Authors' manuscript

Chemosensory Perception in press (July 2011)
The final publication is available on www.springer.com

Short title (running head)

Caffeine-Perceptual Performance

Gustatory Discriminative Norms for Caffeine in Normal Use Point to Supertasters, Tasters and Nontasters

David A. Booth[1], Oliver Sharpe[1] and Mark T. Conner[2]

[1]School of Psychology, College of Life and Environmental Sciences, University of Birmingham, Edgbaston, Birmingham, B15 2TT, U.K.

[2]Institute of Psychology, University of Leeds, Leeds, LS2 9JT, U.K.

Corresponding author's email address: D.A.Booth@Bham.ac.UK (David Booth)

Abstract

Among the early indications of the existence of supertasters, tasters and non-tasters was a trimodal distribution of sensitivities to the taste of low concentrations of caffeine in water. A similar three peaks of prevalence have now been seen in the concentration of caffeine in the individual's usual coffee drink that is perceived as the most preferred, measured by normed multiple discrimination psychophysics. Furthermore, the mode at the lowest concentrations of caffeine in coffee -- the putative supertasters -- was clearer for preference than for bitterness. This was because rated preference for a sample of coffee as a whole can be more sensitive to differences in level of caffeine than rating specifically how bitter it is. This criterion of gustatory performance is independent of the particular scores given by an assessor or the subjective experiencing of any sensations. Hence perception of the tastant in a familiar context is capable of picking out supertasters and non-tasters from the tasters.

Key words: caffeine taste, discriminated levels, taste preference, bitterness,
coffee solids taste

Introduction

This paper reports new indications that sensitivities to differences in level of the taste of caffeine fall into three categories. The evidence comes from a highly economical way of estimating an individual's discriminative performance and ideal point for caffeine in a familiar drink. The least and most differentially sensitive people were both minorities that may correspond respectively to non-tasters of caffeine and to supertasters.

Individual differences in gustatory sensitivity to caffeine

In an early study, the distribution of individuals' sensitivities to the taste of barely detectable concentrations of caffeine in water showed at least two peaks, with 'tasters' of phenylthiourea predominating in the peak at lower concentrations of caffeine (Hall, Bartoshuk, Cain & Stevens 1975). Those observations supported the idea that gustatory receptors for caffeine varied among people similarly to the receptors activated by phenylthiourea and 6-n-propylthiouracil (PROP). However this similarity is not attributable primarily to stimulation of the same receptors among the range of TAS2R bitter-responsive types on the human tongue (Behrens, Foerster, Staehler *et al*. 2007). The genetic variation in sensitivity to PROP has negligible overlap with the genetic variation for caffeine (Hansen, Reed, Wright *et al*. 2006).

Among potential factors covarying with sensitivity to caffeine, a substantial minority appeared to be highly sensitive to PROP and were dubbed "super-tasters" (Bartoshuk 1991). PROP tasters are more likely than PROP non-tasters to reject higher levels of caffeine in coffee (Lee, Prescott & Kim 2008). Lee *et al*. (2008) found fewer clear differences between super-tasters and other tasters of PROP but necessarily the comparisons were between smaller numbers than those between all the tasters and the non-tasters and so had lower power.

Comparisons between studies are complicated by the statistical fact that the proportions of non-tasters and tasters and of ordinary tasters and supertasters are liable to vary across modest-sized samples from a population, over and above differences in prevalence between populations (e.g. Tepper, Koelliker, Zhao *et al*. 2008). Also, studies vary considerably in the tests used. Nevertheless, reports agree in assigning the largest category to tasters and the smallest to supertasters (e.g. Ditcschen & Guinard 2004; Chang *et al*. 2006). The division between non-tasters and tasters is relatively arbitrary, even when a detection task is used, but the conclusion has always been that there are fewer non-tasters.

PROP super-tasters are highly sensitive to other classes of tastant (Miller & Reedy 1990). Hence they are also likely to be distinctive in some general factors in gustatory function. One such is greater numbers of taste buds (Bartoshuk, Duffy & Miller 1994) although this factor is not under the same genetic control as hypersensitivity to PROP (Hayes, Bartoshuk, Kidd & Duffy 2008). Coffee solids contain tastants that are sour such as 4-vinylcatechol oligomers and quinic acid lactones generated during roasting of the beans, as well as many bitter compounds other than caffeine (Frank, Bulmberg, Kunert *et al*. 2007; Frank, Zehentbauer & Hofmann 2006). Hence the taste of a low concentration of caffeine is liable to be masked by both the bitter taste of the roasted coffee and the sour taste that may be confused with it. It is difficult to measure detection performance ('absolute threshold') for one compound in the presence of confusable compounds, in order to distinguish between supertasters and ordinary tasters of caffeine in one of its most common uses. A 'suprathreshold' measure is required. The most objective criterion would be fineness of discriminating (differential acuity) between concentrations of caffeine. In addition, it is conceivable that gustatory sensitivity to caffeine is inversely related to the most preferred concentration in a familiar drink such as coffee, because more sensitive users need less caffeine to bring a mild coffee up to a widely desired strength. The findings presented in this paper support the latter approach.

Normed sensory discrimination

An individual's discriminative performance during a session of tests on a familiar medium is at least as good with a personal internal standard as with an experimenter-provided external standard (McBride & Booth 1986). That was repeatedly demonstrated for discrimination between the level of sucrose or sodium chloride in each sample and the assessor's ideal level in the drink or food (Conner *et al*. 1986, 1987, 1988a,b). Indeed a considerable amount of evidence now indicates that putatively absolute judgments are in fact relative to an internal standard constructed from the series of stimuli thus far tested (Stewart, Brown & Chater 2005). Implicit memories of previously experienced stimulation of combinations of receptors have been invoked to explain olfactory perception (Stevenson & Boakes, 2003). The theory of discrimination between each test sample and a personal template relies on that pattern-recognition mechanism within any sensory modality and also across the senses (Booth & Freeman 1993; Booth, Kendal-Reed & Freeman 2010; Booth, Sharpe, Freeman & Conner, in press; Mobini, Platts & Booth 2011). In this paper, the theory is applied to the mixture of tastants and other sensory stimuli in the assessor's usual drink of caffeine-containing coffee.

