

Author's manuscript

© David A. Booth, University of Sussex, Falmer, Brighton (UK), 12 August 2015

Paper accepted for publication in the *Journal of Sensory Studies* on 18 September 2015

Running title: PREFERENCE VERSUS PLEASURE

“I LIKE IT!” PREFERENCE ACTIONS SEPARATED FROM HEDONIC REACTIONS

DAVID A. BOOTH¹

Food Quality Research Group, School of Psychology, College of Life and Environmental Sciences,
University of Birmingham, Edgbaston, Birmingham B15 2TT, U.K.

¹Corresponding author.

EMAIL: D.A.Booth *at* Bham.ac.UK *OR* D.A. Booth *at* Sussex.ac.uk

Acknowledgements

Thanks are due to Jennifer Schneider and Claire Wainwright for running the experiment and to Suzanne Higgs for agreeing to this use of the laboratory facilities under her direction.

ABSTRACT

In 1952-7, Peryam and colleagues developed nine ordinal phrases of liking and dislike to assess consumers' dispositions to accept or reject a food or drink. They named their questionnaire a Food Preference Scale. Others called it the Hedonic Scale, which means assessment of pleasure, not choice. It is still widely assumed that the word "like" distinguishes felt pleasure from observed wanting to consume the sample. The present quantitative results complement an earlier qualitative finding that preference scores do not provide evidence of the experiencing of pleasure. Rather, "I like it!" simply indicates high acceptance of the sampled variant of a product. Nevertheless, in this experiment, some assessors did also get a convulsive thrill from oral stimulation, as distinct from just enjoying the mouthful, or being pleased by it. However this sensual pleasure came only from strongly disliked levels of stimulation and is probably unique to samples sensed as intensely sweet.

PRACTICAL APPLICATIONS

This experiment's separation of preference from pleasure depended on overcoming practitioners' division between sensory vocabulary and preference scores. Instead of seeking statistical patterns that bridge the supposed gap between sensory concepts and acts of acceptance, sensory studies should design test samples capable of measuring the impact of specified variations in the product range, first on a fully integrative judgment such as match to the personal ideal, or to the most familiar or usual brand. Second, if analytical characterization might help to test the specification, samples can be rated on vocabulary learned in life or in the laboratory, with one anchor on the standard to be matched, such "exactly as I like it" or "just right" (not "like extremely" or "just about right"), and only one other anchor, such as "neither like nor dislike" or "just too wrong to be tolerable." Existing data collection and analysis software are easily adapted this way. 151

Keywords

Sensory Impact. Sensual Pleasure. Product Preference. Sensory Characterization. Psychological Scales.

INTRODUCTION

This paper reports the quantitative results from the first experiment to separate preference from pleasure (Booth and others 2010). The qualitative findings previously reported showed that it is incorrect to use the term ‘hedonic’ for scores of liking, pleasantness, relative acceptance or disposition to choose. The findings presented in this paper were that amount of pleasure does not relate monotonically to strength of preference.

The whole experiment relied on an extension of previous demonstrations of the rapid estimation of each assessor’s ideal point and range of tolerance for a sensed characteristic of a familiar product in a regular usage situation. Such measurement could be central to sensory studies but has so far been neglected. Hence the introduction to this paper takes the space to specify the practicalities from basic scientific principles.

The experimental session had two parts. First, each assessor’s ideal point and upper tolerance limit were estimated using four samples tailored to the individual’s range of acceptance. The second part of the session measured emotional reactions to three samples that varied among subsets of assessors. For each subset, one sample was acceptable to everyone, one was close beyond the tolerance limit of each of those assessors, and the final sample was far into the region of intolerance, to extents that were varied among subsets. The data from this part of the session, reported numerically here, achieved the theoretically innovative result of separating the verbal action of preferring from the expressed reaction of experiencing a pleasure.

Observed Preference or Experienced Pleasure?

It is widely assumed in sensory studies and other research areas that the terms ‘like’ and ‘dislike’ measure the subjective experience of pleasure or displeasure from some conscious sensation generated by the test sample. The commonest research use of these terms is to obtain scores on nine verbal categories from “like extremely” to “dislike extremely.” This layout was designed to provide grades of preference behavior in the laboratory that were valid on choices in the field (Peryam and Haynes 1957). Despite that intention, Peryam and Pilgrim (1957) chose to use the word ‘hedonic’ in the title of their paper as well as the word

‘preferences.’ The reason given in an earlier paper was that the words in that layout were already “commonly called a ‘hedonic’ scale” (Peryam and Giradot 1952).

The word for pleasure in ancient Greece was *hēdonē*. The adjective 'hedonic' and its derivatives became widely used in English in the 19th century for public discussion about bodily pleasures among respectable people. A recent review of research using response layouts containing the word ‘like’ (Lim 2011) pointed out that Peryam’s intentions were overwhelmed by the rapid uptake of the nine ordinal categories but did not acknowledge his reservations about regarding preferences among foods as hedonic reactions. Peryam’s concession to popular jargon has done considerable harm over the subsequent decades (Booth and others 2010).

That confusion of preference with pleasure is not confined to liking. The same strictures apply to the assumption that ratings of pleasantness and unpleasantness assess pleasure and displeasure. “It’s pleasant” is indeed synonymous with “I like it.” This equivalence supports the behavioral interpretation of ‘like’ and ‘dislike’. The words ‘pleasant’ and ‘unpleasant’ are simply latinized forms of ‘pleasing’ and ‘displeasing’. These gerunds imply enthusiastic engagement with an activity that involves the object being assessed -- for example, being “pleased” to eat or drink the item (or being content to buy the food or fragrance). This is not a mere point about words. The verbiage is as it is because it handles reality effectively. When a culturally recognized and personally familiar sort of food item is taken into the mouth, smelt orthonasally, seen directly or in a picture, or merely named, it motivates or de-motivates its purchase and use: the item is accepted or rejected with more or less decisiveness. Whether or not the item also evokes affective experiences such as bodily pleasure or displeasure (or even pain) is an entirely different question, and one that is much more difficult to answer than has generally been acknowledged (Booth 1991, Booth and others 2010; contrast Cabanac 1971, Steiner and others 2001, and many others). Investigators need to stop assuming that scores for liking or pleasantness reflect the amount of pleasure being experienced, until they get evidence in each study from the participants.

Failure to distinguish objective motivation to eat from subjective thrills of pleasure has vitiated the extension to human subjects of a major finding in animal research. Movements of the tongue and jaw elicited by sweetness have been pharmacologically dissociated from the same movements continued into swallowing, which are elicited by all food and drink

(Berridge 1996). Those pre-ingestive movements were dubbed “hedonic.” The unwarranted supposition that ratings of liking or pleasantness show the strength of hedonic reaction (pleasure) has undermined the subsequent human research (e.g., Finlayson and others 2007, Hofmann and others 2010). Higher rated pleasantness merely provides further evidence of wanting to eat or drink the food or drink, i.e. greater motivation to consume (and swallow) the item.

To sum up, preference and aversion are objective dispositions or tendencies to select or to reject the perceived entity. They are the observed conscious or unconscious motivation within public physical or verbal actions, not the private subjectivity of the awareness of an emotional reaction such as a bodily pleasure or physical displeasure or pain.

Preference for What?

How much “I like it” (or dislike the item) can only tell us what actually is attractive (or aversive) to that assessor when those response numbers are shown to vary with the levels of one or more observed features of the tested items. That is, we need to select or manipulate different levels of putative influences so that they vary independently of (orthogonally to) other potential factors, whether known (measured) or unknown (and assumed to be small and/or to vary randomly). If no influence is monitored and implicated in the preferring, then the question arises what scientific meaning or practical use could be validly derived from the bare scores or ranks of liking and dislike.

The degree of preference is not the freestanding number scored, nor need the quantitative judgment have been made on what the investigators intended or assumed to be more or less liked about the item. Preferring is greater acceptance in action or words of the judged object for the sake of those features in which it is perceived by the assessor to differ from the other option(s). Hence a real preference cannot be measured merely by a ruler-like array of words, numbers and points on its own (a ‘scale’ in the jargon sense of that word: Booth, 2009a,b; Lim, 2011). A preference is a position on an intersecting set of mental scales of observed influences on degrees of acceptability. There is evidence of a sensory preference if and only if observed levels of sensed material characteristics of the samples are shown to have affected acceptance on that occasion.

That is to say, the notion of the strength of a preference is empirically and theoretically empty if there is no specification of what is preferred in one test sample over another. Scoring the degree of preference without context treats the objects of choice as ciphers. This is as unproductive as treating the chooser as an empty statistic, rather than as an individual with distinct attributes -- including relative susceptibilities to different influences on choice. Collecting bare preference scores generates the notorious 'fallacies of preference' such as non-transferability of rank order between paired comparisons from three or more samples: assessors could easily be attending to different sets of features in the different pairings, especially when presented in different sequences.

In short, a response such as liking or familiarity measures the individual assessor's way of combining sensed characteristics of the tested materials, and indeed also of attributing concepts brought to the test or included in it (Booth and others 2011d). This contrasts sharply with the presumption, despite a variety of evidence long to the contrary, that sensory vocabulary analyzes out stimulating factors that combine additively.

The experiment reported in this paper varied a single sensory characteristic in a familiar context and measured the levels of either preference or pleasure that it evoked. No sensory vocabulary was made explicit but the samples varied in concentration of a single compound, which was an intense sweetener. Hence the experiment was a particular test of the generic hypothesis that, when a user of a diet cola or a candy, for example, says "I like it!" or expresses a preference in some other way, s/he also has a conscious experience of a pleasurable sensation, such as a physical thrill at the sweetness of the drink, or of the sugar in the chocolate, or whatever.

This test has a wider significance. The popular but dubious hedonic theory of reinforcement postulates that all responses are strengthened by pleasure. On such a basis, it has been claimed that euphoria creates drug addiction and sweetness induces food addiction. If there is no pleasurable thrill from the usual levels of sweetness in foods and drinks, then neither sugars nor sweeteners can be blamed for excessive consumption on the assumption that reward has a hedonic mechanism.

Psychological Scaling of Preference and Pleasure

Pleasure cannot be distinguished from preference without accurate measurement of degrees of acceptance of one or more sensed characteristics of a familiar product. Examples of such accuracy have been widely reported for a long time (e.g., Booth and others 1983, Conner and others 1986, 1988a, Booth 1988b, Booth and Conner 1991, Conner and Booth 1992, Freeman and others 1993, Booth and others 2003a,b). Yet the approach is still not well understood in applied or basic sensory research. Hence this Introduction to the reported experiment has been extended to provide a full explanation of the scientific principles involved in these rapid, precise and therefore economic procedures.

The measurement of a motive, a percept, or any other aspect of mental performance, requires the judged situation to be placed at a point on a causal relationship between that quantitative response and the amounts of the various sources of stimulation that are in fact exerting some influence. Assessors use the distance between a pair of verbal anchors to assess each stimulus selected by the experimenter. Statistical manipulation of the response data on their own, e.g., Thurstone scaling (Ashby and Ennis 2002), 'internal' preference mapping (Jaeger and others 2000), and correlations with ordinal numbering or response positioning on other verbal categories (e.g., 'external' preference mapping: Jaeger and others 2000), can show only the existence of an underlying factor in the responding. It still remains to be ascertained what sorts of stimuli are influencing the mapped responses and how the stimulation-induced processes interact with each other and with the sensed and attributed context (Booth and Blair 1989, Booth and Freeman 1993, Booth and others 2011d). Even the general character of the latent variable is indeterminate without some evidence of specificity from external validation, e.g. whether that response is intentional (motivated, such as a choice) or involuntary (emotional, such as a thrill of pleasure).

In contrast, the present experiment met the requirements for use of a minimum number of data to estimate each panelist's strengths of preference and pleasure at varied levels of sweetener in what they regarded as a high quality drink. There follows a brief overview of the specific conditions for this verbal measurement of both the objective act of preferring and also the subjective reaction of pleasure.

