Multiplicative Models for Combining Information from Several Sensory Experiments: a Bayesian Analysis

Bareng A. S. Nonyane\textsuperscript{a,\,*} and Chris M. Theobald\textsuperscript{b}

School of Mathematics, University of Edinburgh
and Biomathematics & Statistics Scotland,
The King’s Buildings, Edinburgh, Scotland EH9 3JZ, U.K.
\textsuperscript{a} aletta@schoolph.umass.edu
\textsuperscript{b} cmt@bioss.ac.uk

November 7, 2006
Abstract

If data are available from a series of quantitative sensory experiments on the same type of product, and the assessors in these experiments are drawn from a common pool, then it is possible to combine information from the experiments on the relative biases and variability of individual assessors, and to examine the influence of possible temporal effects over the series. Such a combination of information is illustrated using a series of apple-tasting experiments conducted with the main aim of monitoring assessor performance over time. Models which include random effects and multiplicative interaction terms have been used for modelling heterogeneous interaction between assessors and products in individual sensory experiments. Such models are extended here to analyse data from series of experiments. A Bayesian approach is used that allows for adjustment for missing observations and for the use of information on assessors’ previous performance when analysing future experiments. This use of previous information leads to a reduction in the average variance of product differences.

Key Words: Assessor performance; Bayesian inference; Multiplicative model; Sensory studies.
1 Introduction

Sensory experiments are widely used by the food industry for reasons such as establishing consumer preferences for different food and drink varieties on the market and in product development. Usually, series of tasting experiments are conducted with a panel of trained assessors who give their scores or rankings on food products. The performance of the assessors may be evaluated regularly over time, and one might omit from the panel those assessors whose performance is not consistent. The literature has a wide range of methods for evaluating assessor performance: see for example, Naes (1998), Rossi (2001) and McEwan, Hunter, van Gemmert and Leah (2002).

This paper considers the analysis of data from quantitative sensory studies in which the performance of assessors over a sequence of experiments can be examined. The models and methods proposed are illustrated on a series of apple-tasting experiments, which are described in Section 2. Heterogeneity of assessors’ responses to products, and of error variance between assessors, are modelled with a Bayesian multiplicative model, and the individual-specific parameters in this model are used to measure assessor performance relative to others in the panel. Multiplicative models are discussed in Section 3, and extended to series of experiments in Section 4. The apple data are analysed in Section 5. Section 6 discusses the analysis of a future experiment given information about assessor performance from previous experiments, and this is followed by a general discussion.

2 Apple-tasting data

The data studied here were obtained from the Hannah Research Institute in Ayr, Scotland. A series of apple-tasting experiments was conducted over several years with the aim of monitoring assessor performance over time. Data sets from 45 of the experiments spanning the period March 1996 to April 1998 were available for analysis. For these studies, 14 trained female assessors were involved: the number of assessors attending individual experiments varied from 9 to 14 per experiment, and the number of experiments attended from 11 to 43.

In each experiment, several apple varieties were tasted, the choice of varieties depending on the availability in the local market at the time of the experiment. Varieties were sometimes tasted in several versions, with different countries of origin or methods of packaging, and there were some
inconsistencies in recording this information. Thus for consistency we treat the combinations of
varieties and versions as unique to every experiment, and refer to them as products. This results
in between 4 and 16 products per experiment (most often 8), corresponding to between 4 and
11 varieties. The order followed for presenting products to assessors within a tasting session
was based on the rows of Williams Latin-square designs (Williams (1949), MacFie, Bratchel,
Greenhoff and Vallis (1989)). A Williams Latin square is such that every product occurs equally
often in every position, and is preceded by every other product equally often; also each assessor
receives every product the same number of times. The last of these conditions is satisfied even
when the number of assessors is not a multiple of the number of products, as in most of the
experiments considered here. The allocation was replicated 3 times for each assessor (with a few
exceptions), and the replicates were separated by break periods the lengths of which were not
reported.