A major advantage of judgments relative to an internal standard (a 'norm') is that far fewer data are needed to estimate discriminative performance than in traditional methods using a constant external stimulus for comparison (Booth, Thompson & Shahedian 1983; Conner, Land & Booth 1986). Furthermore, the same data provide an estimate also of the level of the discriminated stimulus in the norm (the norm point). The norm used as an anchor for quantitative judgments can be the familiar version of the material (McBride & Booth 1986; Booth, Earl & Mobini 2003; Booth, Mobini, Earl & Wainwright 2003) or the personally most preferred version of it -- the ideal point (Booth *et al*. 1983; Conner & Booth 1988; Conner *et al*. 1986, 1987, 1988a,b). A key measure in this paper is the range of one discrimination unit on either side of the ideal point -- the ideal range.

The present data were previously reported only in summary form within a review paper (Booth, Conner & Gibson 1989). The original analysis showed that the ratings of preference by half of the assessors had better differential acuity for caffeine concentrations in the assessor's usual coffee than the ratings of how "bitter" each sample tasted (Booth *et al*. 1989). The estimated ideal points were not presented in the earlier paper, as they are here.

This paper presents the two basic sorts of psychophysical function available from norm-anchored ratings of preference and intensity. Both the preference norm and the intensity norm showed evidence of three ranges of gustatory sensitivity to differences in level of caffeine. If the lowest observed range of preferred caffeine concentrations encompasses supertasters, the result casts in question the assumption that choices among types of food and drink are directly determined by sensitivities to tastes prior to experience of eating and drinking those materials. Even if coffee tastes extremely bitter to a supertaster, s/he may still come to like the drink, albeit perhaps mildly roasted, in a weak infusion, with added sugar and/or milk or cream, and/or decaffeinated. These effects on the most preferred concentration derived from a function of preference on caffeine level in a food might not show in food preference scores treated in other ways (Hayes *et al*. 2008; Hayes, Wallace, Knopik *et al*. 2011).

Materials and methods

Participants

The assessors were 17 women and 9 men between the ages of 50 and 85 years, and 15 women and 11 men between 20 and 45. Both genders and the two age ranges were recruited in order to test for effects of biological age or generational culture, and any interaction with gender. Each volunteer reported usually drinking at least 5 cups of caffeine-containing instant coffee a week. All 52 assessors were included in analyses of age, gender and composition of usual coffee but five who usually drank ground coffee were excluded from the analyses of cognitive processing in case such experience affected details of performance on the samples of instant coffee tested.

Materials

Unmarked cups were pre-prepared with a standard amount of decaffeinated instant coffee powder and caffeine at one of approximately equal-ratio (1.5 to 2) steps (5, 10, 20, 35, 50, 80, 120, 180, 270 or 405 mg per cup). Each sample to be tasted was made in the presence of the assessor by adding hot water and the assessor's usual amounts of milk and/or sugar or sweetener.

Procedure

Participants were asked to take one or two good sips from each cup of coffee and to rate numerically their preference for that coffee for a whole cupful, from zero meaning "I'd never choose this brand" to nine meaning "I'd always choose this brand." Samples were selected from the pre-prepared set in a way that minimised biases on each assessor's responses. The first sample contained approximately the amount of caffeine that was usual in instant coffee at that time (35 or 50 mg per cup). The second sample was two or three steps below or above the first, and the third sample one or more steps in the other direction, to an extent projected to be rated by that individual at least as far from the most preferred level of caffeine as the second sample had been in the opposite direction: this algorithm was intended to minimise range bias (Conner *et al*. 1987), while also as far as possible pre-empting centring bias (Poulton 1979). Two or three more samples were presented at other levels of caffeine, in an effort to even up the ratings above and below ideal ("always choose") for that assessor while filling any large gaps, thus reducing frequency biases as well as range bias (Parducci 1963).

This set of five or six drinks was then presented once more as freshly made samples, usually in reverse order. This time, however, after again rating degree of preference, the assessor judged how bitter that sample was, from a score of one for "not at all bitter" to nine for "as bitter as my usual coffee" (mid-point of the response format) and appropriately higher scores for greater bitterness. This approach resolved a number of deficiencies in both category scaling and magnitude estimation (Booth 2009; Booth *et al*. 1983).

Measurement of normed discrimination

In our earlier brief report based on these data (Booth *et al*. 1989) discriminative performance and ideal point were calculated by 'unfolding' (Coombs 1964) the assessor's ratings of preference at her/his ideal point and estimating the theoretical function by least squares linear regression (Booth *et al*. 1989; Conner *et al*. 1988a,b). The level of caffeine interpolated at the "always choose" (ideal) rating was the norm point (NP). The function's Weber fraction was the measure of discriminative performance by the judgments of preference. (That differential acuity can also be regarded as the assessor's intolerance of deviations from her/his ideal NP.) Since the Weber fraction is the fractional increase in concentration that has a 50% chance of being discriminated, it is termed here the half-discriminated fraction (HDF).

The theoretical isosceles triangle of preference (or of similarity to usual) cannot be realised when the sensory factor is varied in a less than ideal context (Booth & Shepherd 1988; Booth & Freeman 1993). The peak is truncated, creating a discontinuity in the middle region of the unfolded line (e.g. Figure 5.3 in Booth 1994). (Since the regression line is estimated using squares of deviates, that broadening of the centre of the scatterplot has little effect, unlike observed values nearer the extremes.) The context of the varied factor can be regarded as an integrated stimulus on an orthogonal dimension, creating a cone. Hence the theoretical formula for the rounded peak is a vertical cut through a cone away from its apex. That is, a defective context will turn the isosceles triangle on the varied stimulus into a conic section, i.e. a hyperbola (Booth & Freeman 1973). Hence in this report the data have been fitted to a hyperbolic function, with the tangents at the asymptotes forming the isosceles triangle for an ideal context (Figure 1).

