Animal Behaviour 80 (2010) 1065–1074

Contents lists available at ScienceDirect

Animal Behaviour

journal homepage: www.elsevier.com/locate/anbehav

‘Personality’ in bumblebees: individual consistency in responses to novel colours?

Helene Muller [a,*], Heiko Grossmann [b], Lars Chittka [a]

[a] Research Centre for Psychology, School of Biological and Chemical Sciences, Queen Mary University of London
[b] School of Mathematical Sciences, Queen Mary University of London

ARTICLE INFO

Article history:
Received 4 May 2010
Initial acceptance 17 June 2010
Final acceptance 6 September 2010
Available online 14 October 2010
MS. number: 10-00305R

Keywords:
Bombus terrestris
bumblebee
episodic personality
neophilia
neophobia
repeatability

It is now recognized that many vertebrates and a few invertebrates show individual-specific consistency in their behaviour across time and context, sometimes in ways that can be paralleled with human personality. Our work aimed at assessing behavioural consistency in a social insect: the bumblebee *Bombus terrestris*. We focused on a behavioural dimension commonly used in personality studies: the response of an individual to novelty (neophilia/neophobia spectrum). We used a foraging paradigm to quantify individual bees’ response to novel flower colours and to assess the repeatability of this response over time. As for vertebrates, most individual bumblebees responded to a novel stimulus by increasing the time they spent investigating it compared to known stimuli. Using a new statistical approach, the consistency model, we found that individual bees tended to be consistent in their response to novelty over a few hours but were not consistent in their behaviour over 3 days. We conclude that for the neophilia/neophobia paradigm used here, bumblebee foragers do not fulfil the criteria for animal personality in the common sense of the term. Instead their behavioural response to novelty appears to be plastic, varying on a day to day basis.

The number of publications reporting the existence of animal ‘personality’ has increased dramatically over the past few decades (Heinrich 1976, 1979; Bolnick et al. 2003; Bell 2007; Réale et al. 2007). There are now numerous reports of consistent individual differences in behaviour in species as varied as fishing spiders, *Dolomedes triton* (Johnson & Sih 2007), squid, *Euprymna tasmanica* (Sinn et al. 2008) and blue tits, *Cyanistes caeruleus* (Dingemanse et al. 2004). Some of the best-documented reports of reproducible interindividual differences have been on social bees (Thomson & Chittka 2001; Burns & Dyer 2008), where the genetic basis (Page et al. 1998, Page & Scheiner 2006) and sensory physiology and neurobiology (Spaethe & Chittka 2003; Weidenmüller 2004; Roussel et al. 2009) of some such differences are especially well understood (Thomson & Chittka 2001), explaining, in some cases, the particular specialization of individuals within the colony ‘superorganism’ (Weidenmüller 2004).

Research on animal personality almost invariably assumes that individual consistency indeed exists for the traits in question, and that the challenge for scientists lies in discovering trade-offs or causal links between them and/or with other traits. Indeed, studies often take only a single measurement of an animal’s response to a particular stimulus configuration (e.g. Minvielle et al. 2002; Dochtermann & Jenkins 2007; Hollander et al. 2008). However, research on the species best studied in terms of personality, human beings, demonstrates that while some behavioural traits remain stable over an individual’s adult lifetime in a manner consistent with personality, others vary with mood, hormonal cycles, age and other factors not yet identified. Some such changes might be adaptive in line with developmental stage of an organism or environmental context (Sinn et al. 2008; Roussel et al. 2009) or their individual experience (Bell & Sih 2007), while other changes might vary more unpredictably or be epiphenomena of other processes. This emphasizes the necessity to measure behavioural traits repeatedly and over various timescales (Sinn et al. 2008).

In this paper, we adopt the view that to qualify as a personality trait a certain pattern of behaviour needs to be exhibited in a consistent way over different situations and over time, but that this consistent behaviour may vary between individuals (Briffa et al. 2008; Schuett et al. 2010). Based on this notion of consistency, we introduce a novel statistical model and develop a new approach for assessing intraindividual consistency. We focus on a behavioural dimension commonly used in animal personality research: neophilia/neophobia. Neophilia is defined as attraction to novel objects whereas neophobia is repulsion from novel objects (Martin & Fitzgerald 2005). Interindividual differences on this dimension could be of considerable relevance in bumblebees’ natural foraging ecology. Their floral food resources are far from reliable. Nectar or pollen production by a given flower species might vary across a day (Goulson 2003; Heinrich 2004) and seasonal availability, as well as

* Correspondence: H. Muller, Research Centre for Psychology, School of Biological and Chemical Sciences, Queen Mary University of London, Mile End Road, London E1 4NS, U.K.
E-mail address: h.muller@qmul.ac.uk (H. Muller).

0003-3472/$38.00
doi:10.1016/j.anbehav.2010.09.016

competition with other pollinators, further complicates the task of the foragers. When given the opportunity, bumblebee foragers tend to specialize on a few flower species (Heinrich 1979; Chittka et al. 1999). However, a particular individual's preferred flower species might decline in profitability over various timescales, making it necessary for the forager to seek alternative food sources and more profitable flower species might become available. This therefore leads to the question of how bees keep track of the changes in their environment. Do individual bees vary in their tendency to investigate unknown flowers, and, if so, is this variation generated by consistent interindividual differences rather than chance variation?

We studied the behaviour of individual bumblebees foraging on artificial flowers in the laboratory. This controlled environment allowed us to introduce new flower 'species' and remove previously exploited food sources. We used our new model to investigate to what extent individual bees responded consistently to flowers of novel colours over various timescales. We discuss the implications of the findings from these experiments for regarding neophilia/neophobia as a personality trait.

METHODS

Five colonies (henceforth referred to as colonies A–E, with colony D being mentioned only in the Appendix) of *Bombus terrestris* containing approximately 40 workers each were obtained from Syngenta Bioline (Weert, Netherlands) between September 2008 and April 2009. Queens of this species mate only once in their life; therefore all the foragers within a colony are full sisters (Schmid-Hempel & Schmid-Hempel 2000). Upon delivery, colonies were transferred into bipartite wooden nestboxes (28 × 16 cm and 11 cm high). All tested bees were individually tagged with coloured dots or Opalith number tags (Christian Graze KG, Weinstadt-Endersbach, Germany). The nestbox in which the colony was housed was connected through a plastic tube to a foraging flight arena (120 × 100 cm and 35 cm high) covered with a UV-transparent Plexiglas lid. Bees could be allowed one at a time into this arena using a system of shutters built in the connecting tube. The room in which the colonies were kept had an average ambient temperature of 21 °C. Controlled illumination for laboratory experiments was provided by high-frequency fluorescent lighting (TMS 24F lamps with 4.3 kHz ballasts; Philips, Eindhoven, The Netherlands) fitted with Activa daylight tubes (Osram, Munich, Germany) to simulate natural daylight above the bee flicker – fusion frequency (Dyer & Chittka 2004).