Each assessor was allocated a private booth in which to do the tasting, and she entered the
scores by dragging a mouse pointer along a bar marked from 0 to 100 on a computer screen.
Each product was scored on 11 attributes: fruitiness, sweetness, acidity, bitterness, presence
of perfumed smell, floral sensation, after-taste, persistence, hardness, crunchiness and overall
acceptability of the product. We present analyses only for the sweetness scores which appear to
be roughly normally distributed: some of the other attributes had scores with high proportions
of zeros, for which an assumption of normality would not be appropriate.

A combined analysis of the sweetness attribute from the 45 experiments was carried out in order
to

• model assessor performance across experiments

• adjust mean product effects within experiments for missing assessors

• incorporate information on assessors’ use of the sweetness scale to improve precision in the
  analysis of future experiments which include the same assessors.
3 Multiplicative models for assessor performance in individual experiments

Yates and Cochran (1938) and Finlay and Wilkinson (1963) introduced multiplicative models which are widely used in agricultural trials to model heterogeneity of variety variance between growing environments. Brockhoff and Skovgaard (1994) drew an analogy between variety trials and sensory experiments in which a panel of assessors evaluate products, and proposed the use of multiplicative models for sensory studies.

Consider first unreplicated sensory data in which $y_{ij}$ denotes the attribute score of assessor $i$ for product $j$, and there is no standard measure of attribute intensity to which the scores can be compared. It is often found that there is heterogeneity of product and error variances between assessors. When that is the case, a model which takes this heterogeneity into account by including a multiplicative term may be used. This is given by

$$y_{ij} = \alpha_i + \beta_i \theta_j + \epsilon_{ij},$$

where $\alpha_i$ is the additive effect of assessor $i$, $\theta_j$ is the effect of product $j$, $\beta_i$ is a coefficient for assessor $i$ reflecting her use of the measurement scale and $\epsilon_{ij}$ is a random error term, assumed to be normally distributed with zero mean. Thus, the interaction between assessor and product is modelled by a regression on the unknown product effects with coefficients $\beta_i$. This model can allow for unequal residual variances for assessors, so that $\text{var}(\epsilon_{ij}) = \sigma_i^2$. Theobald and Mallinson (1978) considered model (1), with normally distributed $\theta_j$ and unequal residual variances, in the context of comparative calibration of measuring instruments. They noted a connection with factor analysis models having a single common factor, and emphasised precision, defined as $\pi_i = \beta_i/\sigma_i$, as a measure of the performance of instrument $i$.

The $\beta_i$ were interpreted by Brockhoff and Skovgaard (1994) as measures of assessors’ discriminating ability, and called *assessor expansiveness* by Mead and Gay (1995). If assessor $i$ is good at discriminating she will have low residual standard deviation $\sigma_i$ relative to $\beta_i$, and thus a large value of $\pi_i$.

Brockhoff and Skovgaard (1994) also suggested how this multiplicative model may be generalised to account for the whole design structure when each product is tasted more than once, that is,
in a replicated experiment. Thus, when replication effects are given index \( r \), their model has the form

\[
y_{ijr} = \alpha_i + \beta_i (\theta_j + \gamma_r) + \epsilon_{ijr},
\]

where \( \gamma_r \) is a replication effect and this, like the product effect, is realized through the assessors’ use of the scale.

Frequentist methods for fitting such multiplicative models were explored by Digby (1979), Oman (1991), Gogel, Cullis and Verbyla (1995) and Nabugoomu, Kempton and Talbot (1999).

Smith, Cullis, Brockhoff and Thompson (2003) extended model (2) to allow for more general designs and unbalanced data. Their models may include fixed effects in addition to assessor effects, and random effects other than those for products. The single multiplicative term in (1) is generalised to a sum of such terms, corresponding to a factor analysis model with several common factors. Model fitting is by REML. When the error variances \( \sigma_i^2 \) of assessors are allowed to differ, methods such as REML and maximum likelihood have the disadvantage that some variance estimates may be zero. For example, the data presented in Smith et al. (2003) lead to many zero estimates and to very dissimilar estimates of assessor expansiveness.