--- *Figure 1 about here* ---

A fully algorithmic analysis has recently been programmed (Java applet Co-Pro 2.29), including calculation of cognitive interactions between multiple discrimination functions. Those purely psychological aspects of this set of data are presented elsewhere (Booth, Sharpe & Conner 2011). First, the above calculations from the raw data for each session for an individual participant generate the Weber fraction and the NP for each psychophysical function. These characteristics of assessors in the experiment reported here were compared within or between groups by analysis of ranks or variance as appropriate, using the statistical software package SPSS 16.0.

Results

Better integrative than analytical discrimination

The median half-discriminated fraction (HDF) was numerically lower (better) for one or both sets of integrative judgments of preference than it was for the analytical judgments of bitterness in each of the groups of coffee type, age and gender. This pervasive difference between response conditions was attributable to 34 out the 52 assessors having initial preference HDFs lower than their bitterness HDFs, and 28/52 for the second set of preference ratings. That is, somewhat over half of these assessors perceived differences in levels of caffeine between the samples of their coffee drink without being able to discriminate those differences using the word that normally expresses awareness of the taste of caffeine. These individuals may well have thought about bitterness before the term was introduced by the experimenter, if only because it is common experience that the taste of coffee includes bitterness, even if only very mildly. Yet they did not distinguish disparities in intensity of bitterness coming from caffeine in this context as effectively as they discriminated caffeine levels through variations among the samples of coffee in their liking for each. In other words, when the bitterness HDFs were poorer, there was gustatory perception of differences in level of caffeine by judgments of preference that was subconscious with respect to differences in bitterness (Booth *et al*. 2011). Therefore those with that pattern of performance were categorised as "discriminative subceivers" of the bitterness of caffeine. Those having a lower HDF for bitterness than that for preference, on the other hand, were called "perceivers."
The proportions did not vary markedly with age range, gender or the omission or addition of milk or sweetening to the coffees, with the exception of those adding milk without sweetening who predominantly discriminated better by their preferences than their bitternesses (Figure 2).

--- *Figure 2 about here* ---

The direction of the difference in caffeine discrimination between the integrative and analytical responses was not associated with the proportion of variance in either sort of judgment by an individual that was accounted for by levels of caffeine. About half of both discriminative subceivers and perceivers (i.e., non-subceivers) had a hyperbolic psychophysical function (i.e., relative to norm) with $r^2 < 0.1$, with most of the others in each group having r^2 values in the range 0.47 to 0.81.

Differential acuity of integrative preference and analytical bitter intensity

The best-discriminating 13 of the 52 assessors (up to the lower quartile) in each of the three sets of judgments had half-discriminated (Weber) fractions (HDFs) for caffeine ranging from 0.12 to 0.43 for the first ratings of how often the assessor would choose (prefer) the sipped sample as a drink of coffee, from 0.19 to 0.60 for the second set of these ratings of preference and from 0.14 to 0.57 for ratings of how bitter each sample tasted alongside that second judgment of its acceptability. Considering that these values come from aqueous mixtures of caffeine with many other strongly tasting compounds, such ranges of best discriminative performance compare well with the best HDFs reported for taste intensity ratings of single tastants in pure solutions (around 0.1-0.2: see Conner *et al*. 1988a,b; also compare McBride 1983 for pure solutions with McBride & Booth 1986 for a flavoured mixture of tastants).

Nevertheless, many participants performed poorly at discriminating the level of caffeine in each sample from the level that each implicitly remembered from her or his usual coffee. The median HDFs for the first and second sets of preference ratings were 0.98 and 1.37, and 1.54 for rated intensity of bitterness. These values could be considered to be evidence for real though poor discriminations between the test stimuli and the internal standard, but such ‘just noticeable’ ratios of 2 - 2.5 are dismal compared to the JND values reported for pure solutions. The upper quartile of HDF was 5.2 and 5.3 for first and second acceptances, and 8.1 for bitterness. Thus it seems that the variations in gustatory receptor occupation across concentrations of caffeine had strong competition from compounds produced by the roasting of coffee beans that were at a constant level across the samples for a single assessor (Frank, Zehentbauer & Hofmann 2006). This interpretation is consistent with the likelihood that caffeine and any bitter compound stimulates a profile of the hTAS2R family of receptors, providing a major peripheral contribution to masking, whatever central effects also exist.

Types of coffee

There were three main types of use of the instant coffee in the sample repoprted here -- without any addition (black) or with milk (white), with or without sweetening. As many as 17 (81%) out of the 21 frequent drinkers of unsweetened white coffee had more acute HDFs from preference than from bitterness, reliably above the median prevalence across the three types of coffee, $\chi^2(2) = 6.03$, $P < 0.05$ (Figure 2). In contrast, only 8 (44%) out of 18 exclusive users of instant coffee with milk and some sugar or low-calorie sweetener met this criterion for discriminative subception. (Four out of the 8 drinkers of plain black instant met the criterion but this group was too small for this 50% to distinguish between 40% and 80% or even 100%.) This difference between sweetened and unsweetened coffee drinkers indicates that successful discrimination between levels of caffeine by the integrative judgment is disrupted by the presence of ligands for receptors sensitive to sucrose or the intense sweeteners used (aspartame or saccharin). Some physical process in taste pores on the tongue or even at the receptors for caffeine or coffee roast compounds cannot be excluded. The fat and protein of milk might reduce access to gustatory receptors by particulate-bound or high molecular weight solutes more than to caffeine free in solution (cp. Booth 2005). Nevertheless, central mechanisms for the effect of sweetening are likely as well or instead, such as interactions between afferent fibre patterns for different receptor types and/or between the conceptualisations or phenomenologies of sweetness and the bitterness of caffeine and of roasted coffee. In short, sweeteners might counter all bittering centrally, including that by caffeine.

