 72/65/35 | 18 |
| **Type of rating** | Bipolar analogue | Unipolar 1–5 rating | Bipolar analogue |
| **Content validity?** | No | For psychiatric evaluation | Unipolar 0–4 rating |
| **Construct validity?** | Not applicable | Yes | Yes |
| **Sensitive to normal mood changes over <24 h?** | Maybe | Not established | Not established |

VAMS: Visual Analogue Mood Scale; POMS: Profile of Mood States; VAS: Visual Analogue Scales; ADACL: Activation–Deactivation Adjective Check-List.
Negative Affect Scales\(^{123,124}\); here the dimensions are negative arousal – annoyed/anxious \(v\). contented/serene – and positive arousal – euphoric/elated \(v\). tired/bored. This maps the same space as arousal/hedonic tone, emphasising different axes. Another method is the Geneva Emotion Wheel, which maps emotion words on a circle with axes of hedonic tone (called emotional valence) and perceived low to high control, which is hypothesised to influence mood\(^{117}\).

A final consideration is how many dimensions of mood to measure. Two-dimensional models of mood are usefully simplifying, but using a two-dimensional scale begs the question of whether food might affect components of mood other than arousal or hedonic tone. Moreover, often in nutrition research phenomenological state is not assessed just with a single mood questionnaire, but also with questions regarding subjective physical well-being, and sometimes also with multiple mood questionnaires, including many others in wide use in health-related research not shown to be sensitive to changes over less than 24 h. As discussed above, these different ratings can potentially interact in ways that are not yet understood, but with multiple questionnaires to complete, it is not unusual for designs to require participants to complete several hundred questions in total, leading to test fatigue.

### Inherent biases in the rating process

Whichever rating system is chosen, there are multiple, interacting sources of bias in the use of rating systems to judge mood or anything else. The following are some of the most important and widely recognised biases\(^{1,84,125}\): set-point biases; biases due to the granularity of the scale; biases due to the labelling of the scale; biases due to the items available for rating.

#### Set-point biases

What you initially report limits what you can report next. For example, if you initially rate your hunger 6/7 and then get more hungry, you can only increase your rating by 1 point. If you had initially rated yourself 2/7 you would have more room for rated hunger to increase. Unipolar systems, such as the POMS and ADACL may be particularly problematic.

#### Biases due to the granularity of the scale

When the scale has a lot of points (100-point scales are sometimes used) then participants tend to not use the full scale. When the scale has few points then this may force participant responses into categories. This is a particular problem with even numbers of options, such as the four in ADACL. Recommended solutions are: (a) to use a blank line with clearly labelled end points and ask participants to mark the line according to their mood rating (for example, Visual Analogue Mood Scale). This produces the best rating data, but it can lead to incomplete and uninterpretable responses for less literate, less intelligent participants, who are less familiar with concepts of graphical representation. Consequently, (b) use a seven- or five-point scale, with the points numbered and clearly labelled end points. This produces adequate ratings, and with fewer missing data.

#### Biases due to the labelling of the scale

It is particularly important that the end points are clearly labelled and, for bipolar scales, are convincing opposites. This is problematic for mood ratings because not all possible mood words have unique, convincing opposites. For example, is the opposite of aroused calm, or tired?

#### Biases due to the items available for rating

At one extreme, it is possible to rate mood with a single rating of best to worst mood\(^{103}\), but this forces participants to represent any change in subjective state on this scale; worst mood could be due to many factors such as boredom, depression, low arousal, hypoglycaemia or indigestion. At the other extreme, mood can be rated with as many as seventy-two questions, which requires nuances of judgement that may be unrealistic; can one be more ‘drowsy’ than ‘tired’, and can one simultaneously be ‘alert’?

Ideally, a mood questionnaire should cover all relevant aspects of subjective state. For nutrition research this involves measuring both the main dimensions of transient mood, and common dimensions of subjective physical condition: hunger, thirst, illness and intoxication.

### Towards a solution

To address some of these issues, we have developed a theoretically derived ten-item questionnaire. Bipolar items consist of the two main dimensions of arousal (tired/energetic; restless/relaxed), the main emotions (happy/sad; angry/calm; anxious/composed; disgusted/satisfied), plus items reporting the phenomenology of physical condition (hungry/full; thirsty/not thirsty; intoxicated/sober; ill/well). Pain could be added in relevant studies. The questionnaire is readily usable over periods of hours\(^{44–47}\) and days\(^{126–128}\). Ratings vary across the circadian cycle in consistent and expected ways. For example, energy is higher in the morning than the evening. The questionnaire also has content validity and can be completed quickly without tedium, for example as part of a food and activity diary\(^{126–128}\). Unlike the questionnaires derived by factor analysis, this system is not supposed to have a coherent factor structure and cannot be reduced to two factors, because it was designed to measure the minimum number of orthogonal ratings in the same number of questions. It has not yet been validated against...
established instruments, or by comparing mood-disordered groups with control participants, or by the use of mood-induction procedures.

Conclusions

Despite its importance as a reinforcer of ingestion, mood is hard to define and is inherently subjective. Consequently it is not feasible to produce a definitive procedure for assessing it. However, it is important to distinguish mood from emotion, and in nutrition research it is important to be aware of the distinction between protracted and transient mood, because the problems of assessing the two are quite different, and because eating and drinking are more likely to affect transient mood. Most mood research focuses on protracted mood and some of the mood-rating systems were developed for protracted mood assessment, so may not be as suitable for transient mood assessment. Moreover, unconscious affective and semantic priming can affect both mood and responses on mood-rating systems, making it possible that rated mood bears little relationship to recent phenomenological experience, and that phenomenological mood is not the sole or dominant cause of rated mood, which is also determined by other cognitive factors. Consequently, mood changes may be more important as reasons for food choices than they are as reinforcers of food choices. Individuals may commonly believe that they make choices to affect mood, but they have less awareness of making them on the basis of other pathways.

Transient mood is difficult to measure, so assessment needs to be theoretically considered, brief and administrable quickly, comprehensive, usable, fit for repetitive administration, and administrable under conditions where theoretically irrelevant cognitive factors that can influence mood rating have been considered and controlled. The assessment method should also be validated for use in this way, and not just previously have been used in this way. Despite its popularity, there are reasons to be cautious about the appropriateness of the POMS and other questionnaires derived entirely by factor analysis of large item pools, and there are a variety of other instruments that are worth considering. However, using multiple questionnaires is not recommended.

The minimum standard is to provide a clear rationale for the choice of instrument, based on the requirements of the research, to have a theoretically informed approach to the analysis of mood data, and to be cautious about interpreting ‘significant’ mood differences. It is hoped that these principles will enhance the sophistication of future work using mood assessment in human nutrition research, and there is also a need for further research addressing the question of how and when ingestion affects transient mood, that could inform the choice, development and use of appropriate measurement techniques.

Acknowledgements

We would like to thank the Biotechnology and Biosciences Research Council, The Alcohol Education and Research Council, The Nuffield Foundation, Masterfoods, and Sugar Nutrition UK for their support of our research over the years.

R. H. and M. R. jointly developed and wrote the theoretical ideas in this paper with R. H. taking the lead on its final form. S. L. A. provided further review, editing and critique to ensure that its psychology content was presented appropriately for a nutrition science audience.

There are no conflicts of interest regarding this paper.

References

Measuring mood in nutrition research