Our interest is in extending model (2) to apply to series of experiments on products of the same type, and then to future experiments. We seek also to avoid unstable estimates of error variances and assessor expansiveness.

4 A multiplicative model for series of experiments

Plots of mean sweetness scores over replicates versus products for four of the experiments are shown in Figure 1. The different lines do not correspond to the same assessors in all four plots because attendance was not the same for all experiments. The plots show general consistency between assessors within experiments in terms of the ordering of products on the scale from 0 to 100, but there appear to be differences in mean scores over products and in expansiveness. Within experiments, a model with multiplicative interaction effects, like (1) or (2), is therefore more suitable than an additive model. A combined model for all experiments may also be considered, with common values over experiments for some parameters, such as error variances.
We want to make inferences about product effects in future experiments, so it is necessary to treat these effects as random, as in Smith et al. (2003). We go further and adopt the Bayesian approach which considers all model parameters as randomly sampled from probability distributions. This approach is intended to

- avoid unreasonable estimates of error variances and expansiveness
- avoid the use of large-sample approximations for inferences
- permit the inclusion of expert knowledge into the analysis.

Expert knowledge of similar sensory studies is incorporated by specifying prior distributions for variance components, and for the overall expectation of the responses.

In order to extend model (1) to a multiplicative model for a series of experiments, we allow the product effects and the additive and multiplicative effects of assessors to depend on the experiment. Thus model (1) is generalized to

$$y_{ijk} = \alpha_{ik} + \beta_{ik} \theta_{j(k)} + \epsilon_{ijk}, \quad (3)$$

where $\alpha_{ik}$ is the effect of assessor $i$ in experiment $k$, $\beta_{ik}$ is assessor $i$’s expansiveness in experiment $k$, $\theta_{j(k)}$ is the effect of product $j$ within that experiment and the error terms $\epsilon_{ijk}$ are assumed to be distributed as $N(0, \sigma^2_i)$ independently. The effects $\alpha_{ik}$ for the $i$th assessor are assumed to be normally distributed about her mean effect $\alpha_i$, and similarly the $\beta_{ik}$ for any $i$ are taken to be normal with mean $\beta_i$. Product effects are nested within experiments, so for any $k$ the $\theta_{j(k)}$ are assumed to be normally distributed about a mean product effect $\phi_k$ for experiment $k$. The experiment effects are thus realized via the assessor’s expansiveness, like the replicate effects $\gamma_r$ in (2).

An alternative notation might express $\alpha_{ik}$ as the sum of $\alpha_i$ and an interaction effect $\delta_{ik}$, say, of mean zero, with similar decompositions for the $\beta_{ik}$ and $\theta_{j(k)}$. The parametrization chosen for (3) is an example of hierarchical centring (Gelfand, Sahu and Carlin (1995)), which has been shown to hasten convergence when fitting Bayesian linear models.

Table 1 shows the structure assumed for the parameters in model (3), and also includes the prior distributions. Greek and Roman letters indicate respectively unknown parameters and values.
specifying prior distributions. The subscripts A, B, P, E refer to the populations of assessor effects, assessor expansiveness, product effects and experiment effects, while \( \sim \) means ‘distributed as’ or ‘distributed independently as’, according to the context. The unknown parameters in the model are assumed to be statistically independent \textit{a priori} unless the dependence is made explicit in Table 1.