Preparation and Pretraining

Colour-naïve foragers (i.e. bees that had never encountered any food-providing coloured object) were initially allowed to forage from translucent gravity feeders (von Frisch 1967) which provided unlimited supplies of 50% (w/w) sucrose solution for 2 days; the same concentration was used in all experiments described below. Subsequently, individual foragers were 'pretrained' in the arena (Fig. 1) to use an array of eight translucent artificial 'flowers' (Plexiglas squares of 24 × 24 mm, 4 mm thick and with a well for sucrose solution in its centre, mounted on glass 'stalks' 4.3 cm high). In what follows we refer to the time the bee spent foraging in the arena between two unloading trips to the nest as a 'foraging bout'. Here and in all experiments below, the position of each flower on a 6 × 5 grid (with 14 cm between positions) was randomized by using a computer random generator. We used 15 different random spatial patterns, varying the pattern between bees and between two foraging bouts (this shuffling of the position of the flower was necessary because the bees would otherwise have learned to associate location with reward). The artificial flowers were cleaned with soap and water between foraging bouts to ensure there were

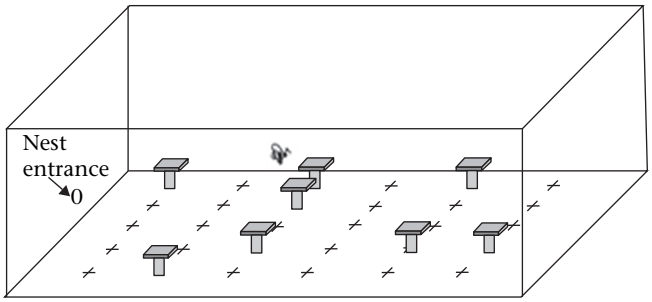

Figure 1. Experimental set-up. The bees were released one at a time in the arena containing the artificial flowers distributed in a random pattern. The sides of the arena were made of wood whereas the top lid was UV-transmitting Plexiglas. The position of the flowers was changed after every foraging bout.

no scent marks left from previous visits (Saleh et al. 2007). To encourage the bees to visit all eight flowers in the arena, we adjusted the amount of sucrose solution available in each flower so that the total volume in all eight flowers matched the stomach capacity of each individual bee. This was achieved by decreasing or increasing the volume of sucrose available in each flower over the three pretraining foraging bouts, until the bees visited all eight flowers. The translucent flowers are achromatic and pretraining to achromatic stimuli has been shown not to affect colour preference during subsequent exposures to coloured stimuli (Giurfa et al. 1995; Raine et al. 2006). For this reason, and because the first three foraging bouts were simply training to the set-up (and adjustment of the quantity of sucrose per flower), these foraging bouts were not included in the data analysis.

Short-term Consistency Test

The main variable measured throughout all our experiments was the feeding latency of the bee, that is, the time elapsed between flight initiation and first probing (proboscis extension) of the well of a flower. Bees were exposed to three colours (which, to bees, appeared UV, blue-green and UV-blue, see Appendix Fig. A1) for three consecutive foraging bouts each (i.e. three foraging bouts on colour A, then three foraging bouts on colour B and finally three foraging bouts on colour C). Because the bees were colour-naïve, each colour appeared to the bee as 'novel' at the beginning of the first foraging bout. Over three foraging bouts, bees learned to associate it with a reward (see Appendix) and so the colour could be regarded as familiar at the end of these three bouts. At foraging bout 4, colour B appeared as 'novel'. The bee then learned to associate it with reward during the course of foraging bouts 4–6. Colour C appeared as 'novel' in foraging bout 7 with gradual familiarization until bout 9.

Just as in the pretraining phase, the arena contained eight rewarded flowers randomly positioned on the grid. We varied the order in which the colours were presented to control for order effects. Twenty-seven foragers were tested from colonies A, B and C. For each colony, 15 bees were presented with the order UV–blue-green–UV-blue, six were given the sequence blue-green–UV–UV-blue and six were given UV-blue–UV–blue-green (thus a total of 81 bees were used in this experiment).

Video recordings of the beginning of each foraging bout (until the first probing of a flower) for 28 bees were used to split the feeding latency into two variables: the first approach time, which is the time between initiation of flight and approach (hovering within 2 cm) of the first flower by the bee; and cumulative investigation time, which is the time the bee spent hovering within a 2 cm radius of individual flowers before the first landing. Sometimes the bee would approach more than one flower before landing, so the sum of the time spent investigating each flower prior to the first landing was used to

produce the cumulative investigation time. Since we were interested in the strength of the relationships between these variables, Spearman rank correlation coefficients were calculated between first approach time and feeding latency and between cumulative investigation time and feeding latency.

To test for a relationship between body size and response to novelty, we measured the maximum thorax width three times per individual and took the average as an estimate of the bee's size (thorax width is the most common measure of body size for bumblebees; Goulson et al. 2002). We measured all tested 27 bees from colony A. We performed a Spearman rank correlation test using the maximum thorax width and the average novelty response as variables. A bee's novelty response to a given colour was calculated by subtracting the feeding latency of the last foraging bout on this particular colour from the feeding latency of the first foraging bout of this colour. The novelty responses for the three colours were then averaged to provide us with a single measure of individual response to novel colours, which we refer to as the average novelty response. By using the difference between the first and last foraging bouts, we accounted for putative differences between bees in terms of overall flying speed.

In a subset of bees (colony A: $N = 7$; colony C: $N = 15$; total of 22 individuals) the identification tags were applied within 48 h of emergence from the pupae, enabling us to determine their age at the time of the experiment. To test for a potential correlation between the response to a novel stimulus and age, a Spearman rank correlation test was performed using the age (days since emergence from the pupa) as one variable and the average novelty response as the other.

Long-term Consistency Test

This experiment aimed at assessing bees' consistency over several days instead of a few hours. It followed the same template as the short-term experiment with individual, colour-naïve, bees being exposed to three colours for three foraging bouts each. However, in this experiment, only one colour per day was presented to the bee. Thus, on day 1, a bee would be tested for three consecutive foraging bouts with colour A. On day 2, the same bee would be tested for three consecutive foraging bouts with colour B and on day 3, she would be tested with colour C in the same way. As for the short-term experiment, we varied the order of the colours and each bee was pretrained with translucent flowers (meaning that she would do three foraging bouts with translucent flowers on day 1 before being exposed to colour A). On subsequent days, each set of three foraging bouts was preceded by one foraging bout with translucent flowers to ensure the bee resumed foraging. Twenty-five bees from colony E were used for this experiment.

Assessment of Behavioural Consistency: the Consistency Model

A frequent difficulty in demonstrating consistency of behavioural traits is to demonstrate consistency statistically, despite the inevitable variation in nearly all behaviours. Indeed, if under the given design, regardless of the colour of the flower (which can influence the bee's behaviour, see Results), the feeding latencies of some bees decreased steadily between the first and third foraging bouts, then this might be explained by some characteristic of the bee which may be regarded as an aspect of her 'personality'. Of course, such a regular pattern is likely to vary between individual bees (see Fig. 2). Thus, it is necessary to identify those bees showing consistent behaviour in terms of a stable pattern of feeding latencies. The procedure we propose amounts to comparing fitted individual profiles of feeding latencies with the empirical data and classifying those bees for which the fitted and actual feeding latencies are in good agreement as being consistent.

To this end, an ANOVA-type general linear mixed-effects model is first fitted to the data. This model takes into account the repeated measurements structure of the data and incorporates several factors likely to affect the response (e.g. the colour of the flower). Moreover, the model captures the crucial aspect of consistent behaviour by allowing every bee to have a different pattern of feeding latencies which, for any given individual, is assumed to stay the same over the three sets of three foraging bouts.