Layouts for Responding

A quantitative response can be spatial or numerical. Spatial positions are often collected on a continuous line, usually with one end anchored on a wording for the apparent absence of the stimulus and the other on a term for an unspecified extreme (a layout widely misnamed a “visual analogue scale”). Such a line can be structured by unlabeled breaks, in the form of dashes, cups, circles or squares. Real numbers are also a continuum, on which the small integers are a highly practiced series of discrete positions. Hence a series of integers can be provided instead of or alongside a row of boxes. If the zero character (0) is assigned to one of two anchors and a finite integer to the other anchor, then the positions of the other integers are at equal ratios as well as equal intervals, at least for integers within the same order such as 0 to 9 (e.g., Bowman and others 2004, Goodchild and others 2008, Booth and others 2011a, Booth 2014). Hence one of the most practical and precise response formats runs from zero (0), though 1, 2, 3 and so on, to a high single-digit integer. That layout was used for all the ratings in the experiment reported in this paper.

Another example of a structured continuum is the 11-point layout of Lawless and others (2010b) with anchors only at the extremes, of the greatest imaginable liking and dislike. The number of points is a secondary consideration, although 0 to 5 may be too few if the psychological distance between the two extreme anchors is large relative to differences between some of the test samples. If the assessor is required to use integers much above 9, on the other hand, there is a risk of ratio biases arising from the switch between the orders of single digits and double digits (as also for moving from ‘percentages’ to three-figure numbers). These issues can only be addressed by comparisons among numbers of unlabeled points between and beyond the same pair of anchor phrases. Statistical differences are uninformative when ratings differ in ways in addition to number of points or an unbroken line.

Two Anchors

There is a fundamental scientific difficulty with the use of more than two verbal anchors on a layout for responses. Each adjacent pair of anchors specifies a psychological scale with a slope and intercept in its linear transform which may be considerably different from those of another pair of anchors. Hence, when the assessor places a judgment about a test sample on a

format with three or more anchors, the question arises of which scale is being used. Is the scaling of responses onto stimuli always based on the two nearest anchors, ignoring those further away? Alternatively, does the scaling use only the outermost anchors? Instead, are there compromise strategies? Do the strategies vary among sets of samples? Is there a fixed generic use of the multiple anchors or are the distances between adjacent anchors determinate only for a particular set of samples and even for each individual panelist? The present experiment avoided all such difficulties by using only two anchors.

The first direct psychological scaling of sensory preferences was characterized and so it was natural for the rating line to use three quantitative anchor phrases, “too little,” “just right” and “too much” (Booth and others 1983, Conner and others 1986, 1987). Hence the psychophysical functions on either side of the middle anchor had to be tested for differences in slope and intercept. Were “too little” and “too much” used by the assessor in ways that placed those two levels of the sensed characteristic at the same psychological distance on either side of the level that was “just right” (i.e., the most preferred concentration in that context: the ideal point)? To avoid distortions of the scaling and hence obtain unbiased estimates of a consumer's ideal point and range of tolerance (from insufficiency to excess), it is also crucial to present samples that spread ratings evenly over the anchored range, but not close to an extreme anchor, and also with the first rating close to ideal (Poulton 1989). Data collected in accord with that bias-minimizing design showed no reliable difference between slopes below and above ideal (Conner and Booth 1992). This principle of co-linearity between pairs of anchors applies whether the degrees of preference are derived from categories (cp. Lawless and others 2010c) or from ratings, i.e. the positions on the line that are specified by each pair of adjacent anchors (cp. Ashby and Ennis 2002, Cordonner and Delwiche 2008).

Contrary to those principles, as with many other quantitative assessments using multiple anchors, the ordinal categories from “like extremely” through “neither like nor dislike” to “dislike extremely” are often treated as equally spaced. This is despite the fact that one of the first published uses of Thurstone scaling of responses, by L. L. Thurstone himself, showed that the nine Peryam/Pilgrim categories were unevenly spaced (Jones and others 1955). Numerical rating without verbal anchors (Stevens 1957, but see Ellermeier and Faulhammer 2000) has repeatedly confirmed those disparities in psychological distance between adjacent phrases (e.g., Moskowitz and Sidel 1971, Cardello and Schutz 2004).

In a paradoxical reversal of the doctrines of magnitude estimation (Stevens 1957), numerical ratings from limited sets of data have been averaged in order to position the anchors on a response line with uneven spacings (cf. Cardello and others 2008, and their citations). That procedure does not address the issue of how the psychological scale specified by each adjacent pair of anchors is used by each assessor of a particular set of samples. Hiding the scores or some of the phrases is no solution either, because assessors may interpret words and numbers between the extreme categories quite differently (Nicolas and others 2010).

Such problems can be avoided by having only unlabeled points between (and beyond) a single pair of anchor phrases. That was the procedure used in the experiment reported here.

Peaked Psychophysical Functions

The plot of a wide enough range of levels of a sensory influence on an individual's strength of preference makes it evident immediately that the function is peaked (Booth 1987). As well as preference being reduced by too little of a good thing, too much of that factor also lowers preference. The common practice of averaging or regressing raw preference scores or ranks across a panel of assessors is therefore a bad mistake. It artifactually flattens preference profiles and response surface umbrellas, in contrast to plots of counts of individuals' peak preferences against levels of the sensed characteristics, e.g. the tasted salt or sugar contents of a food or drink (Booth and others 1983, Conner and Booth 1988; Conner and others 1988b, Freeman and others 1993) or its texture-generating physical properties (Booth and others 2003a,b). Furthermore, the location of a peak of a whole panel's average preference score against a physicochemical measure or a sensory panel score can be arbitrary. For example, if there are two or more divergent regions of popularity, there may be few individual assessors near that peak of average scores, and optimization to that criterion will generate a 'lowest common denominator', highly attractive to no-one. Such criticisms of averaging (or regressing across) bare preference responses have been made many times but the procedure still dominates both research publications and confidential practice.

In contrast, a plot of equal ratios of a sensed material characteristic against an individual's degree of preference in an otherwise ideal context gives a very sharp peak. Theoretically, in

otherwise perfect test samples, the function is an isosceles triangle (Booth & Conner, 1991; Booth and others 1983). These facts about the mental processing that decides the degree of preference for a food sample provide the basis for very rapid determination of the personal ideal point, together with the regions of rejection for intolerable insufficiency or excess of the constituent (Conner and others 1986, 1988a,b). In the present experiment, such minimal data were crucial to the initial identification of a test sample that was close to ideal for the individual assessor, and also of samples close to and far beyond the excessive level that was just rejected by that person. All these values were derived from levels of stimuli and responses for only four samples that presented the varied sensory factor in an otherwise ideal context for each panelist.

Preference versus Aversion

The most important deficiency of the nine phrases of Peryam and Pilgrim (1957) for theoretical and practical purposes is their bipolar disorder, from extreme liking to extreme dislike. Rating from “extremely pleasant” to “extremely unpleasant” has the same limitations. A measure of food preferences should not be concerned with food aversions as well. What is liked in foods having their usual sensed characteristics and conceptualized attributes is often quite different from what is disliked, such as a tainting compound (Mottram 1998) or an named identity that provokes a fear of being made ill (Knibb and others 2001).

In addition, there is a logical asymmetry between “like extremely” and “dislike extremely.” The psychological scale for a preferred constituent is peaked, whereas that for an aversive constituent comes at most to an asymptote at receptor saturation, without ever declining as preference does at excessive levels. Hence, among ordinary foods, preference extends to an absolute maximum, when the sample is ideal or “just right.” The word ‘ideal’ means that no sample could be liked more. That category provides a precise anchor for the preference responses, whereas “extremely liked” does not constrain the panelist between something that is strongly preferred but not ideal and the self-contradictory concept that the sample is more preferred than the ideal version.

Dislike, on the other hand, has no clear maximum in the usual ranges of aversive stimuli. Hence no familiar standard level can be invoked. The phrase “extremely disliked” leaves each panelist free to pick her or his own criterion of extreme. Even “as unpleasant as

imaginable” is vague: it is ambiguous between being as unpleasant as such a sample can be imagined to be (e.g., as bitter as at the saturation of the receptors) and as unpleasant as anything could be (e.g., scaled onto unbearable physical pain).

In addition, the scoring of liking from 6 to 9 in the same set of data as dislike is scored from 4 to 1 can have dire consequences. Averaging such scores across groups has blinded investigators to the difference between an increase in preference and a decrease in aversion. For example, plain tomato soup of modest quality could be somewhat disliked by some people. Adding a little herbal flavoring might make it less intolerable to them. Averaged in with a larger number of people who quite liked the plain version and did not find that the herb flavor improved its quality, this could give a group increase from indifferent to slightly liked. It would not follow that the flavoring increased palatability in any individual, let alone as a generalization across everyone.

For the above reasons, preference and aversion should be assessed separately. One of these two assessments could often be redundant: everywhere in the range of substantial preference is liable to go with zero aversion, and *vice versa*. The present experiment exploited the linearity of preference scales within the tolerable range above ideal and so only degrees of preference needed to be measured, with intolerably high levels of sweetening being estimated by extrapolation of the theoretical straight line.

The panelist’s degree of preference for each sample was expressed in a response placed on a position between ideal (“always choose,” scored as zero) and either just too little or too much at the same position (“never choose,” scored by the investigators as 9 although any number would serve). The stimulus values at points interpolated between of ideal and just rejected are what need to be compared across individuals and panels, not the raw numbers that panelists or investigators assigned to the responses at those particular values. This report provides such a panel-wide profile of sweetener ideals and tolerances in the tested beverage for the market represented.

The present experiment followed the same principle of unipolarity by assessing pleasure and displeasure separately, in identical formats but differing appropriately in the vocabulary of the two anchors. Lack of either polarity of emotional reaction was rated and scored as zero. A degree of either positive or negative affect was assigned a single-digit integer by the panelist

that the investigators used as the score. No word was used that was liable to be ambiguous between pleasure and preference, such as “like” or “pleasant.”

Influences of Sweetener Levels on Preference and Pleasure

A person’s overall liking for a particular food is driven by a distinctive learnt configuration of sensed and symbolic influences (Booth and Freeman 1993, Knibb and others 2001, Moskowitz 2002, Freeman and Booth 2010). This paper reports data from variations in concentration of only an intense sweetener, added to a fruit juice. That manipulation induced differences in the sourness of the juice as well as its sweetness, but there were no variations in texture or osmolarity as with sweetening by a sugar, especially at the high concentrations that might be needed in the present paradigm. Sweetness is no different from any other sensory characteristic or symbolic attribute within the individual’s ideal configuration or norm: the most preferred intensity in a food or drink varies widely among foods and drinks and across people (Conner and others 1986, 1988b).

On the other hand, we and other omnivorous mammals are born with reflexive movements that help mother’s milk down the infant’s throat. As with all reflexes, the response becomes steadily stronger as the stimulus becomes more intense (Thompson and Spencer 1966) -- in this case, “the sweeter the better.” This congenital reflex is suppressed by the learning from infancy onwards of a personal norm for the sweetness of a particular food or drink (together with norm values for all its other perceived features). Nevertheless, if the neural connections of the reflex persist into adulthood, sufficiently strong sweetness could in principle break through the learnt decline in preference as sweetness increases above ideal (Booth 1991, Booth and Shepherd 1988). Clear evidence for such an effect has recently been found: tasters of a popular fruit juice wrote about movements in the mouth that are characteristic of the reflex when the juice was given a revoltingly strong sweetness and may also have been made bitter to some assessors (Booth and others 2010). Furthermore, the sensations of movement were enjoyed more, and overall mood was raised higher, the further above ideal the sweetness was raised (Booth and others 2010). That is, there was a pleasurable reflexive response to the taste of a sweetener at rejected levels. This experimental paradigm had distinguished for the first time a sensual pleasure at sweetness itself from a fully contextualized preference for that taste in a familiar material.

This paper presents previously unreported data from that experiment (Booth and others 2010). The earlier paper related categories of spontaneous wording for felt movements to indicators of preference and pleasure across relevantly selected assessors. Here the correlations between the measures of preference and pleasure are considered across whole panels tested on different levels of unacceptably strong sweetness. Within any one body of data, correlations are statistically independent from frequencies or central tendencies.

Hypotheses

The hypotheses tested in this paper are factual, not methodological. The findings use psychological scales of an influence on a variety of responses in order to advance empirical knowledge. They do not develop procedures for collecting responses. In other words, this paper is ‘psychophysical’: it measures sensitivity to influences. It is not ‘psychometric’: it does not seek multivariate latent factors among diversely worded responses (Booth 1995).