Age and gender differences in effects of rating bitterness of coffee

The present experiment was originally designed to sample older and younger adult age ranges as well as both genders (Booth *et al*. 1989). Hence the change in performance of preference judgments after the investigator had introduced the word “bitter” could be investigated separately in women and in younger people, who both may respond more strongly than the others to some bitter tastants (Bartoshuk *et al*. 1994; Cowart, Yokomukai & Beauchamp 1994). In a larger study, an effect associated with age or gender in these data might prove to have been confounded by differences in prevalence of use of coffee with or without milk or sugar or other sweetener that might mask the taste of caffeine (Keast 2008). Nevertheless, the two age groups and the genders were both fairly evenly spread among users of the three types of coffee reported and no frequency effect approached significance. Hence such confounding of these data is unlikely.

In the second run, assessors under 50 years of age differed reliably from the older assessors in the distribution of HDFs for acceptance (Wald-Walkowitz $Z = -1.77$, $P < 0.04$) because of less diversity in the older group ($P < 0.03$ in the Moses test with one outlier trimmed from each extreme). This wider range of discriminative performance among younger assessors may be attributable to them becoming sensitive to differences in the taste of caffeine after the explicit introduction of the concept of ‘bitter.’ Such an effect was indicated by a lower norm point (implied usual level of caffeine) for the second acceptance than any in the older group. However, the impact on the distribution was not enough for the numerical difference between the two groups’ medians to be reliable (Mann-Whitney $U = 209$, $P < 0.16$).

The second run’s ratings of bitterness (relative to the assessor’s usual coffee) brought out a difference between the genders also in amount of caffeine liked in coffee tasted by a sip. The median norm point was higher in women than in men, Mann-Whitney U = 140, $P < 0.008$, as was the women's norm point that differed most from that in the men, Kolmorogov-Smirnov Z = 1.5, $P < 0.025$. This effect did not necessarily arise from women overall liking their coffee more bitter than men do: for example, on average women might use smaller amounts of coffee solids in a cup and so need more caffeine to match the same overall bitterness preferred by men and women.

A greater mean rank of norm points for second acceptance among older people approached significance (Mann-Whitney U = 234, $P < 0.063$), while norm points for bitterness had lower mean rank in the older group, though far from reliably (U = 269, $P < 0.22$). Yet this pattern could reflect greater popularity of black coffee among the young, rather than worse discrimination among older people. For example, these older people may have preferred what they consciously regarded as mild coffee (especially the women: bitterness norms versus the older men, U = 159, $P < 0.003$), while nevertheless liking higher levels of the 'sharp' or 'bright' bitterness of caffeine [age by coffee type interaction $F(2, 6.3) = 7.3$, $P < 0.025$], as distinct from the 'duller'-tasting flavours of roasted coffee.

However, the above analyses had highly significant inhomogeneities of variance. Analyses of ranked discrimination HDFs and norm points showed no main effects of type of coffee, age group or gender, and none of the interactions apparent in ANOVA approached significance in analyses by ranks or exact probabilities. The prevalences of norm point sensitivities are therefore considered across all the participants in the experiment.

Bimodality of caffeine norm points for bitterness

Variations among assessors in the level of caffeine for usual bitterness indicated effects of peripheral and/or central neural interactions among receptor activations by caffeine and compounds in roasted coffee. Individuals' values of the norm point (NP) for caffeine from ratings of bitterness showed clear signs of a bimodal distribution (Figure 3, middle panel). The trough in incidences between lower and higher concentration NPs was not so deep in the preference NPs at first tasting (Figure 3, leftmost panel). Curiously, it was even less evident in preference after bitterness began to be rated (Figure 3, rightmost panel).

--- *Figure 3 about here* ---

A more complex profile of preference NPs became fully evident when bitterness was also rated (Figure 3, rightmost panel). Simply bimodality for bitterness of caffeine has been reported in descending-staircase thresholds (Hall *et al*. 1975), with the modes spanning the concentrations of caffeine commonly occurring in coffee drinks. In that study and more recently (Lee *et al*. 2008), the bimodality for caffeine correlated with the widely seen bimodality of such values for 6-n-propylthiouracil (PROP). However, the receptors for bitter tastants in the hTAS2R gene family (Behrens *et al*. 2007) which show greatest sensitivity to caffeine are different from those to PROP (Delwiche, Buletic & Breslin 2002) and the correlated bimodalities may come from a supersensitivity having yet a different genetic basis (Hansen *et al*. 2006). Hence that extreme sensitivity to the taste of PROP (and to many other tastants) can sometimes be separated from medium sensitivity among PROP tasters, creating three categories of sensitivity, with non-tasters and supertasters both being minorities of 25-30% (Bartoshuk *et al*. 1994; Tepper & Ulrich 2002).

In the present data, the major mode of NPs was at lower levels of caffeine than the minor mode at higher concentrations (Figure 3). (The peaks of prevalence of NPs at the lower caffeine levels also tended to be sharper, i.e. their HDFs showed finer differential acuity.) This main mode may combine medium tasters with supertasters since it showed signs of splitting off a third category of even lower NPs, to varying degrees in all three panels of Figure 3. The point to note for present purposes is that the ratings of bitterness yielded only a tail of very low NPs (Figure 3, middle panel), whereas preference gave signs of a third peak at the lowest NPs (with finer HDFs too) -- most clearly after bitterness had begun to be rated (Figure 3, rightmost panel). It appears that explicit introduction of the term “bitter” focused the attention of the most taste-sensitive assessors on the differences in bitterness between samples, in fact coming only from caffeine.