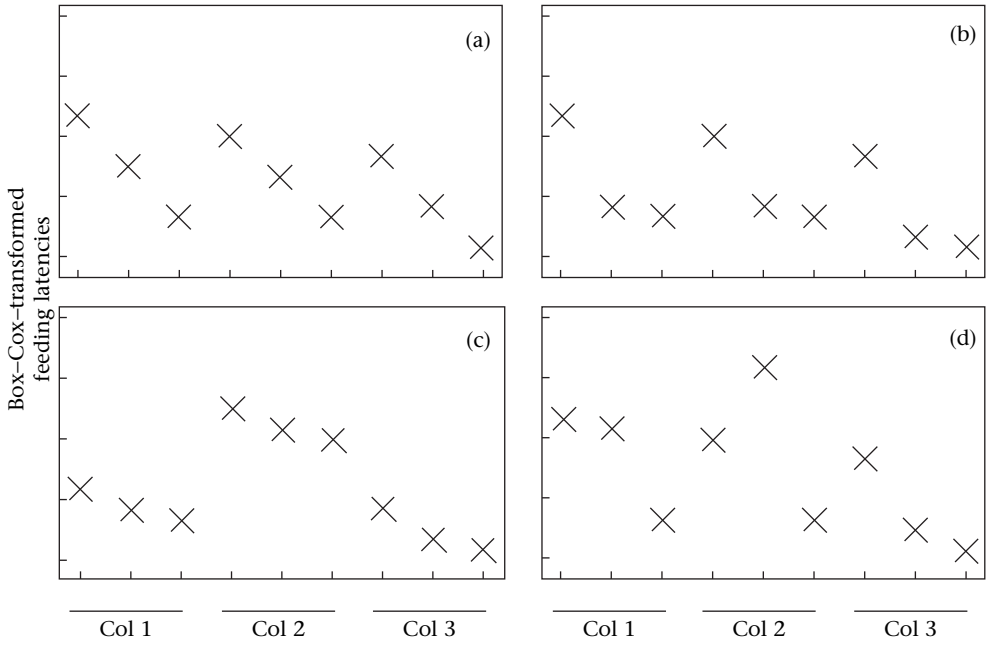

Figure 2. Schematic representation of different patterns for four fictitious bees. 'Col 1', 'col 2' and 'col 3' are the colours used for each set of three foraging bout (each cross represents the Box–Cox-transformed feeding latency for one foraging bout). In (a) the bee shows a gradual decrease in latencies to the stimuli, while in (b) she shows a sharp decrease in her latency to feed from a 'new' stimulus after just one exposure. In (c) the bee has a strong aversion for the second colour. In (d) the bee shows no consistency as her response to the stimuli varies both within and between colours.

Second, a distance measure is used to determine, for each bee separately, how good the fitted and actual feeding latencies agree. By specifying a cutoff value it is then possible to classify all bees for which the discrepancy between the fitted profile and the original observations does not exceed the cutoff as being consistent. A disadvantage of this approach is that specifying the threshold necessarily involves some sort of judgement. One possibility to circumvent this problem is to look at the cumulative distribution function (cdf) of the distances which, for every possible cutoff value x, depicts the proportion of consistent bees. This is useful for exploring how the proportion of consistent bees varies with the threshold. Furthermore, results from different experiments can easily be compared by comparing the corresponding cdfs.

As feeding latencies are typically positively skewed, prior to fitting the model a transformation is employed to make the distribution of the data more normal. Here we used the well-known Box–Cox transformation (Box & Cox 1964) which is a kind of power transformation and depends on a single parameter λ that can be estimated from the data. The basic nested mixed model for the transformed feeding latencies is then

$$Y = \mu + a_i + b_{j(i)} + c_{k(ji)} + \alpha_i + \beta_{m(ji)} + e$$

where Y represents the response and e is the error term. The Greek and Latin letters represent fixed and random effects, respectively:

μ = overall mean,
a_i = random effect of colony i,
$b_{j(i)}$ = random effect for bee j within colony i,
$c_{k(ji)}$ = bee-specific random effect for the kth set of three foraging bouts per bee,
α_l = fixed overall effect of colour l,
$\beta_{m(ji)}$ = bee-specific effect of foraging bout $m = 1, 2, 3$ within every set of three bouts within bee j in colony i. These three parameters per bee are the same for every set k of three bouts and define the bee's response pattern.

This model reflects the fact that sets of three foraging bouts, where each set of three corresponds to a different colour, are nested within individual bees, which in turn are nested within colonies. The repeated measurements nature of the data is accounted for by allowing correlations between (1) the feeding latencies for the three foraging bouts per colour within each bee and (2) the feeding latencies across colours within each individual bee.

The basic model can be refined to separate, for every individual bee, the effect of the first foraging bout within a set of three, which corresponds to the onset of a new colour, from the combined effect of the other two bouts for the same colour. Similarly, it is possible to separate the effect of any of the colours from the combined effect of the other two colours. These modifications correspond to testing predefined orthogonal contrasts as part of the analysis of variance.

Using Genstat Release 10.1 (VSN International Ltd, Hemel Hempstead, U.K.), we fitted this model separately to the data from the short-term and long-term experiments. By fitting the model, we obtained for each bee a profile consisting of nine values. These are the values that give the best fit to the transformed feeding latencies in the three sets of three foraging bouts. To assess how closely the fitted profiles agree with the data, an appropriately standardized version of the usual Euclidean distance can be computed for each bee separately. More specifically, the distance measure we propose is defined as

$$d = \frac{\sqrt{\sum_{i=1}^{9}\left(Y_i - \widehat{Y}_i\right)^2}}{\mathrm{IQR}}$$

where Y_i and $\widehat{Y}_i$ are, respectively, the transformed feeding latencies and fitted values for a single bee, and IQR is the interquartile range of the transformed feeding latencies over all bees. Dividing the Euclidean distance by the interquartile range is similar to standardizing the distance by means of the standard deviation. However, using the IQR makes the proposed distance measure more robust against extreme observations and hence appears to be preferable.

This distance can be regarded as a consistency index (one per bee) with small values of d indicating greater consistency. As explained above, it is a measure of the fit between the predictions of the model and the (Box–Cox-transformed) data and therefore measures the repeatability of a bee's behaviour across colours and foraging bouts.

RESULTS

Short-term Consistency Test

Foragers had longer feeding latencies when the colour of the flowers was unfamiliar than when it was a colour that they had previously experienced (Fig. 3). This appeared to be the case independently of the colour considered or of the position of the colour in the sequence and is confirmed by statistical analysis (see below). There was no significant correlation between the bees' sizes and their average response to novelty (Spearman rank correlation: $r_S = 0.09$, $N = 27$, $P = 0.63$). The same was true for the age of the bees and their average response to novelty (Spearman rank correlation: $r_S = 0.15$, $N = 22$, $P = 0.50$).

Video recordings demonstrated that the observed latency in landing on a novel stimulus was due to the bees' reluctance to land on an unknown flower colour, and not caused by the bee persisting in searching for the previously rewarding stimulus (which would be indicative of dietary conservatism). This is demonstrated by breaking up the feeding latencies into first approach time (time spent flying in

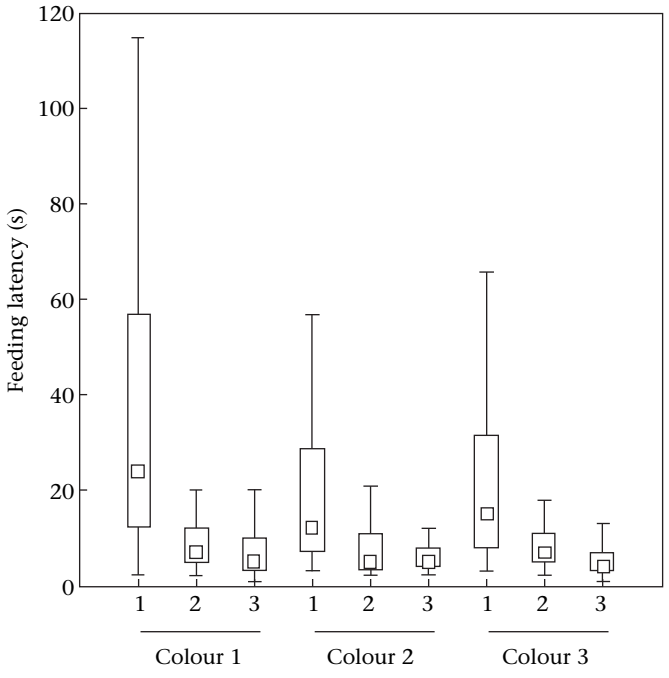

Figure 3. Box plot diagram of the feeding latency of all workers, all colonies and all colours pooled. Small squares represent medians, large rectangles are interquartile ranges and whiskers indicate the ranges of the largest nonoutlier observations. Numbers on the X axis correspond to the foraging bout of the experiment. The colour to which the bee was exposed depended on the order of colours to which she was been subjected.