The first part of the experiment reported here tested the hypothesis that there is a linear relationship between the ratio steps of a sweetener above the panelist’s ideal point for a familiar drink (tested during regular thirst). The strength of preference was placed at a point between the anchors of always choosing to have a drink of the tasted sample and the sample being just too sweet for use as a full drink. Confirmation of the hypothesized linearity was relied on to generate estimates of four concentrations of sweetener that were categorized across the panel as generally near ideal, approaching unacceptably too sweet, definitely unacceptable, or a strength that was completely intolerable in the drink.

The second part of the experiment correlated those group-specified levels of preference or aversion with individual panelists’ performance on each of two indirect assessments of pleasurable affect. One index was improvement in overall good mood. The other indicator was enjoyment of sensations of movement in the mouth or face while tasting the test sample. It was hypothesized that there would generally be no linear relation between preference and pleasure, because they are separate phenomena. A correlation between preference and pleasure was envisaged only at such extreme sweetness (possibly with some bitterness too)

that the decline in preference was associated with suppression also of pleasure and/or its replacement with displeasure.

METHOD

Participants

The panelists were 50 young women and men studying at college or university. The experiment was designed to analyze people's performance regardless of background. Hence no assessment was made of the socio-demographic representativeness of this opportunity sample. Each panelist professed to be familiar with the brand of apple juice used and consented to drinking test samples of juice that varied in normal constituents of the diet, including artificial sweeteners.

Test Samples

Each tested sample was a mixture of equal volumes of fresh juice from pressed apples and a re-sweetened and re-acidulated dialyzate of that juice. Bottles of the apple juice (Copello brand) were purchased at a Sainsbury retailer 2-3 days before use and mixed into a single batch before the preparation of samples for a day of tests. Glucose and malic acid were added back to the dialyzed juice to give a sweetness and sourness that could not be distinguished by the investigators from that of the marketed juice. The intense sweetener varied in concentration among samples was saccharin, which has always had regulatory approval in Europe. The samples were presented at room temperature.

Rating Formats

Preference Behavior

In the first part of the experiment, in 16 panelists, used personally bias-minimized sets of test samples to assess behavioral preference for each juice by marking a box in a vertical array of

15 boxes, under the question “How much would you like to drink a whole glass of juice tasting like this (if you wanted a drink of apple juice at the time)?” The top box was labeled “I’d always want to have a glass of this” (ideal point, IP, scored zero). The only other anchor was next to the eleventh box down, “Just not good enough ever to choose.” Below that rejection point (RP), well away from the row of boxes, was an arrow pointing downward, labeled “even worse to make a drink out of.” Beneath the bottommost box was a line on which to write the number of boxes needed if the drink sample was that bad. The personally tailored selection of sweetener levels endeavored not to go beyond the individual’s RP and so that option was never taken. The score for the rejection point and the number of boxes are of no significance, so long as the primary anchor has a zero score (Booth and Freeman 1993) and there are at least four points between anchors (or a continuous line). The use of only two anchors with one of them scored zero makes linear responding possible, with scores having both interval and ratio properties, i.e. full measurement is achieved.

These panelists were pre-adapted to the format without transfer of response habits other than direction, by assessing a sample of the marketed juice for “sweetness,” “sourness” and “overall taste” on horizontal rows of boxes, otherwise identical to the row for rated choice but with the leftmost box labeled “not at all sweet,” “not at all sour” or “as weak as water” respectively and the anchor “exactly like apple juice” always placed at the array’s middle box.

Pleasurable Experience

The second part of the experiment on other panelists used sweetener concentrations that were fixed across a panel for their general effect on preference. Any pleasure or displeasure evoked by each drink sample was measured indirectly in two distinct ways, one generic and one specific to taste compounds.

The generic method assessed the separate effects of each sweetener level on arrays headed “good mood” (positive affect) or “bad mood” (negative affect). A single bipolar format was avoided because the distinction between an intensification of good mood and a weakening of bad mood is difficult to make within the individual, and impossible in grouped data with means near neutral. The panelist was asked to put one dash on each of two adjacent vertical arrays of eight short vertical lines with substantial spaces between them. The top line of an

array (scored 7) was labeled “as good as imaginable” (left-hand array) or “as bad as imaginable” (right-hand array). The lowest line (scored 0) was labeled “no improvement” for positive affect and “no worsening in how I feel” for negative affect.

Those ratings of general positive or negative affect were preceded by the assessment reported previously, of immediate emotional reaction to any movement felt in the mouth of face (Booth and others 2010). The first question read while tasting a sample was “How nice and/or nasty is the muscular reaction in your mouth?” Below were two horizontal rows of the numbers 0 to 7. The first row had “not at all nice” against the zero and “as nice as I could imagine” against the seven. In the second line, 0 was labeled “not nasty at all” and 7 “as nasty as I could imagine.” The instruction was to circle one of the numbers in each line.

After that, the panelist was asked to use her or his own words to describe the nice and/or nasty mouth movements just rated. Those particular data have been reported in full (Booth and others 2010). Only the group means of ratings of good or bad mood and of nice or nasty movement were given in the earlier paper. Here the complementary information in the data is presented, i.e. correlations between the ratings, and of each rating with the levels of sweetener used in the samples. The ratings of mood came last and then the assessor prepared for the next sample.

Assessment Procedures

In the first part of the main experiment, with 19 panelists divided between two sessions, always/never choose ratings only were elicited for five to seven samples of apple juice that were varied in content of saccharin or malic acid in order to control their sweetness relative to their sourness (Conner and others 1986). The main known stimulus biases across a session on ratings were minimized by personal tailoring of the sequence of concentrations presented (Conner and others 1988a,b). Centering bias (Poulton 1989) was avoided by presenting first a sweetness close to marketed, i.e. the lowest step of either saccharin or malic acid (alternately between successive assessors). Range and frequency biases (Conner and others 1987, Parducci 1965, Risky and others 1979) were minimized by presenting higher steps of saccharin or malic acid alternately until a larger gap in rating appeared on one side of ideal, when a sample was selected that was likely to elicit a rating in the gap. End effects (cp. Lawless and others 2010a) were avoided by selecting the second sample on the other side of

marketed sweetness from the first rating by no more than two steps and then selecting the third and subsequent samples to minimize the risk of a rating beyond “never choose” because under-sweet or over-sweet. Furthermore, the phrase “never choose” was not an end-anchor; options for response were provided well beyond it. It should be noted that first-order sequence biases, i.e. effects of one sample on the rating of the next, have no measurable group effect on this sort of data (Conner and Booth 1992).

From these sets of individual data, a concentration of saccharin (0.1 mg per 100 ml; Step 0.1) was identified that was close to ideal sweetness for this juice (“always choose”) in all the panelists. A higher concentration approached just too sweet more or less closely (2 mg %; Step 2) or was slightly beyond it (3 mg %; Step 3). Still higher concentrations of saccharin were not tested in these initial experiments because the data from the ‘too much’ side of the individuals’ preference triangles demonstrated that they must be intolerably sweet in this apple juice for anybody (Booth and Conner 1992).

In the second part of the experiment with 31 other panelists, just four samples of apple juice were presented in a fixed sequence in order to assess any pleasure or displeasure evoked by the sweetener at the different levels of behavioral preference or aversion. First, the level (Step 0.1) was presented that was close to the various concentrations rated “always choose” by the earlier assessors (their ideal points). An intermediate level close to *never choose* (just intolerably too sweet) was presented second (Steps 2 or 3). The third sample contained a concentration of malic acid instead of saccharin that had been rated by the earlier panelists as over halfway from ideal sweetness to being too little sweet ever to choose; this under-sweet (sour) juice was presented only in order to help to balance the oversweet samples and so the responses are not relevant to the hypothesis and the data are not shown. The final sample was intolerably over-sweet, varied among Steps 8, 4 and 3 in panels of 14, 7 and 10 panelists respectively.

Analysis of Data

The choice ratings and saccharin concentrations from each individual in the initial panels were fitted by least squared deviates regression to the vertical conic section (hyperbola) of contextualized preference (Booth and Freeman 1993) using a software tool programmed by Dr Oliver Sharpe to those specifications for biosocial analysis of cognitive processes and

consumer products (Co-Pro 2.29). The tool's output included each assessor's ideal point (IP) for saccharin in the apple juice while thirsty (the concentration that would be rated "always choose") and the unsigned slope of the isosceles triangle of tangents to the hyperbola (see Figure 1 in Results). From that linear equation above the IP, the excess rejection point for each individual was calculated, i.e. the concentration of saccharin that would have been rated "never choose." The slope and mean square error gave the half-discriminated ratio (HDR) of saccharin concentrations (one plus Weber's fraction; the Tolerance Discrimination Ratio of Conner and others 1988a,b, and the JND of Torgerson 1958). One HDR either side of the IP is the individual assessor's ideal range of the sensed constituent in the sensory, somatic and social context of testing (see Figure 2 in Results).

Inferential statistics were calculated in SPSS 16.0. Graphs not from the output of Co-Pro were constructed in SigmaPlot 11.0.

RESULTS

Measurement of Personal Preferences for Sweetener Levels in Apple Juice

The design was remarkably efficient at estimating each assessor's ideal and excess rejection points using only four samples of juice (Figure 1), in line with previous experience of this approach (Booth and others 1983, Conner and others 1986, 1987). In the two panels combined in Figure 1, only one assessor gave the same rating ("always choose") to each of the four saccharin levels (not shown) and one other gave only two different ratings (leftmost panel of the third row, Figure 1). The other 17 assessors gave three or four distinct ratings of degree of preference. Every assessor in the Figure showed the right-hand limb of the hypothetical peaked function of sweetener concentration ratio, as the design was intended to achieve. Half the assessors had their ideal point within the range of the data. The ideal point for each of the others could be estimated by extrapolation.

Figure 1 about here

The profile of ideal ranges was similarly orderly (Figure 2). The visual appearance of skew to low concentrations and a tail of high concentration is belied by the closeness of the mean [setter: these hyphens are minus signs] (-2.19) to the median (-2.18) of ideal points, i.e. about 0.15

mg of saccharin per 100 ml of juice. It should be noted that the appearance of the extremes of a plot of ideal ranges are dominated by the poorest discriminators between “always choose” and “never choose” -- that is, by those who care least about deviations from their ideal sweetness in apple juice. The finest discriminators -- that is, the pickiest selectors among sweetness levels -- were clustered in the range [setter: the next two hyphens are minus signs] -2.25 to -2.00 (Figure 2). In other words, this middle region of the profile also provided the reliable mode of popularity, close to the mean of the ideals.

Figure 2 about here

Since equal ratios of a physical measure are equally well discriminated in a medium range, the saccharin concentrations have to be on a ratio scale (in logarithms) in order to generate a normal distribution of ideal points (Conner and Booth 1992, Conner and others 1988a,b). Hence the miniscule concentrations at the lower ends of the widest ideal ranges (down to a few micrograms of saccharin in 100 ml) merely mean that no saccharin was needed in the juice formulation that was held constant. Indeed, less of the inherent fructose and glucose, or more of the malic acid in apples, would have been tolerated or even preferred by people with an ideal point around -2.5 (lower mode, Figure 2), perhaps corresponding to those with the low ‘sweet tooth’ of fruit eaters and juice drinkers (Conner and Booth 1988, Conner and others 1988b, Freeman and others 1993, Wansink and others 2006).

The saccharin concentrations extrapolated upwards to be at the “never choose” anchor in these 18 assessors (functions in Figure 1) had a median of 1.2 mg % (quartiles 1.1 and 1.4). Therefore in the second part of the experiment, the concentrations of saccharin presented were 0.001 g/l (Step 0.1) as a near maximally preferred level, Step 2 or 3 as near the point of being too sweet (0.02 g/l being likely to be a little less sweet than just “never choose” and 0.03 g/l liable to be definitely too sweet) and Steps 3, 4 or 8 as much too sweet. The saccharin at Step 8 (0.08 g/l) could have tasted bitter to a substantial proportion of the sample, and perhaps to some even at Step 4. It should be noted that this design does not average preferences scores. (No ratings of preference were made by the subpanels reported in the next sections of Results with Figures 3 to 6.) The Step numbers were surrogates for the central tendencies of narrow distributions of positions on the psychological scale of sweetness preference specified by maximum and minimum dispositions to choose among previous volunteers from the same population (Figures 1 and 2).