The first two entries in the top row of Table 1 express the \( \alpha_{ik} \) and \( \alpha_i \) as normally distributed random effects with expectations \( \alpha_i \) and \( \mu_A \) respectively and variance components \( \sigma_{AE}^2 \) and \( \sigma_A^2 \). The third entry defines the prior distribution for \( \mu_A \) as normal, and the entries in the second row define those for \( \sigma_{AE}^2 \) and \( \sigma_A^2 \). Various choices could be made for the latter distributions, but the most convenient mathematically is that the reciprocals of all the variance components are given gamma or equivalently scaled-\( \chi^2 \) distributions. Following Theobald, Talbot and Nabugoomu (2002), these distributions are conveniently specified using a prior estimate of the variance and a corresponding degrees-of-freedom parameter, larger values of this parameter indicating greater confidence in the estimate. For example, \( \sigma_A^2 \) is given a prior estimate denoted by \( s_A^2 \) with degrees of freedom \( d_A \), and the notation \( \sigma_A^2 \sim IC(s_A^2, d_A) \) means that \( d_A s_A^2 \sigma_A^{-2} \) has the distribution \( \chi^2(d_A) \), so that \( s_A^{-2} \) is interpretable as the expectation of \( \sigma_A^{-2} \) in the prior distribution.

The third and fourth rows in Table 1 relate to the expansiveness parameters, and are similar to the first two rows, except that the \( \beta_i \) are centred around a value of 1. This is necessary in order to fix the scale of these parameters, since in models (1) and (3) multiplying all the \( \beta_s \) by a positive constant and dividing the \( \theta_s \) by the same constant would leave the right hand side unchanged. The product effects are similarly made identifiable by centring their means \( \phi_k \) on zero. The final entry in Table 1 specifies a common prior distribution for the assessors’ residual variances.

5 Combined analysis of apple-tasting experiments

The data on the sweetness of apples from the 45 experiments were combined using means over replicates within combinations of assessor and product rather than using raw scores directly. This simplification causes a slight loss of efficiency because of missing values, but it is the usual procedure for combining data for meta-analysis, such as in analysis of variety trials in the UK. We note, though, that Smith, Cullis and Thompson (2001) advocate analysing individual plot
yields for such trials. From now on, we treat these means as our responses.

To apply the model of Section 3 to these data, we specify the prior mean $m_A$ and variance $r_A^2$ of $\mu_A$ in Table 1 as 33 and 30. The prior estimates of the variance components and their degrees of freedom are given in Table 2. These values were obtained from an expert who had been involved in the programme of sensory studies carried out at the Hannah Research Institute (E A Hunter, personal communication). The few degrees of freedom for some of the variances indicate his lack of experience with analysing series of experiments rather than individual ones.

The Bayesian analysis was carried out using the WinBUGS program, which is freely available from http://www.mrc-bsu.cam.ac.uk/bugs. Using an iterative process known as Markov Chain Monte Carlo (Gilks, Richardson and Spiegelhalter (1996)), it generates long sequences of values drawn from the joint posterior distribution of the unknown parameters, that is their distribution given the prior distribution and the information in the data. These sequences can then be used to approximate the posterior distributions of individual parameters.

Summaries of the posterior distribution may include expectations and standard errors for individual parameters: such summaries are given for the parameters specific to the assessors in Model (3) in Table 3. The posterior expectations for the mean effects $\alpha_i$ of assessors have a range of about 50, suggesting large differences in the perception of sweetness. Thus the absence of certain assessors from a particular experiment could substantially influence the mean scores for products. The large standard error for assessor 12 is due to the fact that she left the panel of assessors after the first 16 experiments. The estimates of precision for assessors 1, 2, 5, 6, 7, 8 and 13 suggest that they have high discriminating abilities. Using the estimates of expansiveness or standard deviation alone leads to different judgements about which assessors are the best discriminators.

Systematic effects of experiment number might arise from changes in the source or composition of products. Such effects for apple sweetness may be investigated by plotting the posterior expectations of the mean product effects within experiments $\phi_k$ against the experiment numbers $k$, which are in date order. Considering these effects rather than the simple mean scores for experiments corrects for the absence of particular assessors from some of them: the posterior expectations of the individual product effects $\theta_{j(k)}$ are corrected in the same way. The plot is
given in Figure 2. A downward trend from the first experiment is followed by a sharp increase
around the twentieth, and then a slight decline to the last. No firm explanation can now be
found for the sudden increase, but it occurs at the start of a set of seven experiments in each
of which only four products were tasted. A partial auto-correlation plot revealed no significant
auto-correlation between the posterior expectations of the $\phi_k$.