There was one further indication that bitterness and preference differed in the modulation of processing of the gustatory stimulation from caffeine by stimulation from other tastants in coffee. The main mode of the NPs was at higher levels of caffeine for preference on both occasions than it was for bitterness (Figure 3; note the variation across the three graphs in the range of concentrations of caffeine on the x-axis). This is consistent with the taste of caffeine being augmented by the taste of other compounds in ratings of bitterness, while differences in levels of caffeine had stronger control of ratings of preference. That is, bitterness was more susceptible than preference to masking of the gustatory stimulation from caffeine, opening the possibility of subception of differences in level of caffeine (Booth et al., 2011a,b).

Discussion

It might be asked of the present approach, "why attempt to do psychophysics in such a complex system?" Caffeine in water is physically much simpler than caffeine in coffee and so might be thought to be simpler to investigate perceptually. However, purity comes with artificiality, whereas familiarisation with a recurring complexity creates a 'template' or 'norm' in memory which can be used to anchor behavioural or analytical ratings. To put the point another way, if linear responding is acquired to stable affordances (Gibson, 1977), it could provide a powerful tool for dissecting out the processes operative during action on the complex reality. This configural view implies that levels of caffeine could be discriminated better in a familiar mixture using an internal norm than in pure solution without one (Booth, 1995). Indeed, in visual psychophysics, an internal standard performs better than the traditional external standard (Morgan *et al*. 2002; Nachmias 2006). In chemosensory perception, comparisons using an internal ideal point were as discriminating as intensity with an external standard across groups assessing concentrations of a familiar breakfast drink (McBride & Booth 1986).

Nevertheless, within any group, assessors vary greatly. More than half the participants in this experiment discriminated between levels of caffeine better by ratings of how much they liked each sample of coffee (choice preference) than they did by ratings of how bitter the samples were (analytical description). Since both responses measure the differential sensitivity to caffeine in the same complex medium, how can they differ? The answer must come from what happens after the purely sensory processing of the caffeine and other tastants in the drink.

Differences in the bitterness of caffeine are liable to be masked by the sensing of substances generated in the roasting of coffee. These might be clearly bitter but astringent or sour can also be termed bitter (O'Mahony *et al*. 1979). Hence, to distinguish between levels of caffeine in such a context requires the assessors to have developed rather refined 'analytical' strategies. These might be based on extensively varied relevant sensory experience – not necessarily conceptually formulated into descriptive analysis. They might arise from much discussion of relevant situations, such as choices among coffee houses or types of coffee, with little attention to specific sensed material characteristics as distinct from broad concepts such as strength of brew or extent of roasting. Commercial approaches to sensory descriptive analysis are meant to bring together the sensing and the describing but, as far as we know, none of our assessors had been trained this way. (Some might say that life even had not trained them in appreciation of coffee, else they would not be drinking instant!)

In particular, it is unlikely that the assessors had previously faced the task of comparing drinks of coffee that differed in contents of caffeine. A few might have tried a decaffeinated brand but that is likely to have had other differences from their usual caffeinated brand, such as in blend of beans and even residual aroma if at that time an organic solvent was used to extract the caffeine. Hence, those assessors whose gustatory sensitivity to caffeine was sufficient for them to sense some difference between the tested levels of caffeine are very likely to have assimilated the difference to a generic concept of overall strength of flavour or even perhaps to unspecified differences in overall quality of the brand, as might relate to method of drying the coffee extract and to retail price.

In that case, it would not be surprising if some assessors had no thought of the strength of taste when judging the likelihood of their choosing the "brand" that each test drink was presented as. However, if the taste as such of coffee was in mind, then the term "bitter" was likely to be easily accessible. Nevertheless, the term itself is not readily brought into use during spontaneous assessment of liking for a coffee drink because, even though the personally preferred strength of coffee is largely some extent of the drink's characteristic bitterness (even if very low), bitterness is often a complaint rather than a merit.

In short, the questions about the present approach could be formulated as the fitness for purpose of sensual versus conceptual appreciation of coffee. How bitter the coffee is may do better in a sensual mood. The coffee I like overall may do better when I am in a conceptualising mode. Maybe those participants who had a precise norm for their whole concept of their coffee are those who discriminated between levels of caffeine better by preference for the coffee than by description of its bitterness.

At the very least, this report illustrates an improvement on the usual approach to investigating the role of the genetics of taste or of family environment in individual differences in likings for foods. Assessing only the strengths of preferences for foods is a poor way of understanding the role of PROP bitterness in food choices. The levels of the tastant of interest should be varied in each food of interest and the preference scores related to those concentrations using the learnt optimum as an internal standard. When the design is theoretically appropriate, only a few samples are required to test each person on each food (Conner *et al*. 1986; Booth *et al*. 2010). In the present experiment, testing of just five samples enabled the measurement of each ideal point and discrimination unit. Those data were sufficient to yield the grouped profiles reported here and to investigate the mental processes producing the apparently unconscious tasting of disparities in caffeine concentrations among those in whom preference was more sensitive than bitterness (Booth *et al*., 2011b).

Acknowledgments

The collection of data was supported by an SERC CASE with Unilever Colworth. The new analyses reported here were partly supported by two grants to Prof. Booth from the Agri-Food Committee of the Biotechnology and Biological Sciences Research Council (UK).

There is no conflict of interest in the conduct of this research or in this publication

References

Bartoshuk LM (1991) Sweetness: history, preference and genetic variability. Food Technol 45(11):108, 110, 112-113.

Bartoshuk LM, Duffy VB, Miller IJ (1994) PTC/PROP tasting: anatomy, psychophysics and sex effects. Physiol Behav 56:1165-1171.

Behrens M, Foerster S, Staehler F, Raguse JD, Meyerhof W (2007) Gustatory expression pattern of the human TAS2R bitter receptor gene family reveals a heterogenous population of bitter responsive taste receptor cells. J Neurosci 27:12630-12640.