Measurement of Pleasure and Displeasure

Pleasure and displeasure could be measured indirectly in two ways in the present sets of samples of apple juice with saccharin added. A generic measure was the immediate effect of tasting the sample on mood. A sweet-specific measure was the affective evaluation of impressions of movement in the mouth or face while tasting a sample.

These two measures had considerable cross-validity, especially in positive affect (pleasure). The correlations between the two sorts of rating were highest in the panel tasting very strong saccharin (Step 8), where the shared variance (r^2) reached 74% (top left cell in Table 1 and Figure 3). Good mood was reliably correlated with a nice feeling of movement also in the subpanels presented juice with lower levels of the strongest saccharin (across the row in Table 1 and down the column in Figure 3). In those circumstances, however, there was a tendency for nice movement to occur without an improvement in mood. That is, the pleasurable impressions of muscular reactions to sweetness were sometimes elicited but this reaction was too focused to affect overall mood. Also there were signs of categorical rather than graded behavior with these more moderate ranges of saccharin: mood particularly was rated as either very good (5 or 6 out of 7 for “extremely good”) or as neutral or slightly good (0 to 2).

Table 1 and Figure 3 about here

Similarly with displeasure, these two measures showed the best cross-validity with the most extreme top concentration of saccharin (second row of Table 1 and right-hand column of Figure 3). Overall, the ratings of displeasure tended to be lower than the ratings of pleasure, as to be expected with a high-quality drink and the taste of sweetness. Indeed, when all three concentrations of saccharin were tolerable to many assessors, the great majority of ratings of the felt movement were zero displeasure and worsening of bad mood rated zero or 1 out of 7.

Relationships between Preference/Aversion and Pleasure/Displeasure

Preference and Mood

The second part of the experiment did not repeat the measurements of preference, in the interests of rapid collection of wordings of felt movements in the mouth and face (Booth and others 2010) and then the ratings of affect that are analyzed by correlations in this paper.

Instead, each subpanel of assessors tasted a different top concentration of saccharin. The result was that panels varied greatly in the relationship between predicted preference/aversion and observed expression of pleasure/displeasure.

When a very high concentration of saccharin (Step 8) was included among the three levels (left-most column of data in Table 1), the decline in choice with excess of sweetness represented by that Step number correlated reliably with mood, both negatively with good mood and also positively with bad mood (third and fourth cells). Including this exceedingly sweet sample (probably bitter too for some) induced good and bad moods that shared almost 50% of their variance (fifth cell, leftmost column, Table 1), a stronger correlation than of either good or bad mood with preference. The scatterplots make plain what was going on (Figure 4).

Figure 4 about here

The near-ideal Step 0.1 of sweetener was very cheering to almost everyone (top left plot in Figure 4), while Step 3 (just at or beyond “never choose”) was less so. In most assessors, neither of these steps induced any bad mood (mid-left scatterplot, Figure 4).

In contrast, on tasting Step 8 some assessors expressed little or no good mood and they were all put into a bad mood -- many strongly so. These data do not distinguish between a revoltingly extreme sweetness and a bitter taste of the saccharin at this level, or even a mixture of both for some assessors. Clearly though, any innate pleasure at sweetness had to contend with some powerfully aversive reactions.

The two scatterplots were almost mirror images, accounting for their high correlation. At the same time, Step 8 induced a large deviation from linearity of mood with choice, giving those more modest correlations.

The other two subpanels showed no reliable correlations between choice preference and mood (third and fourth data lines, Table 1). The assessors tested on a less extreme but still strongly rejected concentration of saccharin (Step 4) produced a trough of good mood that mirrored a peak of bad mood at Step 2 (Figure 5), generating a small negative correlation between good and bad mood (Table 1). A good number of assessors were cheered by tasting the near ideal sweetness (Step 0.1; top left plot in Figure 5). A minority showed the sign of

innate pleasure in improved good mood on tasting Step 4. Most experienced little or no bad mood either near ideal or at reflex-inducing rejected sweetness. However, on tasting the level of sweetness approaching “never choose,” half the panel rated worsening of mood at 5 out of 7 for “extremely bad.” That is, they expressed considerable displeasure.

Figure 5 about here

The preference-mood functions were effectively flat when all three levels of sweetener were ideal or tolerable (Figure 6). Bad mood was rare in this subpanel but, at each of these Steps of saccharin, a substantial proportion rated moderately large improvement in good mood (4 to 6 out of 7 for “extremely good”). This finding is consistent with a correlation between preference and pleasure arising only when an aversively tasting sample is included.

Figure 6 about here

Movement Affect and Drink Sweetness Preference

There were many close parallels between the effects on general mood and effects on the enjoyment of feelings of movement in the mouth or face (right-hand columns of Figures 4-6). This detailed concordance provides further cross-validation of these two sorts of index of pleasure and displeasure.

That broad validity makes a disparity all the more striking. “Nice” and “nasty” feelings of movement were uniquely highly correlated (80% shared variance) in the panel tasting saccharin that was intolerably sweet (Step 4) but far less extreme than Step 8 and yet also not within the tolerated range as Step 3 more often would have been (last cell of middle data column of Table 1; bottom right-hand plot of Figure 5). It is therefore of some importance that this measure also showed no relationship between preference and pleasure in this panel (the two cells above in Table 1). Hence, even with separate unipolar assessments, a relationship between motivation and emotion is an artifact of including a highly aversive sample in the set.

The high correlations between niceness and nastiness of movement in the other two subpanels were less meaningful because they arose largely from categorical behavior: a movement was rated as either both substantially nice and not at all nasty, or nasty and not nice (bottom right of Figures 4 and 6). Nevertheless, the categorization of pleasure or

displeasure in the felt movement sustained the shape of the function of choice preference that was seen with mood (compare right-hand with left-hand upper two plots of Figures 4-6).

DISCUSSION

Separation of Preference and Pleasure

When the tested samples of apple juice included only mildly intolerable levels of over-sweetness, there was no correlation between the direct measure of preference and either of the two indirect measures of pleasure. That lack of a monotonic relationship is evidence that there is a phenomenon of pleasurable experience induced by the taste of a sweetener which is distinct from a preference shown for one level of sweetness over another within a high-quality beverage.

Those ratings of preference relative to ideal for variants of a particular food or drink provide a highly economical way of measuring the strength of influence of any factor in food choice behavior (Conner and others 1986). In the present experiment, just four samples were sufficient to estimate nearly all the untrained panelists' ideal points and excess rejection points for the sweetener in a familiar juice. The one exception was indifferent to levels of sweetener in the range tested and might have produced a measurable preference function if a more strongly sour or sweet sample had been included.

In contrast to such ease of measuring an influence on preference, there are inevitable difficulties in assessing pleasure. No known wording by itself can pick out pleasure from preference, and so that job must be done less directly. The logically simplest evidence for any hypothetical construct is converging operations from two different effects of that mediating variable. Furthermore, each of those indicators of pleasure has to be disconfounded from preference under the conditions of testing. This report illustrates one way of running such an experimental design.

It may be questioned whether it is worth tackling these experimental complexities. Sheer curiosity is of course legitimate but its pursuit may be difficult to resource. Is genuinely hedonic assessment of any practical value? Expected sensory factors are likely to dominate most choices of foods, but sometimes emotions may play a part (Macht 2008). In some circumstances, the links between the senses and the emotions may be critical.

Food and Mood

Hence one reason to investigate experiences of pleasure as distinct from the disposition to accept is clarify the emotional functions of food consumption. Experiencing a pleasure could create or boost an overall good mood. Also, pleasure might counter a bad mood if one existed at the time. Indeed, one of this paper's indirect assessments of pleasure was the strength of "good" mood. A separate assessment of the intensity of "bad" mood might improve understanding of the overall affective state.

In addition, it may be worth attempting to differentiate among types of positive affect (Sauter 2010). However, the respondent's actual mood is not identified simply by a word that the investigator uses (Goodchild and others 2005). Once again, experimental designs are needed that use converging operations on products having known variations.

A possibility worth further investigation is that sensual pleasure comes from elicited movements, not just straight from the stimulation. The second indirect assessment of pleasure in this paper was enjoyment of those particular movements that become more likely with increasing stimulation, or even of just an impression that the movements occurred (Booth and others 2010). In the present case of strong sweetness, the sensual pleasure might therefore be based at least partly on sensing or imagining the ingestive movements of the tongue that are stimulated in the human baby by sugar (Rosenstein and Oster 1988). Analogously, displeasure might be generated from grimaces at bitter or sour tastes (Rosenstein and Oster 1988). The convergence of this implicit index of pleasure with the mood ratings in the present experiment supports such conjectures, as also does the analysis of the verbalized concepts of movement in the mouth and face (Booth and others 2010).

Strategies for Practice

The success of the method for separating preference from pleasure presented here and earlier (Booth and others 2010) has wide implications for sensory studies. These considerations range from the design of response scales to the selection of products to test.

Response Categories

It is over sixty years since Peryam and colleagues proposed that scores for food preference in the market be elicited by nine ordinal categories ranging from very strong liking to very strong dislike (Jones and others 1955, Peryam and Girardot 1952, Peryam and Haynes 1957, Peryam and Pilgrim 1957). The layout was immediately popular and remains the most commonly used tool for the assessment of preference in both academic and commercial research, either in its original form or modified in various ways (e.g., Bartoshuk and others 2006, Schutz 1965, Schutz and Cardello 2001; see review by Lim 2011). Nevertheless there are major difficulties in constructing usefully quantitative preference functions from the responses made in any bipolar multiple-category format. These issues were detailed in the Introduction in order to justify the responses used in the present experiment. Additional practical implications are considered now.

The evidence and argument in this paper provide clear pointers to the most effective use of Peryam's categories for the measurement of food preference behavior. The bipolar mid-category of "neither like nor dislike" could be the lower of two anchor points for an otherwise unlabelled array of degrees of preference. To serve as the one other anchor, the phrase "like extremely" needs to be changed to refer explicitly to the logical maximum of liking, the ideal point, e.g. "could not be liked more" or "maximally liked." That phrase would anchor the ratings on the personally most preferred variant of the assessed sort of food in the use that the test session simulates. Even if some other factors in a test sample are below optimum for a panelist, the psychophysics of preference provides the means to estimate that ideal point and how far any tested or untested sample is from it (Booth and Freeman 1993, Booth and others 2003a,b).

Multiple anchor phrases are good only for surveys of relative popularity of directly competing products. Such data must be analyzed only as counts of respondents at each

anchor. It would be no more difficult for consumers to place items in positions between two phrases such as “could not like more” and “neither like nor dislike.” Then preferences can be measured on psychological scales by simply programmed calculations. Separate placing of responses between “greatest dislike imaginable” and “neither liked nor disliked” can screen for any incidence of antagonism to a product in the population represented by the panel. More realistic pairs of anchors for preference invoke the act of selection of the rated sample from the alternatives considered by the assessor, such as “I’d always choose one like this in this context” and “I never would.”

Wanting a food or drink (appetite for the item) is a contextualized preference. That is, the attractiveness of a food increases with closeness to a usual situation for eating it. The attraction also decreases as that item and other foods are eaten. Hence if a substantial number of items are consumed in a session, aversiveness or at least a transient loss of preference should also be screened for. “I’m full” need not refer to a sensation of stretch within the abdomen; it may simply mean that the speaker wants to stop eating whatever is available (Booth 1976, Booth and others 1982). Young children have to learn such conventions for satiety and its expression (Laurier and Wiggins 2011). In the same way, they learn which foods in which eating situations have the sensed characteristics and the conceptualized attributes to be selected right then. Both rumbling or distended stomachs and also sensual pleasure or displeasure can be by-the-by (Booth and Blair 1979). Consumption tests need to be monitored for accumulating satiety in anybody, as well as initial unacceptability in some.

Hence for theoretical or practical use of data on preferences, the questioning should at the very least focus on eating or shopping. Better, the main factors in a particular situation requiring the decision should be specified (Cardello and others 2000). The measurement of preference would be much more effective if the whole situation, material and social, were mimicked as closely as possible. Assessment of liking in the abstract carries the unrealistic preconception of a food with a context-free ‘palatability’, or indeed a fixed satisfaction or satiating power (Booth 1990b). Hence, both to make scientific sense and for application in health or commerce, a measure of the strength of the disposition to select a food needs to incorporate measures of the momentary influences on the observed response.