6 Analysis of future experiments

By using the knowledge gained about the assessors from a combined analysis of data from
several experiments, as described in Sections 4 and 5, it should be possible to make more precise
comparisons of product effects in future experiments which use the same product type and pool
of assessors, and to correct estimated product effects for missing assessors. To adapt model (3)
for future experiments, the product effects are assumed to have the same structure as those for
the past series, with mean product effects drawn from the same population as the past ones.

We use $f$ to replace the index $k$ identifying the experiment and $e$ to replace $j$ as the product
index, so that the future responses by assessor $i$, denoted by $y_{ief}$, satisfy

$$y_{ief} = \alpha_{if} + \beta_{if} \theta_{e(f)} + \epsilon_{ief}.$$  \hspace{1cm} (4)

By analogy with the fifth row of Table 1, future product effects $\theta_{e(f)}$ are then assumed to be
drawn from a Normal population $N(\phi_f, \sigma^2_{PE})$, where $\phi_f$ is itself drawn from $N(0, \sigma^2_E)$.

It would also be possible to extend this model further to a future experiment that included
some assessors who did not take part in the past experiments, so long as their mean effects,
expansiveness and variances could be assumed to be drawn from the same distributions as for
the existing assessors.

To illustrate the possible benefit from including information provided by the analysis of previ-
ous experiments with the same assessors, we compare analyses of data from a single experiment
(treated as a future one) first ignoring information on model parameters from the other exper-
iments and then including this information. Since the main interest in sensory experiments is
likely to be to establish differences between products, we measure the reduction in the average
variance (over all pairs of products) of product differences in the posterior distribution. If $a$
and \( b \) denote any two of \( p \) products assessed in the future experiment then the average posterior variance of the \( p(p-1) \) product differences is given by

\[
\frac{1}{p(p-1)} \sum_{a \neq b} \text{var}(\theta_{a(f)} - \theta_{b(f)}) ,
\]

where the variance relates to the posterior distribution. Theobald et al. (2002) show that (5) may be simplified to

\[
\frac{2}{p-1} \left\{ p \sum_{e=1}^{p} \text{var}(\theta_{e(f)}) - p \text{var}(\bar{\theta}_f) \right\} ,
\]

where \( \bar{\theta}_f \) denotes the mean of the \( \theta_{e(f)} \). Thus the inclusion of past information on assessors is measured by examining the reduction in (6).

For the data on the sweetness of apples, the effect of including past information is illustrated by using the entire series as past data and one experiment which all 14 of the assessors happened to attend as the future experiment. The data from this experiment are thus used twice, but their influence on the posterior distribution from the combined analysis is slight since it arises from only one of 45 experiments.

The comparison under the heading ‘Expert prior distribution’ in Table 4 was made assuming the prior distribution specified in Section 5. In order to examine the robustness of this comparison to changes in the prior distribution, a second comparison was made after dividing by 5 the degrees of freedom corresponding to all the variance components: this is shown under the heading ‘Diffuse prior distribution’. As expected, the latter prior distribution gives a higher average posterior variance of product differences, but in both cases including information from other experiments results in a reduction of over 50\% in the average variance of the product differences.

7 Discussion

We consider a situation in which quantitative sensory data are available from several experiments on similar products with overlapping sets of assessors. In such a situation, it should be possible to extend the usual analysis of individual experiments by combining the information they contain on assessor effects and variance components. Such extensions include
• examining any pattern in the scores over time

• adjusting estimates to account for missing assessors

• using the information gained on assessors in the past to obtain more precise analyses of future experiments.