Booth DA (1994) Psychology of nutrition. Psychology Press, Hove UK

Booth DA (2005) Perceiving the texture of a food: biomechanical and cognitive mechanisms and their measurement. In: E. Dickinson (ed) Food colloids: interactions, microstructure and processing. Royal Society of Chemistry, Cambridge, pp. 339-355.

Booth DA (2009) The basics of quantitative judgment. How to rate the strength of appetite for food and its sating. Appetite 53:438-441.

Booth DA, Freeman RPJ (1993) Discriminative feature integration by individuals. Acta Psychol 84:1-16.

Booth DA, Shepherd R (1988) Sensory influences on food acceptance - the neglected approach to nutrition promotion. BNF Nutr Bull 13(1):39-54.

Booth DA, Conner MT, Gibson EL (1989) Measurement of food perception, food preference, and nutrient selection. Ann NY Acad Sci 561:226-242.

Booth, D.A., Earl, T., & Mobini, S. (2003). Perceptual channels for the texture of a food. Appetite 40:69-76.

Booth DA, Higgs S, Schneider J, Klinkenberg I (2010) Learned liking versus inborn delight. Can sweetness give sensual pleasure or is it just motivating? Psychol Sci 21:1656-1663.

Booth DA, Kendal-Reed MS, Freeman RPJ (2010) A strawberry by any other name would smell as sweet, green, fruity and buttery. Multisensory cognition of a food aroma. Appetite 55:738-741.

Booth DA, Mobini S, Earl T, Wainwright CJ (2003) Consumer-specified instrumental quality of short-dough cookie texture using penetrometry and break force. J Food Sci 68(1):382-387.

Booth DA, Sharpe O, Conner MT (2011a) Discrimination without description. Are the differences conceptualised or fully subconscious? http://epapers.bham.ac.uk Booth (downloaded on 1 February 2011)

Booth DA, Sharpe O, Freeman RPJ, Conner MT (2011b) Insight into sight, touch, taste and smell by multiple discriminations from norm. Seeing Perceiv in press.

Booth DA, Thompson AL, Shahedian B (1983) A robust, brief measure of an individual's most preferred level of salt in an ordinary foodstuff. Appetite 4:301-312.

Chang WI, Chung JW, KimYK, Chung SC, Kho HS (2006) The relationship between phenylthiocarbimide (PTC) and 6-n-propylthiouracil (PROP) taster status and taste thresholds fo sucrose and quinine. Arch Oral Biol 51:427-432.

Conner MT, Booth DA (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 MT, Haddon AV, Booth DA (1986) Very rapid, precise measurement of effects of constituent variation on product acceptability: consumer sweetness preferences in a lime drink. Lebens-Wiss u -Technol 19:486-490.

Conner MT, Land DG, Booth DA (1987) Effects of stimulus range on judgments of sweetness intensity in a lime drink. Br J Psychol 78:357-364.

Conner MT, Booth DA, Clifton VJ, Griffiths RP (1988a) Individualized optimization of the salt content of white bread for acceptability. J Food Sci 53:549-554.

Conner MT, Haddon AV, Pickering ES, Booth DA (1988b) Sweet tooth demonstrated: individual differences in preference for both sweet foods and foods highly sweetened. J Appl Psychol 73:275-280.

Coombs CH (1964) A theory of data. Wiley, New York.

Cowart BJ, Yokomukai Y, Beauchamp GJ (1994). Bitter taste in aging -- compound-specific decline in sensitivity. Physiol Behav 56:1237-1241.

Delwiche JF, Buletic Z, Breslin PAS (2002) Clustering bitter compounds via individual sensitivity differences: evidence supporting multiple receptor-transduction mechanisms. In: Given P, Paredes D (eds) Chemistry of taste: mechanisms, behaviors and mimics. American Chemical Society, Washington DC, pp. 65-77.

Ditschun TL, Guinard J-L (2004) Comparison of new and existing methods for the classification of individuals according to 6-n-propylthiouracil (PROP) taster status. J Sens Stud 19:49-170.

Frank O, Blumberg S, Kunert C, Zehentbauer G, Hofmann T (2007) Structure determination and sensory analysis of bitter-tasting 4-vinylcatechol oligomers and their identification in roasted coffee by means of LC-MS/MS. J Agric Food Chem 55:1945-1954

Frank O, Zehentbauer G, Hofmann T (2006) Bioresponse-guided decomposition of roast coffee beverage and identification of key bitter taste compounds. Eur Food Res Technol 222:492-508.

Gibson JJ (1979) Ecological approach to visual perception. Houghton Mifflin, New York.

Hall MJ, Bartoshuk LM, Cain WS, Stevens JC (1975) PTC taste blindness and the taste of caffeine. Nature 253:442-443.

Hansen JL, Reed DR, Wright MJ, Martin NG, Breslin PAS (2006) Heritability and genetic covariation of sensitivity to PROP, SOA, quinine, HCl and caffeine. Chem Senses 31:403-413.

Hayes JE, Bartoshuk LM, Kidd JR, DuffyVB (2008) Supertasting and PROP bitterness depends [*sic*] on more than the TAS2R38 gene. Chem Senses 33:255-265.

Hayes JE, Wallace MR, Knopik VS, Herbstman DN, Bartoshuk LM, Duffy VB (2011) Allelic variation in TAS2R bitter receptor genes associates with variation in sensations from and ingestive behaviors toward common bitter beverages in adults. Chem Senses 36:311-319.

Keast, R.S.J. (2008). Modification of the bitterness of caffeine. Food Qual Pref 19:465-472.

Lee YM, Prescott J, Kim KO (2008) PROP taster status and the rejection of foods with added tastants. Food Sci Biotechnol 17:1066-1073.