Preference as the First Sensory Measure

Whether characterized or not, degree of preference or relative acceptance is an integrative measure of sensory impact. It reflects the actual interactions among sensed factors.

Moreover, these mental interactions include the context of conceptual attributes that each panelist brings to the test session. In addition, the investigator generates attributions during instructions or in labels like those used in conjoint analysis, including sensory expectations. For example, identical drinks of coffee, varied only in caffeine content, were liberated from conceptualization as varying in bitterness by asking panelists to rate each “brand” of coffee (Booth and others 2011b,c).

In short, preferences are no less sensory than are sensory intensities. Indeed, for many purposes, the integrative sensitivity of preference is much more useful than the putatively analytical sensitivity of a sensory vocabulary. There is certainly no need to bridge a supposed gap between sensory concepts and acts of acceptance by collecting large amounts of unfocused data to be mined for statistical patterns.

Characterized preferences are no different from characterized intensities. The use of a verbal concept may improve sensitivity to a sensed factor if that concept provides a good description, but it could worsen sensitivity if another verbal concept works better for that panelist (Booth 2011a, Booth and others 2011a, Booth and Freeman 2014).

Selection of Items for Testing

The impact of sensing can only be assessed by picking out samples that contrast in levels of each ingredient or instrumental measure independently of others. This is the basic principle of causal analysis. Hence a sensory study conducted for business purposes should be designed to test specific hypotheses about consumers’ action on the bundle of sensed and attributed features that constitute the (sub)branded product under investigation.

The test samples need to be selected into a set capable of measuring the impact of hypothesized variations in the product range, first on a fully integrative judgment such as match to personally ideal, most familiar or usual brand, and then on user or expert generated vocabulary if some analytical characterization might help to test an hypothesis. Neither the

number of samples nor the size of the analytical vocabulary should be larger than needed to fill any gaps in the evidence on the originally formulated hypotheses that have been generated by personal preference or familiarity.

The present results show that a correlation between preference/aversion and pleasure/displeasure can be an artifact of including an aversive test sample. When only marketable samples were tested, there was no correlation between the usual preference and individual pleasure. That fact is important in itself if consumers' actions are what matter, rather than unrelated phenomenology. However, the observation also has strategic importance, especially if testing has not been tailored to panelists. The panel's data should not be grouped for analysis without screening for anomalous panelists or test products. If ranks or quantities range from preference to aversion, aversion to one or more of the samples may go unnoticed. Liking and dislike should be rated separately, so that the situation is plain at the panel level as well as in each panelist.

Dangers from the Unique Impact of Sweetness

The present results strengthen the evidence that, without properly conducted sensory optimization, there can be serious dangers to both users and producers of foods and drinks containing intense or bulk sweeteners. We are born with a reflex that draws material into the mouth with increasing vigor as its sweet taste get stronger, without limit. Ordinarily "the sweeter the better" is suppressed by the universal mechanism of learning to like whatever becomes familiar in physiologically and socially congenial circumstances. A particular food or drink consumed in a usual situation is most accepted when its sweetness and every other feature are at the levels remembered from previous occasions (Booth and Freeman 1993, Booth and others 2011b).

In newborn infants, the sweetness of human milk elicits the unlearned reflex which helps to ensure that fluid goes straight from the breast to the throat. Yet the baby probably learns rapidly to prefer breastmilk's level of sweetness (and levels of sodium salts, viscosity, warmth and aroma) over either less or more. Certainly the weanling acquires a preference for the level of each taste in each food provided (Harris and Booth 1987).

If child or adult consumers are not presented with extra sweet drinks or foods, they will not come to like them (Conner and Booth 1988, Conner and others 1988b, Wansink and others 2006). Worse, if they are presented with sufficiently strong sweetness, the infantile milk acceptance reflex is liable to break through. Hence, even the young adults whose experience of natural levels of sugars and acids has taught them that extremely sweet apple juice is revolting, as observed in the present experiment (Booth and others 2010), may be inveigled into trying a super-sweet drink again because of an oral thrill from that taste.

Enjoyment without Thrills

Finally, those who would like sensual pleasure to be the usual experience with any food or drink should consider the implications of the fact that evidence of the experience of pleasure was obtained in this experiment with intolerably over-sweet samples. Of course, strong expressions of pleasure in food can be appropriate in some social situations. There is no doubt either that the activity of eating can be intensely enjoyable. The question posed by this report is whether eating is normally accompanied by physical thrills. A climactic experience with every mouthful might be difficult to cope with!

Eaters and drinkers do not need sensual fireworks to choose among their foods and beverages and to be happy with them. Hedonism about food and drink is particularly unattractive if it relies on such strong sweetening that the infantile reflex overrides the normal dislike for that extreme. Great enjoyment and full satisfaction can be obtained with an educated palate. Sweetness should take its adult place among other sensory characteristics, each at the level most liked by the individual for a particular product in use within a familiar situation.

REFERENCES

- ASHBY, F.G., and ENNIS, D.M. 2002. A Thurstone-Coombs model of concurrent ratings with sensory and liking dimensions. *J. Sensory Studies* 17, 43-59.
- BARTOSHUK, L.M., SNYDER, D.J., and DUFFY, V.B. 2006. Hedonic gLMS: valid comparisons for food liking/disliking across obesity, age, sex and PROP status. *Chem. Sens.* 31, A50.
- BERRIDGE, K.C. 1996. Food reward: brain substrates of wanting and liking. *Neurosci. Biobehav. Rev.* 90, 36-42.
- BOOTH, D.A. 1976. Approaches to feeding control. In *Appetite and food intake* (T. Silverstone, ed.) pp. 417-478, Abakon / Dahlem Konferenzen, West Berlin, F.R.G.
- BOOTH, D.A. 1988a. Practical measurement of the strengths of actual influences on what consumers do: scientific brand design. *J. Market Res. Soc. (U.K.)* 30, 127-146.
- BOOTH, D.A. 1988b. Estimating JNDs from ratings. [Abstract: AChemS-10] *Chem. Sens.* 13 (4), 675-676.
- BOOTH, D.A. 1990. Designing products for individual customers. In *Psychological bases of sensory evaluation* (R.L. McBride & H.J.H. MacFie, eds.) pp. 163-193, Elsevier Applied Science. London, UK
- BOOTH, D.A. 1991. Learned ingestive motivation and the pleasures of the palate. In *The hedonics of taste* (R.C. Bolles, ed.) pp. 29-58, Erlbaum, Hillsdale NJ.
- BOOTH, D.A. 1995. Cognitive processes in odorant mixture assessment. *Chem. Sens.* 20, 639-643.
- BOOTH, D.A. 2009a. Lines, dashed lines and “scale” ex-tricks. Objective measurements of appetite versus subjective tests of intake. *Appetite* 53, 434-437.
- BOOTH, D.A. 2009b. The basics of quantitative judgment. How to rate the strength of appetite for food and its satiation. *Appetite* 53, 438-441.
- BOOTH, D.A. 2014. Measuring sensory and marketing influences on consumers' choices among food and beverage product brands. *Trends Food Sci. Technol.* 35 (3), 129-137.
- BOOTH, D.A. 2015. Scientific measurement of sensory preferences using stimulus tetrads. *J. Sensory Stud.* *in press*.
- BOOTH, D.A., and BLAIR, A.J. 1989. Objective factors in the appeal of a brand during use by the individual consumer. In *Food acceptability* (D.M.H. Thomson, ed.) pp. 329-346, Elsevier Applied Science, London, U.K.

- BOOTH, D.A., and CONNER, M.T. 1991. Characterisation and measurement of influences on food acceptability by analysis of choice differences: theory and practice. *Food Qual. Pref.* 2, 75-85.
- BOOTH, D.A., and CONNER, M.T. 2009. Salt in bread. *J. Food Sci.* 74(3), vii-viii.
- BOOTH, D.A., and FREEMAN, R.P.J. 1993. Discriminative feature integration by individuals. *Acta Psychol. (Amst.)* 84, 1-16.
- BOOTH, D.A., and SHEPHERD, R. 1988. Sensory influences on food acceptance - the neglected approach to nutrition promotion. *BNF Nutr. Bull.* 13(1), 39-54.
- BOOTH, D.A., HIGGS, S., SCHNEIDER, J., and KLINKENBERG, I. (2010). Learnt liking versus inborn delight. Can sweetness give sensual pleasure or is it just motivating? *Psychol. Sci.* 21, 1656-1663.
- BOOTH, D.A., MATHER, P., & FULLER, J. (1982). Starch content of ordinary foods associatively conditions human appetite and satiation, indexed by intake and eating pleasantness of starch-paired flavours. *Appetite* 3, 163-184.
- BOOTH, D.A., MOBINI, S., EARL, T., and WAINWRIGHT, C.J. 2003a. Consumer-specified instrumental quality of short-dough cookie texture using penetrometry and break force. *J. Food Sci.: Sens. Nutr. Qual. Food* 68, 382-387.
- BOOTH, D.A., MOBINI, S., EARL, T., and WAINWRIGHT, C.J. 2003b. Market-optimum instrumental values from individual consumers' discriminations of standard sensory quality of the texture of short-dough biscuits. *J. Food Qual.* 26(5), 425-439.
- BOOTH, D.A., O'LEARY, G., LI, L., and HIGGS, S. 2011a. Aversive viscerally referred states and thirst accompanying the sating of hunger motivation by rapid digestion of glucosaccharides. *Physiol. Behav.* 102, 373-381.
- BOOTH, D.A., SHARPE, O., and CONNER, M.T. 2011b. Discrimination without description of differences. Implicit or fully subconscious? <http://epapers.bham.ac.uk/682/> (accessed 18 March 2015)
- BOOTH, D.A., SHARPE, O., and CONNER, M.T. 2011c. Differential gustatory sensitivity to caffeine in normal use points to supertasters, tasters and non-tasters. *Chemosens. Perc.* 4, 154-169.
- BOOTH, D.A., SHARPE, O., FREEMAN, R.P.J., and CONNER, M.T. 2011d. Insight into sight, touch, taste and smell by multiple discriminations from norm. *See. Perc.* 24, 485-511.
- BOOTH, D.A., THOMPSON, A.L. and SHAHEDIAN, B. 1983. A robust, brief measure of an individual's most preferred level of salt in an ordinary foodstuff. *Appetite* 4, 301-312.

- BOWMAN, S.J., BOOTH, D.A., PLATTS, R.G., and the UK Sjögren's Interest Group. 2004. Measurement of fatigue and discomfort in primary Sjögren's syndrome using a new questionnaire tool. *Rheumatol.* *43*, 758-764.
- CABANAC, M. 1971. Physiological role of pleasure. *Science* *193*, 1103-1107.
- CARDELLO, A., and SCHUTZ, H.G. 2004. Research note. Numerical scale-point locations for constructing the LAM (labeled affective magnitude) scale. *J. Sensory Stud.* *19*, 341-346.
- CARDELLO, A., LAWLESS, H.T. and SCHUTZ, H.G. 2008. Effects of extreme anchors and interior label spacing on labeled affective magnitude scales. *Food Qual. Pref.* *19*, 473-480
- CARDELLO, A.V., SCHUTZ, H.G., SNOW, C., and LESHER, L. 2000. Predictors of food acceptance, consumption and satisfaction in specific eating situations. *Food Qual. Pref.* *11*, 201-216.
- CONNER, M.T., and BOOTH, D.A. 1988. Preferred sweetness of a lime drink and preference for sweet over non-sweet foods, related to sex and reported age and body weight. *Appetite* *10*, 25-35.
- CONNER, M.T., and BOOTH, D.A. 1992. Combining measurement of food taste and consumer preference in the individual: reliability, precision and stability data. *J. Food Qual.* *15*, 1-17.
- CONNER, M.T., BOOTH, D.A., CLIFTON, V.J., and GRIFFITHS, R.P. (1988a). Individualized optimization of the salt content of white bread for acceptability. *J. Food Sci.* *53*, 549-554.
- CONNER, M.T., HADDON, A.V., and BOOTH, D.A. (1986). Very rapid, precise measurement of effects of constituent variation on product acceptability: consumer sweetness preferences in a lime drink. *Lebens.-Wiss. u. -Technol. [Food Sci. Technol.]* *19*, 486-490.
- CONNER, M.T., HADDON, A.V., PICKERING, E.S., and BOOTH, D.A. 1988b. Sweet tooth demonstrated: individual differences in preference for both sweet foods and foods highly sweetened. *J. Appl. Psychol.* *73*, 275-280.
- CONNER, M.T., LAND, D.G., and BOOTH, D.A. 1987. Effects of stimulus range on judgments of sweetness intensity in a lime drink. *Br. J. Psychol.* *78*, 357-364.
- CONNER, M.T., PICKERING, E.S., BIRKETT, R.J., and BOOTH, D.A. 1994. Using an individualised attribute tolerance model in consumer acceptability tests. *Food Qual. Pref.* *5*, 225-232.