the arena before approaching the first flower) and cumulative investigation time (hovering close to the new stimulus). We found a significant correlation between the feeding latency (all colours, all colonies pooled) and the first approach time (Spearman rank correlation: $r_S = 0.28$, $N = 28$, $P < 0.001$) and between the feeding latency and the cumulative investigation time (Spearman rank correlation: $r_S = 0.77$, $N = 28$, $P < 0.001$). Because the mean cumulative investigation time ($X_{cit} + SD = 9 + 21.7$ s) was much more strongly correlated with the feeding latency ($X_{fl} + SD = 23 + 66.0$ s) than the first approach time ($X_{fat} + SD = 4 + 10.7$ s), we conclude that most of the observed variation in the response to the new colour is explained by the amount of time the bees spend investigating (hovering close to) the new colour. The first approach time explains much less of the feeding latency. If the bee was merely ignoring the new stimulus while actively searching for the known stimulus, then we would expect her to spend most of her feeding latency flying in the arena instead of hovering close to the new stimulus. We would expect the first approach time to explain much more of the feeding latency than the cumulative investigation time and we observed the opposite phenomenon. This suggests that feeding latency is an appropriate measure of a bee's response to a novel stimulus.

ANOVA

The maximum likelihood estimate of the parameter λ of the Box–Cox transformation for the feeding latencies was equal to −0.4. Table 1 shows the analysis of variance for the transformed data. This accounts for 80.6% of the variability in the transformed latencies as measured by the coefficient of determination R^2. Residual plots indicate that the distribution of the residuals is close to normal. We adopted a Bonferroni adjustment of the 1% significance level to account for the fact that we performed multiple tests. In total, there are $k = 5$ tests and hence an effect is regarded as being significant only if the P value is smaller than $\alpha = 0.01/k = 0.002$. This approach is very conservative, but is intended to avoid drawing conclusions that may not stand up in replication studies.

With this standard, there is clear evidence for interindividual variability in the feeding latencies between bees ('bees within colonies level', Table 1). There is also a strong overall effect of colour, and, independently of the effect of colour, there is a strong intraindividual effect of the position of the foraging bout within the three replications of the same colour ('bouts within sets within bees', Table 1).

A refined analysis, which includes a preplanned comparison for distinguishing between the effect of the first bout and the combined effect of the remaining two bouts for any set of three bouts, splits the sum of squares and the degrees of freedom for bouts in Table 1 into two independent components without changing any of the other rows of the table. The corresponding test reveals that the already reported effect of the position of the foraging bout is due to the difference in the feeding latencies for the first and the other bouts (ANOVA: $F_{81,324} = 7.14$, $P < 0.001$), that is the effect is due to the onset of a new colour, irrespective of what that colour is. The P values for the other tests in Table 1 are also comparatively small. Although the corresponding effects are not regarded as being significant by our standard, these P values confirm that colonies and sets within bees should be included as blocking factors in the analysis.

Table 1
Analysis of variance for the short-term experiment

Source	*df*	Sum of squares	Mean square	*F*	*P*
Colony level	2	2.42	1.21	4.00	0.02
Bees within colonies level	78	23.61	0.30	3.08	<0.001
Sets within bees within colonies level					
Colour	2	3.60	1.80	18.32	<0.001
Sets within bees	160	15.72	0.10	1.45	0.003
Units level					
Bouts within sets within bees	162	46.21	0.29	4.20	<0.001
Residual	324	22.01	0.07		
Total	728	113.56			

Consistency model

By fitting the ANOVA model, we obtained individual profiles of nine fitted values for each bee (Fig. 4, Appendix Fig. A2). For every individual bee, this profile was superimposed on her Box–Cox-transformed feeding latencies and the consistency distance was calculated. For the 81 bees, the consistency distance varied between 0.13 and 2.13 with a mean of 0.78 and SD of 0.41.

The solid line in Fig. 5 represents the cumulative distribution function for the distances. For every possible cutoff value x on the horizontal axis it gives the proportion of bees in the experiment for which the discrepancy between the fitted profile and the data is smaller than or equal to the threshold x and which hence would be regarded as behaving consistently. An advantage of using the cdf is that the sensitivity of the classification can be easily explored. For example, if we adopt a threshold of $x = 0.65$ for the distances, the proportion of consistent bees, as given by Fig. 5, is 0.44 or 36 bees. For $x = 0.60$ and $x = 0.70$, the proportions from the figure are 0.36 and 0.49 corresponding to 29 and 40 bees, respectively. Since the total number of 81 bees in the experiments is known, no information is lost by using the cdf and every proportion can be converted back to the underlying number of bees.

Long-term Consistency Test

The analysis of the long-term experiment paralleled that for the short-term experiment except that no test for the colony effect could be performed since all bees were from the same colony. Here the maximum likelihood estimate of the parameter λ for the Box–Cox transformation was equal to −0.5. Table 2 presents the analysis of variance of the transformed feeding latencies. The analysis accounted for 75.7% of the total variation in the transformed latencies and the distribution of the residuals as judged by residual plots was again close to normal. As before, a Bonferroni adjustment of the 1% significance level was carried out; the number of tests is equal to $k = 4$ with a corresponding $\alpha = 0.01/k = 0.0025$.

Over the longer timescale, interindividual differences between bees could not be detected (Table 2). Similarly, there is no significant effect of colour (Table 2). There is, however, strong evidence for differences between sets of foraging bouts and an intraindividual effect of the position of the foraging bout within the sets of three bouts under any given colour (Table 2). As in the short-term experiment, the refined analysis shows that this position effect is due to the onset of a new colour (ANOVA: $F_{25,100} = 4.51$, $P < 0.001$).

Since sets of foraging bouts were performed on different days, the significant effect of sets indicates that between-day variation is probably the most important factor governing the bees' foraging behaviour. Notwithstanding, the effect of novelty of a colour still prevails, as indicated by the significant result for bouts within sets and the significant position effect for the first bout within a set.

To assess consistency, we obtained individual profiles of fitted values and corresponding distances, as in the short-term experiment. Here the distances varied between 0.52 and 2.00 with a mean of 1.05 and SD of 0.42. The corresponding cumulative distribution function is shown by the dotted line in Fig. 5. As before, by using the cdf the effect of different threshold values of x can be explored. For example, for a threshold of $x = 0.65$ the proportion of consistent bees is equal to 0.12 or three of 25 bees. Similarly, for $x = 0.60$ and $x = 0.70$ the corresponding proportions are 0.8 and 0.24 which correspond to two and six bees, respectively.