We have shown how these extensions can be achieved in the context of a multiplicative model for normally distributed scores. Such models can be expected to fit better than additive ones when the assessors differ in how much of the recording scale they use, but fitting multiplicative models in the usual frequentist framework can lead to zero estimates of variance components and unstable estimates of assessor expansiveness. The Bayesian framework offers a solution to these problems while incorporating expert knowledge into the analysis, and it deals with missing data and prediction for future experiments.

The effects of missing assessors on product means can be corrected by using assessors’ responses in the experiments that they did attend. Incorporating information from past analyses into future ones is also seen to increase precision by reducing the average variance of product differences.

One objection to Bayesian modelling is that choosing appropriate prior distributions may be difficult. In cases where prior information seems unreliable, it may be given a low weight by using small degrees of freedom for the variance parameters, and this is better than ignoring prior information completely.
Acknowledgements

This work was done as part of the first author’s PhD study which was funded by the Cecil Renaud Charitable and Educational Trust Fund, South Africa. This work was also supported by funds from the Scottish Executive Environment and Rural Affairs Department. The apple-tasting data were kindly provided by the Hannah Research institute in Ayr, Scotland. We are grateful to the late Rob Kempton for suggesting this study and to E A Hunter for his advice on the analysis of apple-tasting data.
References


Figure captions

Figure 1: Plots of mean sweetness scores against apple varieties for four experiments

Figure 2: Plot of posterior expectations of mean product effects within experiments against experiment number
Figure 1: Plots of mean sweetness scores against apple varieties for four experiments
Figure 2: Plot of posterior expectations of mean product effects within experiments against experiment number.
Table 1: Structure of the parameters in multiplicative model (3) for combined data. The notation $\sigma^2 \sim IC(s^2, d)$ means that $d s^2 \sigma^{-2}$ has the distribution $\chi^2(d)$.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_{ik} \sim N(\alpha_i, \sigma^2_{AE})$</td>
<td>$\alpha_i \sim N(\mu_A, \sigma^2_A)$ $\mu_A \sim N(m_A, r_A^2)$</td>
</tr>
<tr>
<td>$\sigma^2_{AE} \sim IC(s^2_{AE}, d_{AE})$</td>
<td>$\sigma^2_A \sim IC(s^2_A, d_A)$</td>
</tr>
<tr>
<td>$\beta_{ik} \sim N(\beta_i, \sigma^2_{BE})$</td>
<td>$\beta_i \sim N(1, \sigma^2_B)$</td>
</tr>
<tr>
<td>$\sigma^2_{BE} \sim IC(s^2_{BE}, d_{BE})$</td>
<td>$\sigma^2_B \sim IC(s^2_B, d_B)$</td>
</tr>
<tr>
<td>$\theta_{j(k)} \sim N(\phi_k, \sigma^2_{PE})$</td>
<td>$\phi_k \sim N(0, \sigma^2_E)$</td>
</tr>
<tr>
<td>$\sigma^2_{PE} \sim IC(s^2_{PE}, d_{PE})$</td>
<td>$\sigma^2_E \sim IC(s^2_E, d_E)$ $\sigma^2_i \sim IC(s^2, d)$</td>
</tr>
</tbody>
</table>
Table 2: Values defining the prior distributions of variance components for model (3) applied to the apple-tasting data.

<table>
<thead>
<tr>
<th>Variance</th>
<th>Estimate</th>
<th>Degrees of freedom</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_A^2$</td>
<td>264</td>
<td>7</td>
</tr>
<tr>
<td>$\sigma_B^2$</td>
<td>0.14</td>
<td>9</td>
</tr>
<tr>
<td>$\sigma_E^2$</td>
<td>30</td>
<td>5</td>
</tr>
<tr>
<td>$\sigma_{AE}^2$</td>
<td>66</td>
<td>4</td>
</tr>
<tr>
<td>$\sigma_{BE}^2$</td>
<td>0.0225</td>
<td>5</td>
</tr>
<tr>
<td>$\sigma_{PE}^2$</td>
<td>40</td>
<td>15</td>
</tr>
<tr>
<td>$\sigma_i^2$</td>
<td>50</td>
<td>30</td>
</tr>
</tbody>
</table>
Table 3: Expectations and standard errors of the posterior distributions for the parameters of model (3) that relate to individual assessors in the apple-tasting data.