- CORDONNER, S.M., and DELWICHE, J.F. 2008. An alternative method for assessing liking. Positional relative rating versus the 9-point hedonic scale. *J. Sensory Stud.* 23, 284-292.
- ELLERMEIER, W., and FAULHAMMER, G. 2000. Empirical evaluation of axioms fundamental to Stevens's ratio-scaling approach: I. Loudness production. *Perc. Psychophys.* 62, 1505–1511.
- FINLAYSON, G., KING, N., and BLUNDELL, J.E. 2007. Is it possible to dissociate ‘liking’ and ‘wanting’ for foods in humans? A novel experimental procedure. *Physiol. Behav.* 90, 36-42.
- FREEMAN, R.P.J., and BOOTH, D.A. 2010. Users of ‘diet’ drinks who think that sweetness is calories. *Appetite* 55, 152-155.
- FREEMAN, R.P.J., RICHARDSON, N.J., KENDAL-REED, M.S., and BOOTH, D.A. 1993. Bases of a cognitive technology for food quality. *Br. Food J.* 95(9), 37-44.
- GOODCHILD, C.E., PLATTS, R.G., TREHARNE, G.J., and BOOTH, D.A. 2005. Excessive negative affect and deficient positive affect in anxiety and depression: balancing the valences in the Hospital Anxiety and Depression Scale (HADS). *Health Psychol. Update* 14(2), 45-50.
- GOODCHILD, C.E., TREHARNE, G.J., BOOTH, D.A., KITAS, G.D., and BOWMAN, S.J. 2008. Measuring fatigue among women with Sjögren’s syndrome or rheumatoid arthritis: a comparison of the Profile of Fatigue (ProF) and the Multidimensional Fatigue Inventory (MFI). *Musculoskel. Care* 6, 31-48.
- HARRIS, G., and BOOTH, D.A. 1987. Infants' preference for salt in food: its dependence upon recent dietary experience. *J. Reprod. Infant Psychol.* 5, 97-104.
- HOFMANN, W., VAN KONINGSBRUGGEN, G.M., STROEBE, W., RAMANATHAN, S., and AARTS, H. 2010. As pleasure unfolds: hedonic responses to tempting food. *Psychol. Sci.* 21, 1863-1870.
- JAEGER, S.R., WAKELING, I.N., and MACFIE, H.J.H. 2000. Behavioural extensions to preference mapping: the role of synthesis. *Food Qual. Pref.* 11, 349-359.
- JONES, L.V., PERYAM, D.R., and THURSTONE, L.L. 1955. Development of a scale for measuring soldiers’ food preferences. *Food Res.* 20, 512-520.
- KNIBB, R.C., SMITH, D.M., BOOTH, D.A., ARMSTRONG, A.M., PLATTS, R.G., MACDONALD, A., and BOOTH, I.W. 2001. No unique role of nausea attributed to eating a food in the recalled acquisition of sensory aversion to that food. *Appetite* 36, 225-234.
- MACHT, M. 2008. How emotions affect eating: a five-way model. *Appetite* 50, 1-11.

- LAURIER, E., and WIGGINS, S. 2011. Finishing the family meal. The interactional organisation of satiety. *Appetite* 36, 53-64.
- LAWLESS, H.T., CARDELLO, A.V., CHAPMAN, K. W., LESHNER, L.L., GIVEN, Z., and SCHUTZ, H.G. 2010a. A comparison of the effectiveness of hedonic scales and end-anchor compression effects. *J. Sensory Stud.* 25, Supplement 1, 18-34.
- LAWLESS, H.T., POPPER, R., and KROLL, B.J. 2010b. A comparison of the labeled magnitude (LAM) scale, an 11-point category scale and the traditional 9-point hedonic scale. *Food Qual. Pref.* 21, 4-12.
- LAWLESS, H.T., SINOPOLI, D., & CHAPMAN, K.W. 2010c. A comparison of the labeled affective magnitude scale and the 9-point hedonic scale and examination of categorical behavior. *J. Sensory Stud.* 25, Supplement 1, 54-66.
- LIM, J. 2011. Hedonic scaling. A review of methods and theory. *Food Qual. Pref.* 22, 733-747.
- MCBRIDE, R.L., and BOOTH, D.A. (1986). Using classical psychophysics to determine ideal flavour intensity. *J. Food Technol.* 21, 775-780.
- MOSKOWITZ, H. 2002. Knowledge building and optimization strategies for a product used in different carriers. *J. Sensory Stud.* 17, 327-349.
- MOSKOWITZ, H.R., and SIDEL, J.L. 1971. Magnitude and hedonic scales of food acceptability. *J. Food Sci.* 36, 677-680.
- MOTTRAM, D.S. 1998. Chemical tainting of foods. *Int. J. Food Sci. Technol.* 33, 19-29.
- NICOLAS, L., MARQUILLY, C., and O'MAHONY, M. 2010. The 9-point hedonic scale. Are words and numbers compatible? *Food Qual. Pref.* 21, 1008-1015.
- PARDUCCI, A. 1965. Category judgment: a range-frequency model. *Psychol. Rev.* 72, 407-418.
- PERYAM, D.R., and GIRADOT, N.F. 1952. Advanced taste test method. *Food Eng.* 24(7), 58-61, 194.
- PERYAM, D.R., and HAYNES, J.G. 1957. Prediction of soldiers' food preferences by laboratory methods. *J. Appl. Psychol.* 41, 2-6.
- PERYAM, D.R., and PILGRIM, F.J. 1957. Hedonic scale method of measuring food preferences. *Food Technol.* 11(September), 9-14.
- POULTON, E.C. 1989. *Bias in quantifying judgments*, Laurence Erlbaum Associates, Hove, U.K.

- RISKEY, D.R., PARDUCCI, A., and BEAUCHAMP, G.K. 1979. Effects of context on judgments of sweetness and pleasantness. *Perc. Psychophys.* 26, 171-176.
- ROSENSTEIN, D., and OSTER, H. 1988. Differential facial responses to four basic tastes in newborns. *Child Dev.* 59, 1555-1568.
- SAUTER, D. 2010. More than happy: the need for disentangling positive emotions. *Curr. Dir. Psychol. Sci.* 19, 36-40.
- SCHUTZ, H.G. (1965). A food action scale for measuring food acceptance. *J. Food Sci.* 30, 365-374.
- SCHUTZ, H.G., and CARDELLO, A.V. 2001. A labeled affective magnitude (LAM) scale for assessing food liking/disliking. *J. Sensory Stud.* 16, 117-159.
- STEINER, J.E., GLASER, D., HAWILO, M.E., and BERRIDGE, K.C. (2001). Comparative expression of hedonic impact: affective reactions to taste by human infants and other primates. *Neurosci. Biobehav. Rev.* 25, 53-74.
- STEVENS, S.S. (1957). On the psychophysical law. *Psychol. Rev.* 64, 153–181.
- THOMPSON, R.F., and SPENCER, W.A. (1966). Habituation: a model phenomenon for the study of neuronal substrates of behavior. *Psychol. Rev.* 73, 16–43.
- TORGERSON, W.S. 1958. *Theory and methods of scaling*. John Wiley, New York NY.
- WANSINK, B., BASCOUL, G., and CHEN, G.T. 2006. The sweet tooth hypothesis. How fruit consumption relates to snack consumption. *Appetite* 47, 107-110.

TABLE 1. CORRELATIONS OF RANKS FOR SCATTERPLOTS IN FIGURES 3-6.

Correlation (Spearman's ρ) between ratings of pleasure and displeasure or between one of those scores and Step numbers of sweetener that corresponded to scores for likelihood of choice ("always choose" to "never choose", and worse than never, i.e. degrees of rejection).

Pair of variables	Choices (Sweetener Steps)		
	Always (0.1), Reject (3), Aversive (8)	Always (0.1), Just Never (2), Strong Reject (4)	Always (0.1), Just Never (2), Reject (3)
Good Mood vs Nice Movement	0.86***	0.74***	0.50**
Bad Mood vs Nasty Movement	0.77***	0.45*	0.67***
Choice vs Good Mood	-0.46**	-0.20	0.08
Choice vs Bad Mood	0.43**	0.03	-0.18
Mood, Good vs Bad	-0.69***	-0.30	-0.02
Choice vs Nice Movement	-0.37*	-0.15	0.18
Choice vs Nasty Movement	0.44**	-0.03	-0.26
Movement, Nice vs Nasty	-0.80***	-0.89***	-0.76***
<i>N</i>	14	7	10

*** $P < 0.0005$, ** $P < 0.005$, * $P < 0.05$, all two-tailed.

Titles of Figures, with subtitles and footnotes

FIG. 1. INITIAL MEASUREMENT OF INDIVIDUALS' PREFERENCES

Each graph shows the raw data from one panelist fitted to the theoretical equation.

Note. Each assessor's four data-pairs were fitted to a right hyperbola (vertical conic section) by Co-Pro software (Booth and others 2011). Continuous line (apple green): least squares fit. Broken lines (sky blue): tangents to the fitted hyperbola. Horizontal axis: concentration of saccharin in the tasted sample in logarithms to the base 10 of grams per liter (equal ratios). Vertical axis: first (and only) response (R1) to each sample, with "always choose" scored as zero and "never choose" scored as -9. One panelist rated all four samples as "always choose" and so those data have been omitted from this Figure.

FIG. 2. FREQUENCIES OF INDIVIDUALS' DISCRIMINATED IDEAL POINTS

The number of panelists (N = 18, Fig. 1) whose ideal range (ideal point \pm one discrimination unit, HDR) included the indicated concentration of saccharin (in units of \log_{10} g/l of tasted juice).

Note . This frequency polygon does not have predetermined bins. The unit count for a particular panelist begins on the left at one HDR below IP and finishes on the right at one HDR above IP. Hence that person's IP for saccharin in this context was at the unmarked concentration at the middle of that line for one increment in count.

FIG. 3. CROSS-VALIDITIES OF GENERAL AND SWEET-SPECIFIC AFFECT

Scatterplots of individuals' ratings of general affect (good or bad mood) and sweet-specific affect (nice or nasty feeling of orofacial movement) in three subpanels (top, middle and bottom rows). Compare with correlations in Table 1.

[Note, over the page]

[Figure 3's legends continued]

Note. As indicated in brackets above each graph, each subpanel received a near-ideal level of saccharin (Step 0.1) and level fairly close to just rejected (Step 2 or 3). The subpanels varied in the tested level above just rejected, from as far as Step 8 (top row) to as close as Step 3 (bottom row), with Steps in concentration ratios of about 1.2.

Left-hand column of graphs: pleasurable affect. Right-hand column: displeasure.

Horizontal axes: change in general mood on tasting each sample, from no change (scored as zero) to “extreme” change (scored as 7).

Vertical axes: how “nice” or “nasty” was any feeling of movement in the mouth (or face) on tasting the sample for one level of sweetness, from no affect (0) to “extreme” (7).

Data points: two-letter ID codes of panelists.

FIG. 4. RELATIONSHIPS BETWEEN PLEASURE OR DISPLEASURE AND ACCEPTED OR REJECTED SWEETENER LEVELS, WITH ONE EXTREMELY STRONGLY REJECTED LEVEL

Raw data from the individuals' data in the uppermost subpanel of Figure 3 (N = 14), who were presented saccharin sweetener at Step 0.1 (near ideal), Step 3 (sweeter than the rejection point) or Step 8, a great way beyond rejection and also liable to taste bitter.

Note. Under these conditions (Booth and others 2010, and Figure 3 in this paper), pleasure (as distinct from preference) was indexed by positive affect (“good” mood) and by a concept of movement in the mouth or face that is pleasurable (“nice”), and displeasure by “bad” mood and a “nasty” felt movement.