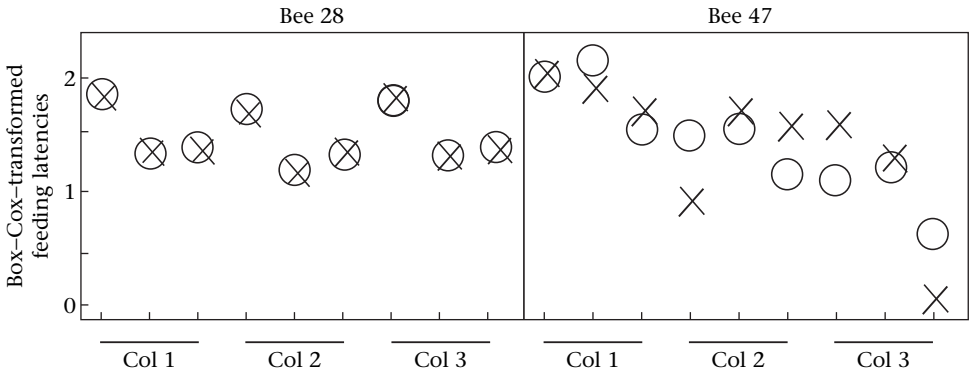

Figure 4. Examples of individual profiles generated by the consistent bee model for two representative bees. Open circles are fitted values from the model and crosses are actual Box–Cox-transformed feeding latencies. 'Col 1', 'col 2' and 'col 3' are the colours used for each set of three foraging bouts (each cross represents the Box–Cox-transformed feeding latency for one foraging bout). Bee 28 is a good example of a very consistent bee. She has a consistency distance of 0.22, reflecting a good match between fitted and Box–Cox-transformed feeding latencies. By contrast, bee 47 has a consistency distance of 2.1, reflecting a poor match between fitted and Box–Cox-transformed feeding latencies.

Exact binomial tests for the threshold values $x = 0.60$, $x = 0.65$ and $x = 0.70$ of the null hypothesis that the proportion of consistent bees does not exceed 0.05 give P values of 0.358, 0.127 and 0.001, respectively. In the short-term experiment, exact binomial tests for each of the three thresholds reject the null hypothesis that the proportion of consistent bees is smaller than or equal to 0.20 with a P value of <0.001. Thus while for these values of x the proportions of consistent bees were clearly greater than 0.20 in the short-term experiment, over longer terms the proportion can only be shown to be above 0.05 for $x = 0.70$.

If we compare the cdfs in Fig. 5, it is obvious that the proportion of consistent bees in the short-term experiment is greater than the corresponding proportion in the long-term experiment regardless of the choice of threshold. One of the reviewers of this paper raised the question of whether greater variability between the three colonies in the short-term consistency test than in the long-term consistency test might have affected our findings. We reanalysed these data colony by colony and compared the results with those from the long-term consistency test (see Appendix Fig. A3). Visual inspection of the individual colonies' curves in the short-term consistency tests reveals the same pattern (and the same difference with the long-term consistency test) as for the pooled colonies (Appendix Fig. A3). Moreover, for each colony from the short-term experiment we tested whether the proportion of consistent bees (when using, as above, a threshold value of $x = 0.65$ for classifying a bee as consistent) was larger than 0.10. For each of the three colonies we obtained P values of <0.004 when testing the null hypothesis that the proportion was smaller than or equal to 0.10. In the long-term consistency test there was no significant evidence that the proportion of consistent bees was at least 0.05. This demonstrates that our results also hold for each colony in the short-term experiment separately and are not due to greater genetic variability in the combined sample.

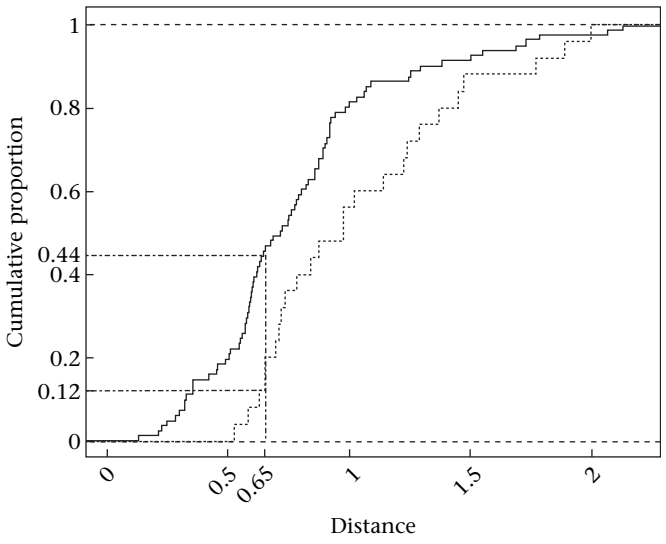

Figure 5. Cumulative distribution function, cdf, for the distances between Box–Cox-transformed feeding latencies and ANOVA-fitted model values. The solid line represents the cdf for the short-term experiment (within a day) whereas the dotted line represents the cdf for the long-term experiment (between days). The cdf for the short-term experiment is based on 81 bees, while the cdf in the long-term experiment is based on 25. For every possible cutoff value x on the horizontal axis the cdf line gives the proportion of bees in the experiment for which the discrepancy between the fitted profile and the data is smaller than or equal to the threshold x and which hence would be regarded as behaving consistently. As an example, at a threshold of $x = 0.65$, the proportion of consistent bees is 0.44 in the short-term experiment and 0.12 in the long-term experiment. If bees are considered to behave consistently when they have a distance value of 0.65 or below, then 36 of 81 bees (0.44) are regarded as consistent in the short-term experiment and three of 25 bees (0.12) are consistent in the long-term experiment.

In conclusion, the extent to which bees behaved consistently within a day was much greater than it was when tested over several days. This finding can be corroborated by testing whether the proportions for a given threshold x under the two conditions are equal. A nonparametric test that can be used for this purpose is Fisher's exact test. For example, for the threshold $x = 0.65$ the proportions of consistent bees in the short-term and long-term experiments are equal to 36 of 81 and three of 25, respectively, for which Fisher's exact test gives a P value of 0.004, providing clear evidence that the proportions are different. The same conclusion is reached when the equality of the proportions derived from other threshold values x is tested. For example, for $x = 0.60$ and $x = 0.70$ the P values are 0.010 and 0.037.

Table 2
Analysis of variance for the long-term experiment

Source	*df*	Sum of squares	Mean square	*F*	*P*
Bees level	24	3.23	0.13	1.23	0.26
Sets within bees level					
Colour	2	0.28	0.14	1.26	0.29
Sets within bees	48	5.25	0.11	2.18	<0.001
Units level					
Bouts within sets					
Within bees	50	6.85	0.14	2.74	<0.001
Residual	100	5.00	0.05		
Total	224	20.61			

The reported P values increase with x and for larger values of the threshold it is not always possible to demonstrate that there is a statistically significant difference between the corresponding proportions of consistent bees. Comparisons of proportions derived from a threshold $x > 0.70$ do not seem to be very meaningful, however, since in view of the narrow range of the scale on which the distances are measured, basing the classification of bees on such a large threshold would appear to be too liberal. This assessment also seems to be supported by visual inspection of the profiles and the corresponding distances in Appendix Fig. A4.

DISCUSSION

We have shown that, like many species of vertebrates, bumblebees respond to novel objects by investigating them extensively before first accepting them as a food source. When confronted with a flower of a new colour, the vast majority of bees spent time hovering closely around it, presumably to enable visual exploration of the novel stimulus. This investigative behaviour was longest during the first encounter with the new stimulus and then drastically decreased during subsequent encounters. Such behaviour is similar to the response described for many species of vertebrates (Heinrich 1995; Mettke-Hofmann et al. 2002) and is commonly quantified along a spectrum from neophilia (attraction for novel stimuli) to neophobia (repulsion for novel stimuli; Martin & Fitzgerald 2005). Our study therefore confirms anecdotal reports of 'novelty response' (Heinrich 1976) or neophobia (Forrest & Thomson 2008) in bumblebees. Additionally, our finding that the flower colour in itself had an effect on the bee's response to a novel flower is consistent with earlier studies which showed that bees have innate colour preferences (Lunau et al. 1996).