<table>
<thead>
<tr>
<th>Assessor</th>
<th>Mean effect $\alpha_i$ (SE)</th>
<th>Expansiveness $\beta_i$ (SE)</th>
<th>SD $\sigma_i$ (SE)</th>
<th>Precision $\pi_i$ (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>48.6 (2.29)</td>
<td>1.26 (0.15)</td>
<td>9.92 (0.46)</td>
<td>0.127 (0.017)</td>
</tr>
<tr>
<td>2</td>
<td>26.6 (2.14)</td>
<td>1.01 (0.14)</td>
<td>7.18 (0.38)</td>
<td>0.141 (0.020)</td>
</tr>
<tr>
<td>3</td>
<td>30.1 (2.39)</td>
<td>1.31 (0.16)</td>
<td>13.48 (0.64)</td>
<td>0.097 (0.013)</td>
</tr>
<tr>
<td>4</td>
<td>47.8 (2.13)</td>
<td>1.31 (0.16)</td>
<td>12.46 (0.56)</td>
<td>0.106 (0.014)</td>
</tr>
<tr>
<td>5</td>
<td>65.1 (2.32)</td>
<td>1.19 (0.15)</td>
<td>8.18 (0.43)</td>
<td>0.146 (0.021)</td>
</tr>
<tr>
<td>6</td>
<td>57.2 (2.15)</td>
<td>0.78 (0.14)</td>
<td>5.41 (0.28)</td>
<td>0.146 (0.027)</td>
</tr>
<tr>
<td>7</td>
<td>40.4 (2.44)</td>
<td>1.03 (0.15)</td>
<td>7.69 (0.43)</td>
<td>0.135 (0.022)</td>
</tr>
<tr>
<td>8</td>
<td>35.8 (2.10)</td>
<td>1.44 (0.16)</td>
<td>10.17 (0.51)</td>
<td>0.142 (0.015)</td>
</tr>
<tr>
<td>9</td>
<td>28.1 (2.22)</td>
<td>0.80 (0.14)</td>
<td>9.66 (0.43)</td>
<td>0.083 (0.015)</td>
</tr>
<tr>
<td>10</td>
<td>60.8 (2.16)</td>
<td>0.65 (0.13)</td>
<td>6.49 (0.29)</td>
<td>0.101 (0.021)</td>
</tr>
<tr>
<td>11</td>
<td>15.6 (2.05)</td>
<td>0.82 (0.14)</td>
<td>9.58 (0.49)</td>
<td>0.085 (0.016)</td>
</tr>
<tr>
<td>12</td>
<td>22.2 (4.16)</td>
<td>0.54 (0.21)</td>
<td>6.68 (0.53)</td>
<td>0.081 (0.032)</td>
</tr>
<tr>
<td>13</td>
<td>53.2 (2.04)</td>
<td>0.81 (0.12)</td>
<td>5.19 (0.26)</td>
<td>0.156 (0.025)</td>
</tr>
<tr>
<td>14</td>
<td>47.9 (2.23)</td>
<td>0.93 (0.15)</td>
<td>9.19 (0.45)</td>
<td>0.102 (0.018)</td>
</tr>
</tbody>
</table>
Table 4: Average posterior variances of product differences from analysing a single apple-tasting experiment ignoring or including data from the other experiments: two prior distributions are assumed.

<table>
<thead>
<tr>
<th>Analysis ignoring other experiments</th>
<th>Expert prior distribution</th>
<th>Diffuse prior distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analysis including data on other experiments</td>
<td>4.49</td>
<td>6.09</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Analysis ignoring other experiments</th>
<th>Expert prior distribution</th>
<th>Diffuse prior distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analysis including data on other experiments</td>
<td>9.51</td>
<td>12.41</td>
</tr>
</tbody>
</table>