Bottom pair of graphs: at the extremely unpleasant level of saccharin (Step 8), both indexes of pleasure and displeasure were negatively correlated.

Key to data points and vertical axes: see Note to Figure 3.

FIG. 5. RELATIONSHIPS BETWEEN PLEASURE OR DISPLEASURE AND ACCEPTED OR REJECTED SWEETENER LEVELS, WITH ONE STRONGLY REJECTED LEVEL
Ratings by individuals in the middle panel of Figure 3 (N = 7), presented saccharin at Step 0.1 (near ideal), 2 (near rejection) and 4, well beyond rejection.

Note. Bottom pair of graphs: at the clearly rejected level of saccharin (Step 4), pleasure and displeasure were negatively correlated but good and bad mood were not.

Keys to data points and vertical axes are given in the Notes to Figures 4 and 3.

FIG. 6. RELATIONSHIPS BETWEEN PLEASURE OR DISPLEASURE AND ACCEPTED OR REJECTED SWEETENER LEVELS, WITH ONE LEVEL JUST REJECTED.
Ratings from bottom subpanel in Figure 3 (N = 10), presented saccharin at Step 0.1 (near ideal), 2 (near rejection) and 3, just beyond rejection.

Note. Bottom pair of graphs: at the level of saccharin on the border liking and dislike (Step 3), neither pleasure and displeasure nor good and bad mood were substantially correlated.

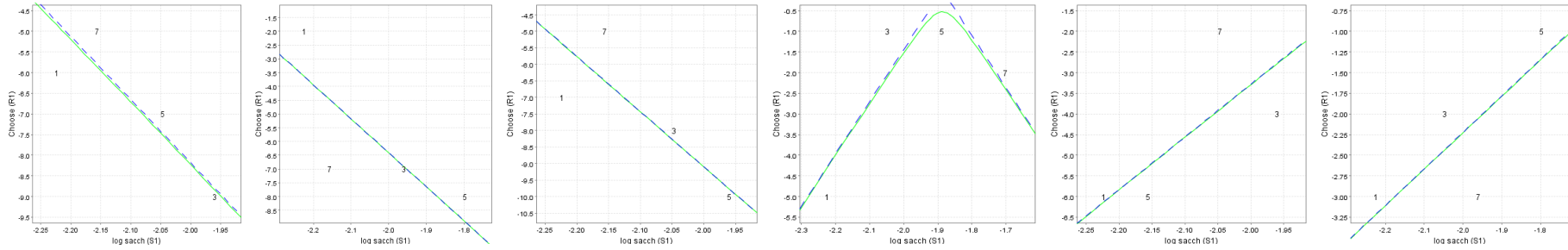
Key in Notes to Figures 3 and 4.

Last line of the Portrait page before Landscape – do NOT delete this line.

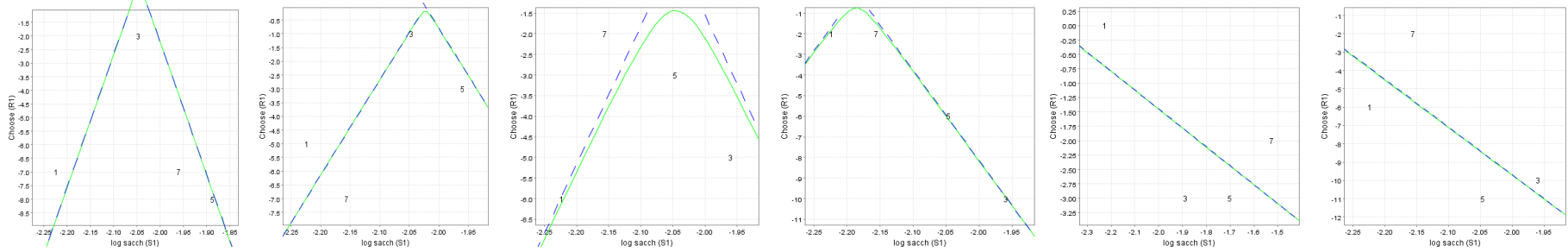
1 Figure 1

2
3
4

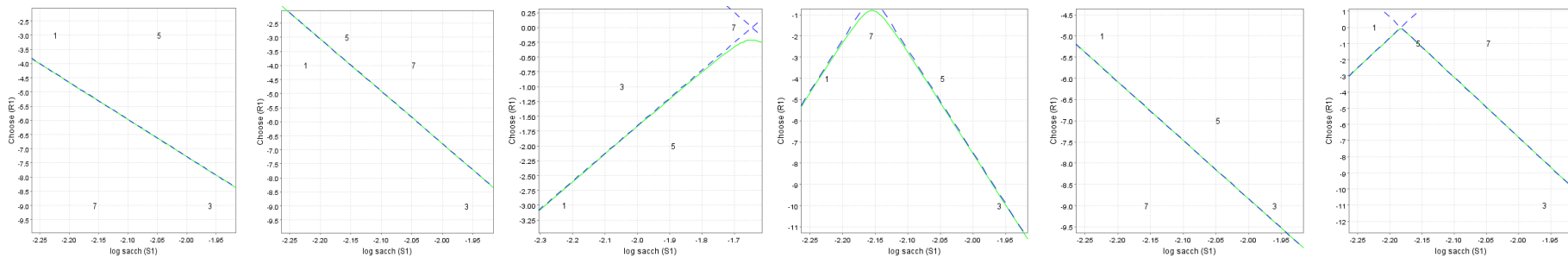
5



6



7



8

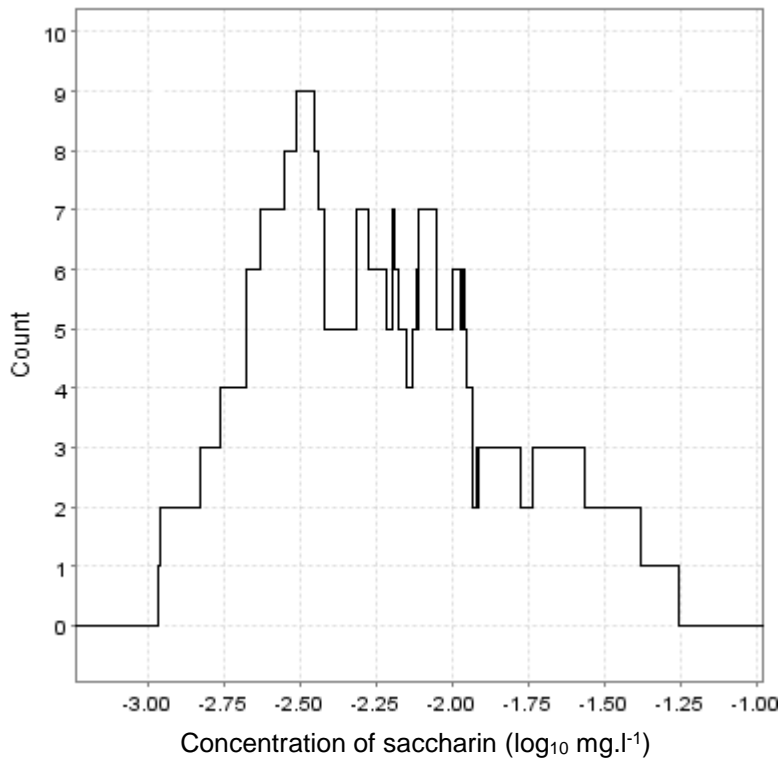
9

10

11

last line before return to Portrait – do NOT delete

12 Figure 2
13
14
15
16
17
18
19

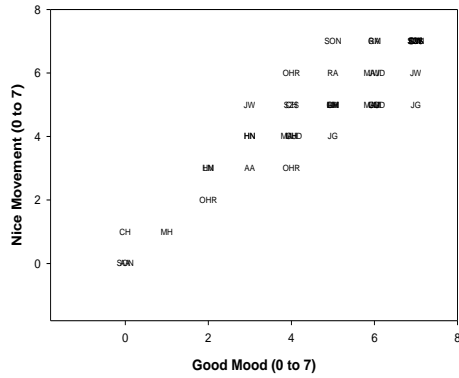


20
21

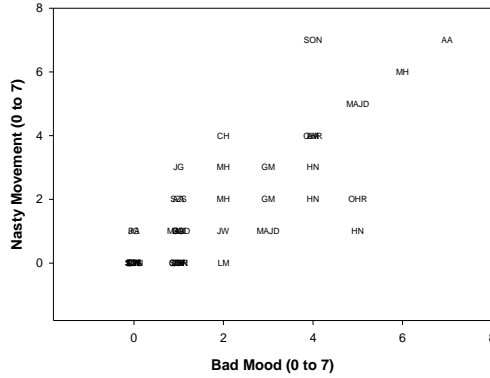
22 Figure 3

23
24
25
26
27

GOOD MOOD vs. NICE MOVEMENT
(Sweetener Steps 0.1, 3, 8)

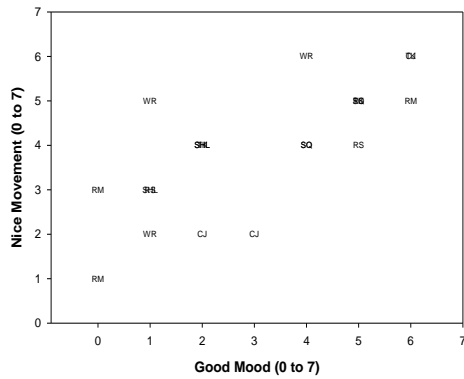


BAD MOOD vs. NASTY MOVEMENT
(Sweetener Steps 0.1, 3, 8)

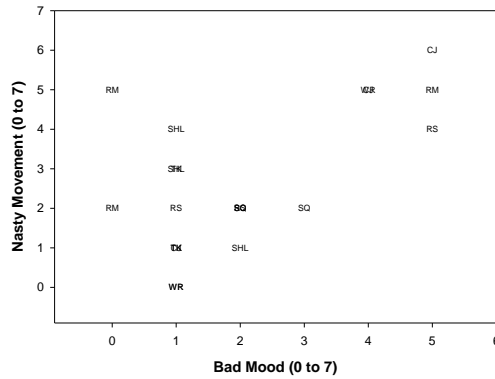


28
29
30

GOOD MOOD vs. NICE MOVEMENT
(Sweetener Steps 0.1, 2, 4)

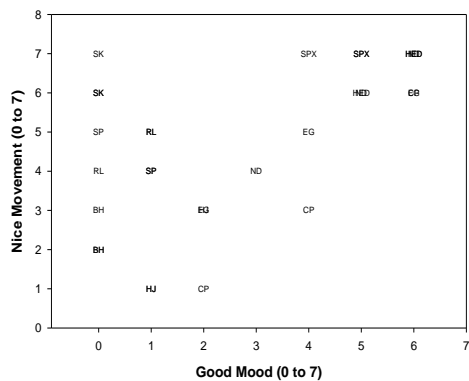


BAD MOOD vs. NASTY MOVEMENT
(Sweetener Steps 0.1, 2, 4)

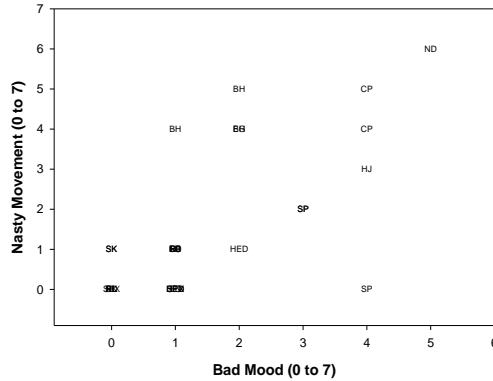


31
32
33

GOOD MOOD vs. NICE MOVEMENT
(Sweetener Steps 0.1, 2, 3)



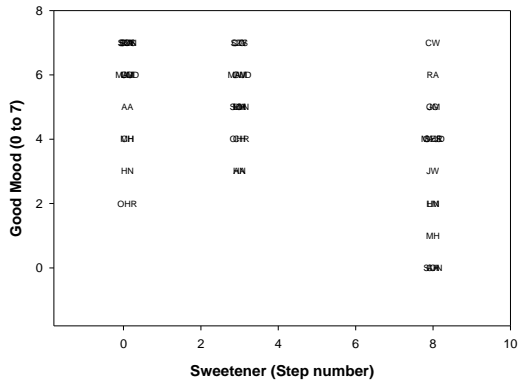
BAD MOOD vs. NASTY MOVEMENT
(Sweetener Steps 0.1, 2, 3)