Sampling new food sources can be risky. While doing so, foragers might spend valuable time on unprofitable flower species (e.g. orchids mimicking nectar-producing flowers; Jersakova et al. 2006) or risk being attacked by ambush predators lurking on flowers (such as crab spiders; Ings & Chittka 2008). Following this view, 'neophilic' bees might take greater risks in terms of both predation and foraging efficiency than 'neophobic' bees. Nevertheless, the gains from sampling new food sources are potentially high. As Chittka et al. (1999), Mettke-Hofmann et al. (2002) and Martin & Fitzgerald (2005) pointed out, exploiting known resources is only advantageous if the foraging environment changes little over time. Bumblebees' foraging environment, however, appears to be ever-changing, across and within days (Goulson 2003; Heinrich 2004). Therefore, sampling new food sources will often be rewarding in the long-term. By keeping themselves up-to-date with the resources available in their environment, bumblebee workers run less risk of suffering from the depletion of their current food source. It is conceivable that temporally variable conditions in terms of predation threat and foraging conditions could maintain the variability in terms of the 'neophilia/neophobia' gradient observed here, or indeed explain a lack of selection towards behaviour consistency in this dimension. Additionally, variability in foragers' response could be beneficial at the colony level (Muller & Chittka 2008). 'Bet hedging' strategy has often been invoked in bees to account for intracolony variability in traits such as foraging speed – accuracy trade-offs (Burns & Dyer 2008) and could well explain some of the variability in response to novelty observed here.

Our results suggest that bumblebees' behavioural responses towards novel objects are not consistent enough to qualify formally as a 'personality trait' in the common use of the term. Many bees showed reasonable consistency (repeatable response to novelty) over a few hours, but very few bees could potentially be described as consistent when the experiment was repeated over 3 days. Individual bumblebees therefore appear to be inconstant in their response to novelty over periods longer than a few hours. The observed decrease in 'consistent bees' across time is unlikely to be explained by a developmental process (Sinn et al. 2008), because the change in response to novel targets was not predictable from worker age. It can also not be explained by an adaptive response to variation in environmental context, for example changes in predation threat or flower profitability, because all these parameters were kept constant in our experiments. Thus our results are more comparable to the more unpredictable variation in mood variability in humans (although hormonal changes sometimes predict such variation) or to the recent study by Pronk et al. (2010) which came to similar conclusions in the gloomy octopus, *Octopus tetricus*. Pronk et al. (2010) found that individual octopuses exhibited marked repeatability in response to various visual stimuli within a day but not between days, and termed this phenomenon 'episodic personality'. Both our results and those of Pronk et al. (2010) emphasize the need to investigate personality traits repeatedly and, ideally, over various timescales appropriate to the animals' life span. Single 'snapshots' of an individual's behaviour responses (or even repeated measurements over short timescales) might be misread to indicate individually predictable responses when such responses might fluctuate over longer timescales. As in our study, such variation might not have easily identifiable environmental inducers, or internal contributing factors such as those correlating with age.

Acknowledgments

We thank Sana Nasrullah for helping with preliminary work and Syngenta Bioline (Weert, Netherlands) for providing the colonies used in our experiments.

References

Bell, A. 2007. Future directions in behavioural syndromes research. *Proceedings of the Royal Society B*, **274**, 755–761.

Bell, A. & Sih, A. 2007. Exposure to predation generates personality in threespined sticklebacks (*Gasterosteus aculeatus*). *Ecology Letters*, **10**, 828–834.

Bolnick, D., Svanback, R., Fordyce, J., Yang, L., Davis, J., Darrin Hulsey, C. & Forister, M. 2003. The ecology of individuals, incidence and implications of individual specialization. *American Naturalist*, **161**, 1–28.

Box, G. E. P. & Cox, D. R. 1964. An analysis of transformations (with discussion). *Journal of the Royal Statistical Society B*, **26**, 211–252.

Briffa, M., Rundle, S. & Fryer, A. 2008. Comparing the strength of behavioural plasticity and consistency across situations: animal personalities in the hermit crab *Pagurus bernhardus*. *Proceedings of the Royal Society B*, **375**, 1305–1311.

Burns, J. & Dyer, A. G. 2008. Diversity of speed-accuracy strategies benefits social insects. *Current Biology*, **18**, R953–R954.

Chittka, L. 1992. The colour hexagon: a chromaticity diagram based on photoreceptor excitations as a generalized representation of colour opponency. *Journal of Comparative Physiology A*, **170**, 533–543.

Chittka, L., Thomson, J. D. & Waser, N. M. 1999. Flower constancy, insect physiology, and plant evolution. *Naturwissenschaften*, **86**, 361–377.

Dingemanse, N. J., Both, C., Drent, P. & Tinbergen, J. 2004. Fitness consequences of avian personalities in a fluctuating environment. *Proceedings of the Royal Society B*, **271**, 847–852.

Dochtermann, N. & Jenkins, S. 2007. Behavioural syndromes in Merriam's kangaroo rats (*Dipodomys merriami*): a test of competing hypotheses. *Proceedings of the Royal Society B*, **274**, 2346–2349.

Dyer, A. G. & Chittka, L. 2004. Fine colour discrimination requires differential conditioning in bumblebees. *Naturwissenschaften*, **91**, 228–231.

Forrest, J. & Thomson, J. D. 2008. Pollinator experience, neophobia and the evolution of flowering time. *Proceedings of the Royal Society B*, **276**, 935–943.

von Frisch, K. 1967. *The Dance Language and Orientation of Bees*. Cambridge, Massachusetts: Harvard University Press.

Giurfa, M., Núñez, J., Chittka, L. & Menzel, R. 1995. Colour preferences of flower-naive honeybees. *Journal of Comparative Physiology A*, **177**, 247–259.

Goulson, D. 2003. *Bumblebees: Their Behaviour and Ecology*. Oxford: Oxford University Press.

Goulson, D., Peat, J., Stout, J., Tucker, J., Darvill, B., Derwent, L. & Hughes, W. 2002. Can alloethism in workers of the bumblebee, *Bombus terrestris*, be explained in terms of foraging efficiency? *Animal Behaviour*, **64**, 123–130.

Heinrich, B. 1976. The foraging specializations of individual bumblebees. *Ecological Monographs*, **46**, 105–128.

Heinrich, B. 1979. 'Majoring' and 'minoring' by foraging bumblebees, *Bombus vagans*: an experimental analysis. *Ecology*, **60**, 245–255.

Heinrich, B. 1995. Neophilia and exploration on juvenile common ravens, *Corvus corax*. *Animal Behaviour*, **50**, 695–704.