McBride RL (1983) A JND-scale category-scale convergence in taste. Perc Psychophys 34:77-83.

McBride RL, Booth DA (1986) Using classical psychophysics to determine ideal flavour intensity. J Food Technol 21:775-780.

Miller IJ, Reedy FE (1990) Quantification of fungiform papillae and taste pores in living human subjects. Chem Senses 15:281-294.

Mobini S, Platts RG, Booth DA (2011) Haptic signals of texture while eating a food. Multisensory cognition as interacting discriminations from norm. Appetite 56:386-393.

Morgan MJ, Watamaniuk SN, McKee SP (2000) The use of an implicit standard for measuring discrimination thresholds. Vision Res 40:2341-2349.

Nachmias J (2006) The role of virtual standards in visual discrimination. Vision Res 46:2456-2464.

O'Mahony M, Goldenberg M, Stedmon J, Alford J. (1979) Confusion in the use of the taste adjectives 'sour' and 'bitter.' Chem Senses 4:301-318.

Parducci A (1963) Range-frequency compromise in judgment. Psychol Monog 77:1-50.

Poulton EC (1979) Models for biases in judging sensory magnitude. Psychol Bull 86:777-803.

Stevenson RJ, Boakes RA (2003) A mnemonic theory of odor perception. Psychol Rev 110:340-364.

Stewart N, Brown GDA, Chater N (2005) Absolute identification by relative judgment. Psychol Rev 112:881-911.

Tepper BJ, Ulrich NV (2002). Influence of genetic taste sensitivity to 6-n-propylthiouracil (PROP), dietary restraint and disinhibition on body mass index in middle-aged women. Physiol Behav 75:305-312.

Tepper BJ, Koelliker Y, Zhao L, Ullrich NV, Lanzara C, d'Adamo P, Ferrara A, Ulivi S, Esposito L, Gasparini P (2008) Variation in the bitter-taste receptor gene TAS2R38, and adiposity in a genetically isolated population in Southern Italy. Obesity 16:2289-2295.

Captions to Figures

Figure 1 Normed psychophysical functions from concentration of caffeine to rated acceptance before and after mention of bitterness and to rated bitterness in assessors 24 and 46, who drank coffee daily without sugar (46 without milk either). Each data point gives the sequence of coffee samples in the first run as the first numeral and the second run as the second. Continuous line: conic section (hyperbolic function) fitted to the data by least squares. Broken lines: tangents to the hyperbola intersecting (isosceles triangle) at the norm used in those quantitative judgments. The HDDs were (for assessor 24) 1.17, 1.39 and 1.48 and (for 46) 1.20, 1.88 and 5.29, i.e. Weber fractions of 17 or 20%, 39 or 88% and 48 or 430%.

Figure 2 Proportion (%) of each subset of participants whose initial integration was more (left-hand bar) or less (right-hand bar) discriminative than their later analysis, split by age group, gender and the absence or presence of added material obscuring or countering the stimulus. Left side: age ranges in years (each N = 26). Middle: female (N = 32) or male (N = 20). Right side: plain coffee (unsweetened, unmilked: UU; N = 8); unsweetened coffee with milk (UM; N = 21); sweetened coffee with milk (SM; N = 18). * $P < 0.05$.

Figure 3 Frequency polygons of Norm Ranges (each assessor's Norm Point $\pm$ Half-Discriminated Disparity) of caffeine (mg/cup) shown by ratings of bitterness at the start of the second session (middle graph) and acceptance in the second and first sessions (left- and right-hand panels, respectively). For the purposes of graphic presentation only, two of the 52 assessors have been excluded from the middle panel because of extremely large HDDs in the second session. Note that the range of the horizontal axis varies among the three conditions and so the distance between -5 and +5 log mg of caffeine per cup has been approximately equated across the graphs.