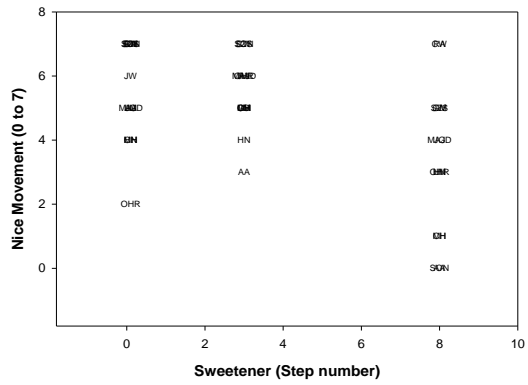
34

35 Figure 4
 36
 37
 38
 39
 40
 41
 42

Sweetener Steps vs. Good Mood

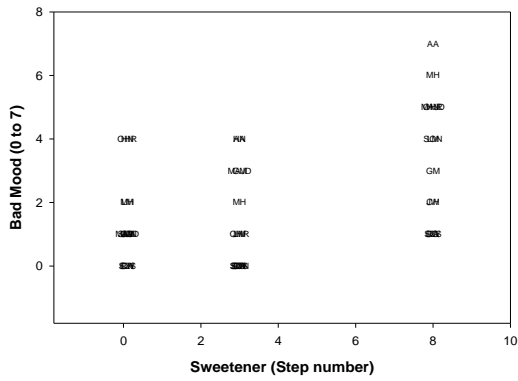


Sweetener Step vs. Nice Movement

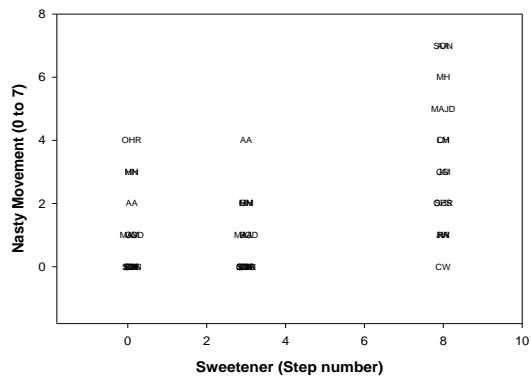


43
 44
 45

Sweetener Steps vs. Bad Mood

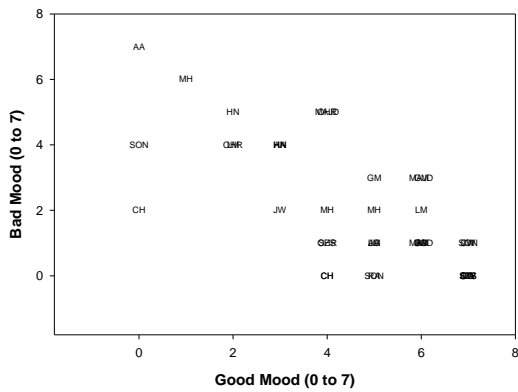


Sweetener Steps vs. Nasty Movement

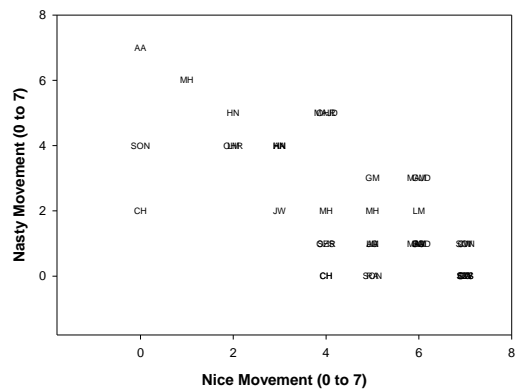


46
 47
 48

GOOD vs. BAD MOOD: top Step 8



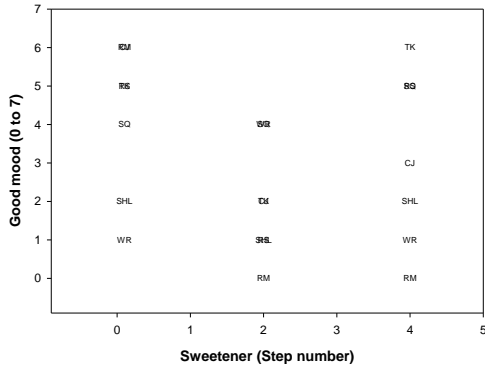
NICE vs. NASTY MOVEMENT: top Step 8



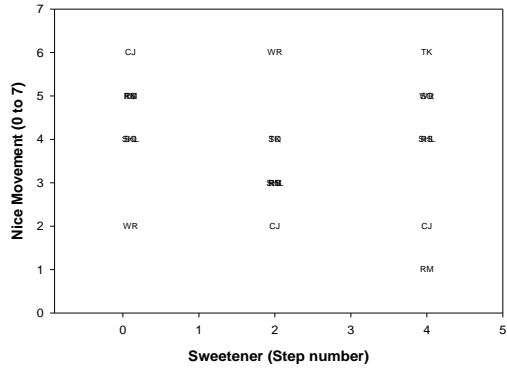
49

50 Figure 5
 51
 52
 53
 54
 55
 56

Sweetener Steps vs. Good Mood

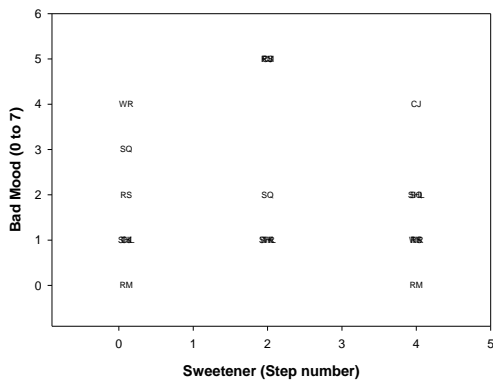


Sweetener Steps vs. Nice Movement

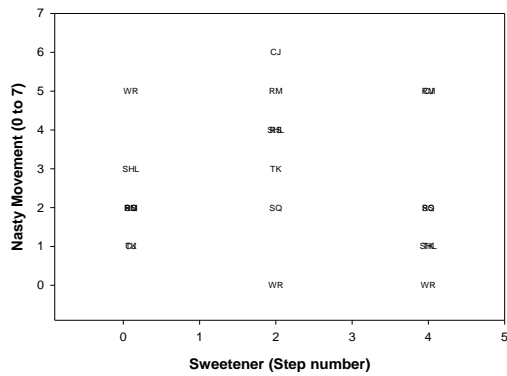


57
 58
 59

Sweetener Steps vs. Bad Mood

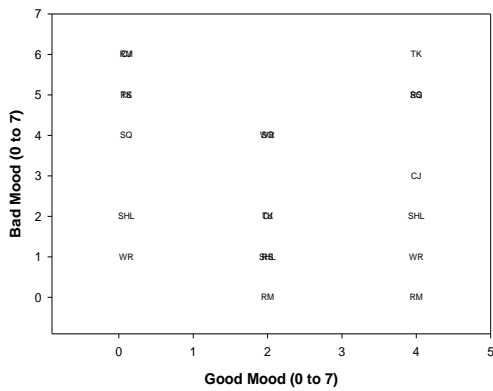


Sweetener Steps vs. Nasty Movement

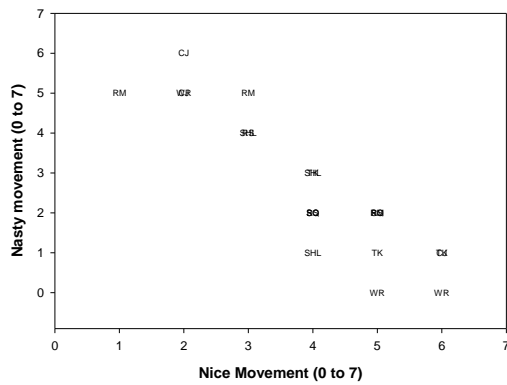


60
 61
 62

GOOD vs. BAD MOOD: top Step 4



NICE vs. NASTY MOVEMENT: top Step 4



63
 64

65 Figure 6

66

67

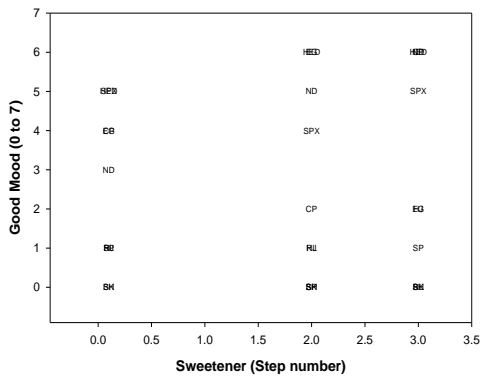
68

69

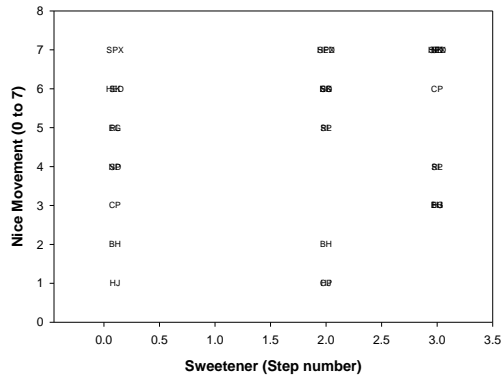
70

71

Sweetener Steps vs. Good Mood



Sweetener Steps vs. Nice Movement

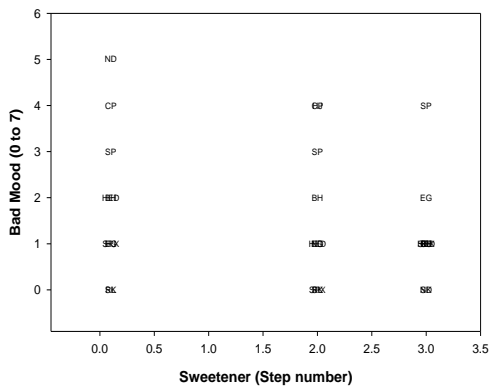


72

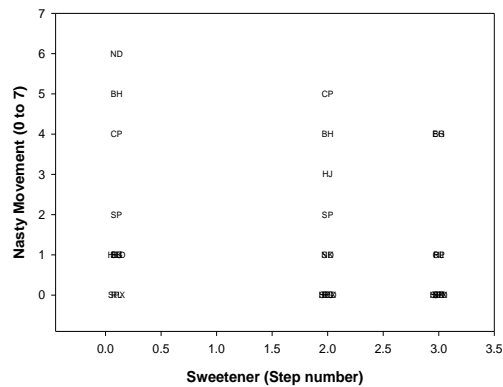
73

74

Sweetener Steps vs. Bad Mood



Sweetener Steps vs. Nasty Movement:

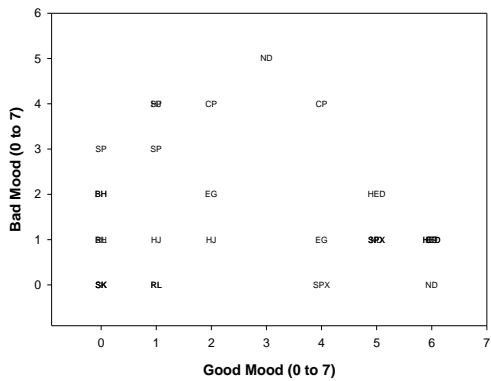


75

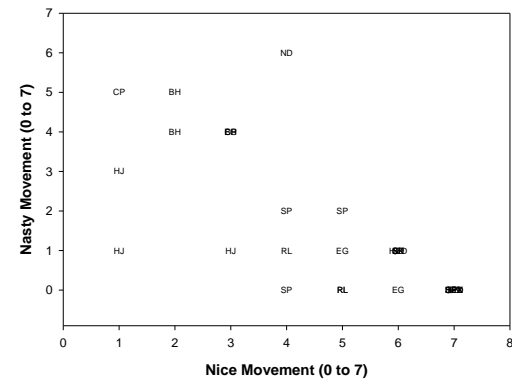
76

77

GOOD vs. BAD MOOD: top Step 3



NICE vs. NASTY MOVEMENT: top Step 3



78

79

80

81