Heinrich, B. 2004. *Bumblebee Economics*. Cambridge, Massachusetts: Harvard University Press.
Hollander, F., Van Overveld, T., Tokka, I. & Matthysen, E. 2008. Personality and nest defence in the great tit (*Parus major*). *Ethology*, **114**, 405–412.
Ings, T. & Chittka, L. 2008. Speed-accuracy tradeoffs and false alarms in bee responses to cryptic predators. *Current Biology*, **18**, 1520–1524.
Jersakova, J., Johnson, S. & Kindlmann, P. 2006. Mechanisms and evolution of deceptive pollination in orchids. *Biological Reviews*, **81**, 219–235.
Johnson, J. C. & Sih, A. 2007. Fear, food, sex and parental care: syndrome of boldness in the fishing spider, *Dolomedes triton*. *Animal Behaviour*, **74**, 1131–1138.
Lunau, K., Wacht, S. & Chittka, L. 1996. Colour choices of naive bumble bees and their implications for colour perception. *Journal of Comparative Physiology A*, **178**, 477–489.
Martin, L. & Fitzgerald, L. 2005. A taste for novelty in invading house sparrows, *Passer domesticus*. *Behavioral Ecology*, **16**, 702–707.
Mettke-Hofmann, C., Winkler, H. & Leisler, B. 2002. The significance of ecological factors for exploration and neophobia in parrots. *Ethology*, **108**, 249–272.
Minvielle, F., Mills, A. D., Faure, J. M., Monvoisin, J. L. & Gourichon, D. 2002. Fearfulness and performance related traits in selected lines of Japanese quail (*Coturnix japonica*). *Poultry Science*, **81**, 321–326.
Muller, H. & Chittka, L. 2008. Animal personalities: the advantage of diversity. *Current Biology*, **18**, R961–R963.
Page, R. E. & Scheiner, R. 2006. The development and evolution of division of labor and foraging specialization in a social insect. *Current Topics in Developmental Biology*, **74**, 253–285.
Page, R. E., Erber, J. & Fondrk, M. K. 1998. The effects of genotype on response thresholds to sucrose and foraging behavior of honey bees (*Apis mellifera* L.). *Journal of Comparative Physiology A*, **182**, 489–500.
Pronk, R., Wilson, D. R. & Harcourt, R. 2010. Video playback demonstrates episodic personality in the gloomy octopus. *Journal of Experimental Biology*, **213**, 1035–1041.
Raine, E., Ings, T., Dornhaus, A., Saleh, N. & Chittka, L. 2006. Adaptation, genetic drift, pleiotropy, and history in the evolution of bee foraging behavior. *Advances in the Study of Behavior*, **36**, 305–354.
Réale, D., Reader, S., Sol, D., McDougall, P. & Dingemanse, N. J. 2007. Integrating temperament within ecology and evolution. *Biological Review*, **82**, 291–318.
Roussel, E., Carcaud, J., Sandoz, J.-C. & Giurfa, M. 2009. Reappraising social insect behavior through aversive responsiveness and learning. *PLoS One*, **4**, e4197.
Saleh, N., Scott, A., Bryning, G. & Chittka, L. 2007. Distinguishing signals and cues: bumblebees use general footprints to generate adaptive behaviour at flowers and nest. *Arthropod–Plant Interactions*, **1**, 119–127.
Schmid-Hempel, R. & Schmid-Hempel, P. 2000. Female mating frequencies in *Bombus* spp. from Central Europe. *Insectes Sociaux*, **47**, 36–41.
Schuett, W., Tregenza, T. & Dall, S. R. X. 2010. Sexual selection and animal personality. *Biological Reviews*, **85**, 217–246.
Sinn, D. L., Gosling, S. & Moltschaniwsky, J. 2008. Development of shy/bold behaviour in squid: context-specific phenotypes associated with developmental plasticity. *Animal Behaviour*, **75**, 433–442.
Spaethe, J. & Chittka, L. 2003. Interindividual variation of eye optics and single object resolution in bumblebees. *Journal of Experimental Biology*, **206**, 3447–3453.
Thomson, J. D. & Chittka, L. 2001. Pollinator individuality: when does it matter. In: *Cognitive Ecology of Pollination* (Ed. by L. Chittka & J. D. Thomson), pp. 191–213. Cambridge: Cambridge University Press.
Weidenmüller, A. 2004. The control of nest climate in bumblebee (*Bombus terrestris*) colonies: interindividual variability and self reinforcement in fanning response. *Behavioral Ecology*, **15**, 120–128.

APPENDIX

Choice of Colours for the Consistency Tests

Since we are interested in the bee's response to novel colours, we have to ensure that the bees are able to distinguish between the three colours used. Indeed, a lack of response to a novel colour could simply mean that the bees do not perceive the colour as differing from the previous one. To this end, we selected colours that, following the hexagon model of bee colour vision (Chittka 1992; Fig. A1), should be easily distinguishable for bees. We used red, cream and pink artificial 'flowers' of the same shape as those used for pretraining. To bees, these flower types appeared as UV, blue-green and UV-blue, respectively, and were roughly equidistant in a bee's colour space (Fig. A1). However, we also wished to ascertain experimentally that bees could easily discriminate the colours.

To assess the bees' ability to discriminate between the colours used, we tested 10 foragers from colonies A, B, C and six from colony D. The experiment consisted of a training phase of three foraging bouts followed by a test. Individuals were first exposed to eight randomly positioned flowers of the colour A ('known' colour) for three consecutive foraging bouts. During the colour discrimination test, the arena contained 16 randomly distributed flowers, eight of which were of the same colour as in the immediately preceding foraging bout ('known colour') and eight of which were of the novel colour. These tests were unrewarded; all flowers contained 10 μl of water to mimic the visual appearance of sucrose solution. The number of flowers of each type visited was recorded as correct (landing on the 'known' colour) or incorrect (landing on the 'novel' colour) and the number of correct and incorrect choices out of the first 10 flowers chosen was used as a measure of the ability of the bees to discriminate between colours. Different individuals were used for the colour discrimination tests and the experiments testing responses to novel colours.

We found that bees are able to discriminate between the three colours we used in our experiments. Only 5% of the bees tested made more than 30% incorrect choices when having to discriminate between UV and blue-green (average percentage of incorrect choices $\pm$ SD $= 12 \pm 13\%$) and the same was true for when bees had to discriminate between blue-green and UV-blue (average percentage of incorrect choices $\pm$ SD $= 8 \pm 13\%$). All the bees tested for discrimination abilities between UV and UV-blue made less than 20% incorrect choices (average percentage of incorrect choices $\pm$ SD $= 7 \pm 10\%$). We found that the observed number of correct choices for all colour discrimination tests were significantly different from chance (binomial test: UV versus blue-green: $P < 0.0001$; blue-green versus UV-blue: $P < 0.0001$; UV versus UV-blue: $P < 0.0001$). This demonstrates that bees are able to discriminate well between the relevant colours after only three foraging bouts.