Figure 1

“choose” in 1st run **“bitter” in 2nd run** **“choose” in 2nd run**

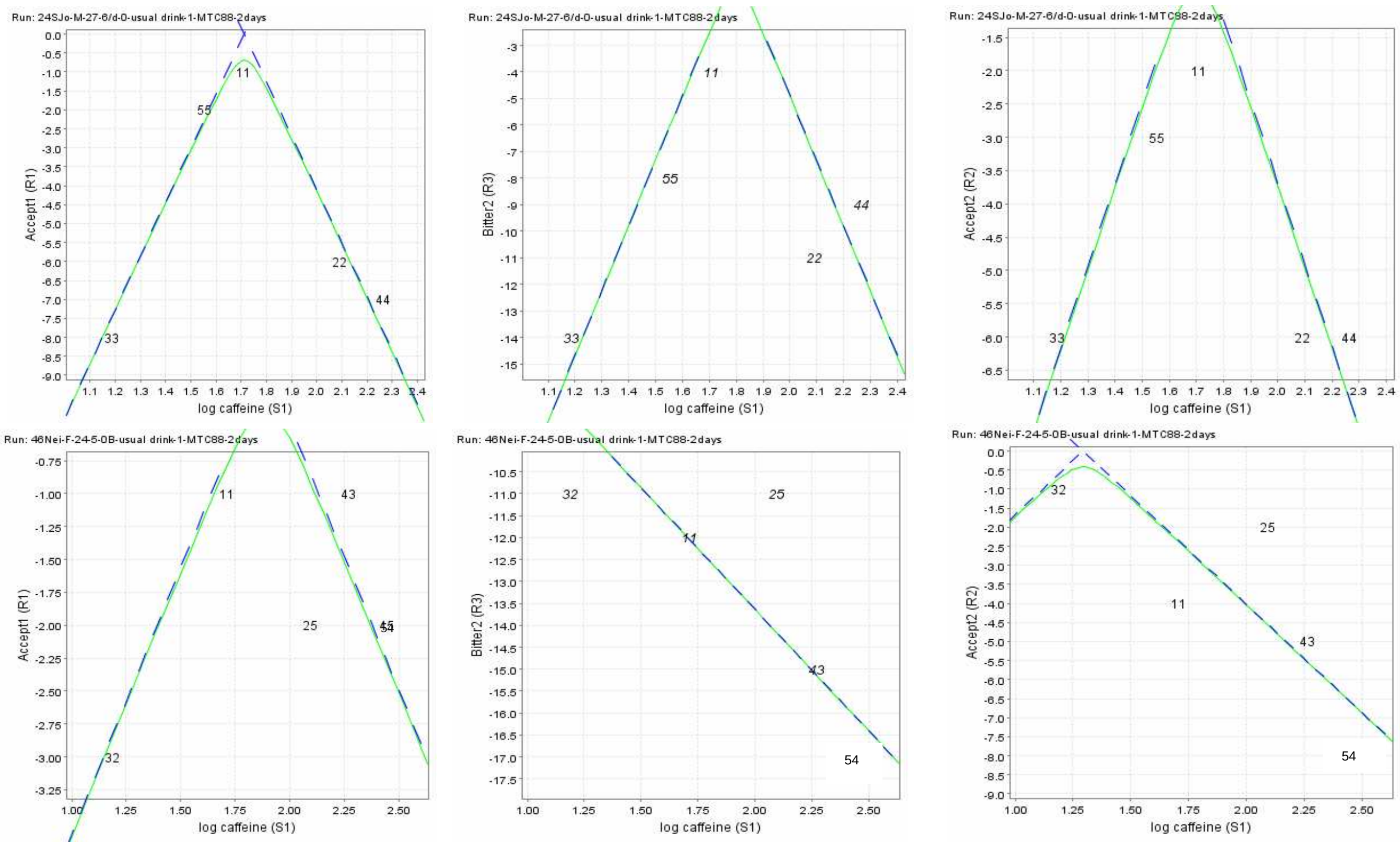


Figure 2

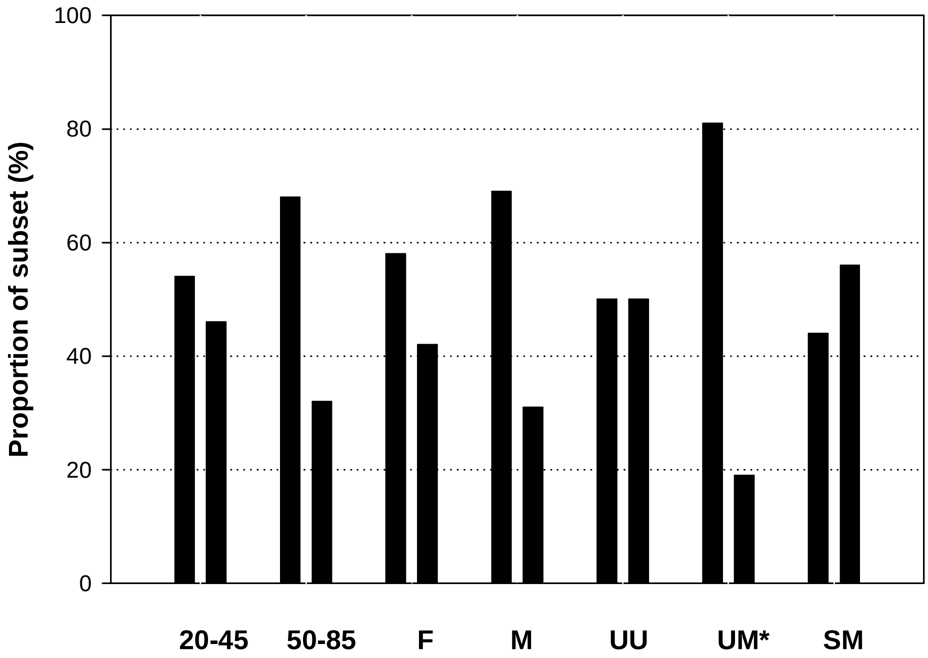
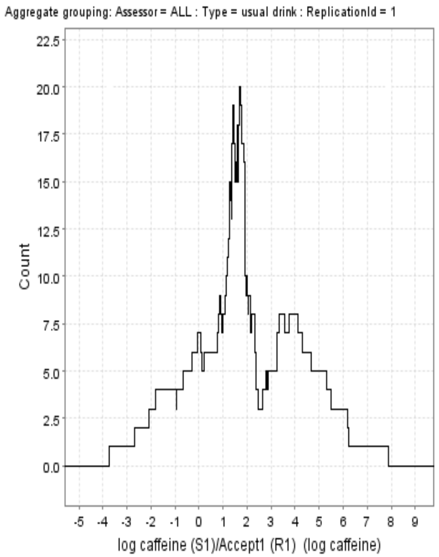
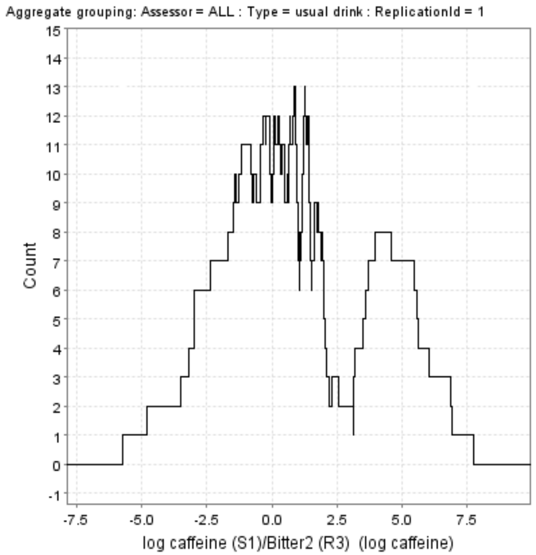


Figure 3

"choose" in 1st run



"bitter" in 2nd run



"choose" in 2nd run