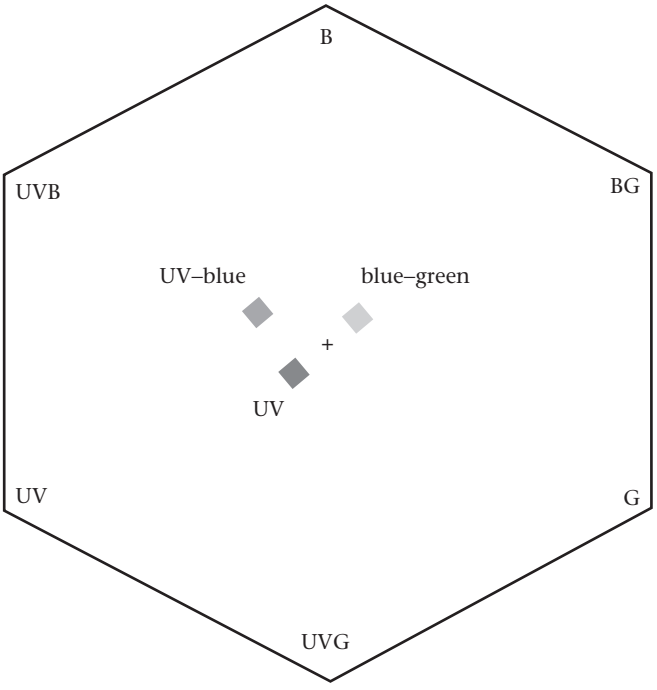

Figure A1. Loci of the artificial flower colours in the colour hexagon. Loci are calculated according to the relative stimulation of the three receptor types (UV, blue, green) elicited by the stimulus (Chittka 1992). The colour hexagon coordinates were obtained using spectrophotometer readings from 300 to 700 nm (i.e. including the ultraviolet, UV, range). The angular position (measured from the centre) in this colour space indicates bee-subjective hue, which is in turn determined by the relative excitations of bees' UV, blue and green receptors. A colour locus in the lower left portion of the colour hexagon indicates a 'bee-UV' colour, a colour locus in the 'up' direction denotes 'bee-blue' and so forth. Distances between colour loci indicate discriminability; the total distance between the centre and any of the corners of the colour hexagon equals unity, and distances above 0.1 are typically well distinguishable. Euclidian distances between the colours used here are 0.23 between UV and UV-blue, 0.24 between UV-blue and blue-green and 0.19 between UV and blue-green, and are therefore predicted to be well distinguishable by bees. B = blue, BG = blue-green, G = green, UVG = UV-green and UVB = UV-blue.

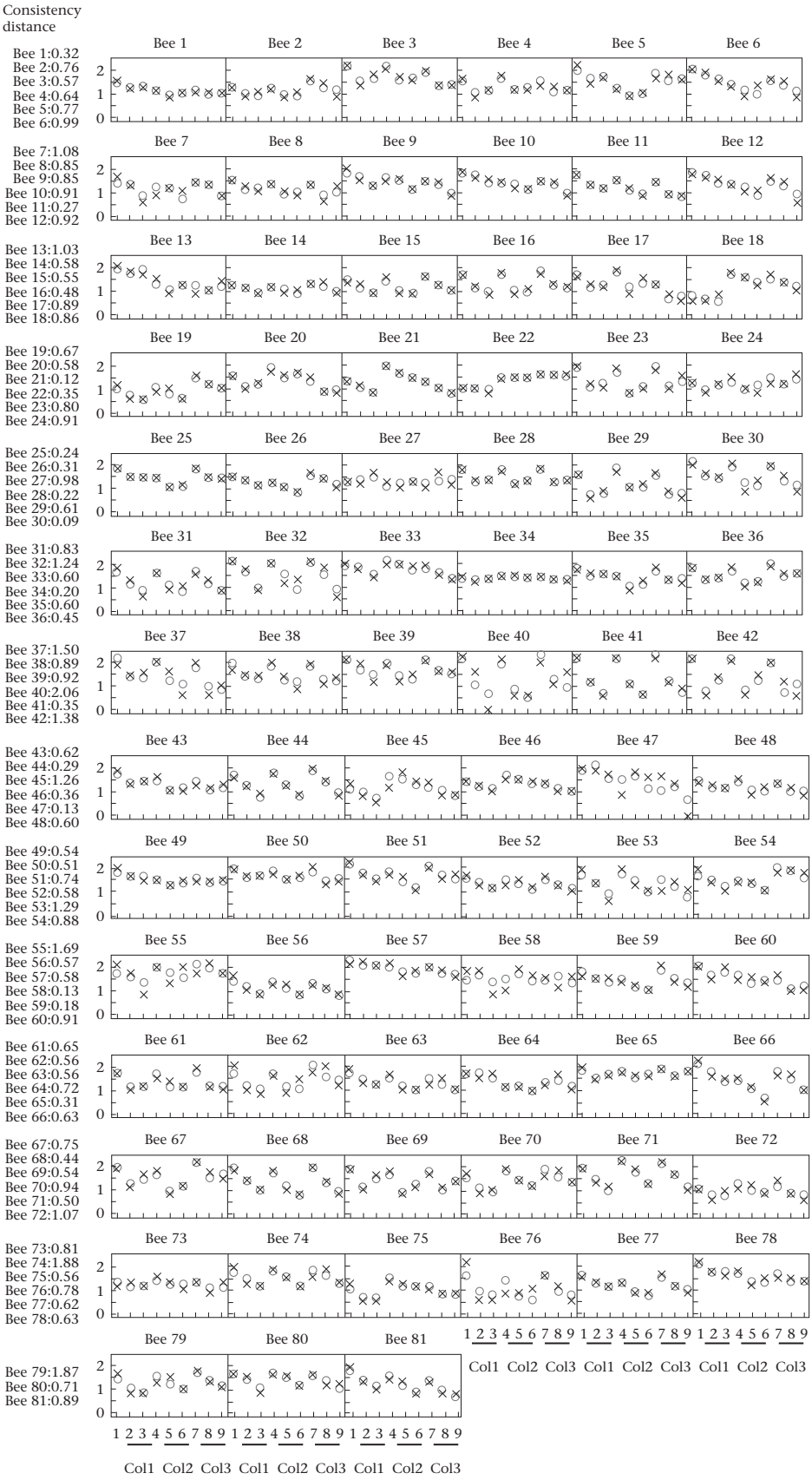

Figure A2. Individual profiles for 81 bees from the short-term experiment generated using the consistency model. Open circles are fitted values from the model and crosses are actual Box–Cox-transformed data. The consistency distance values represent the difference between the fitted values and Box–Cox-transformed data. Numbers 1–9 on the horizontal axis are the foraging bouts and 'col1', 'col2' and 'col3' are the colours used for each set of three foraging bouts.

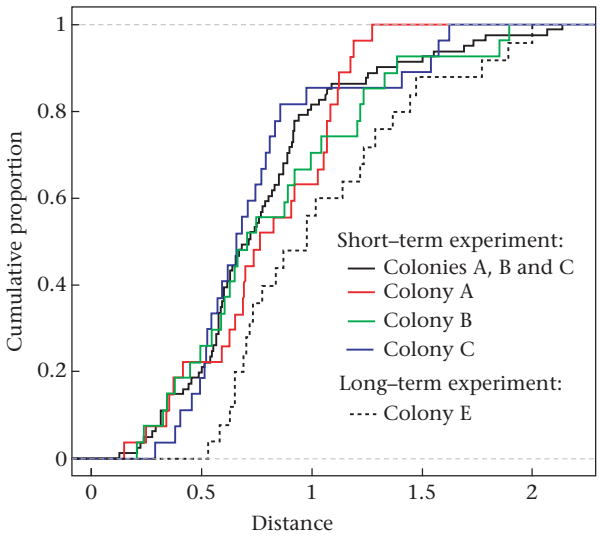

Figure A3. Cumulative distribution function for the distances between Box–Cox-transformed feeding latencies and ANOVA-fitted model values for the three colonies of the short-term experiment and the single colony of the long-term experiment. The solid lines represent the cdfs for each colony of the short-term experiment (within a day) whereas the dotted line represents the cdf for the long-term experiment (between days). For each of the three colonies in the short-term experiment we fitted exactly the same model as in the long-term experiment. Based on these separate analyses we plotted the cumulative proportions of bees for each threshold value for each colony in the short-term experiment, as in Fig. 5, and compared the resulting curve with the one for the colony in the long-term experiment. All the curves for the individual colonies in the short-term experiment are shifted to lower distance values than the data from the long-term experiment, indicating greater consistency in every colony used in the short-term experiment compared to the long-term experiment. This indicates that our findings are not caused by greater genetic heterogeneity in the pooled sample.

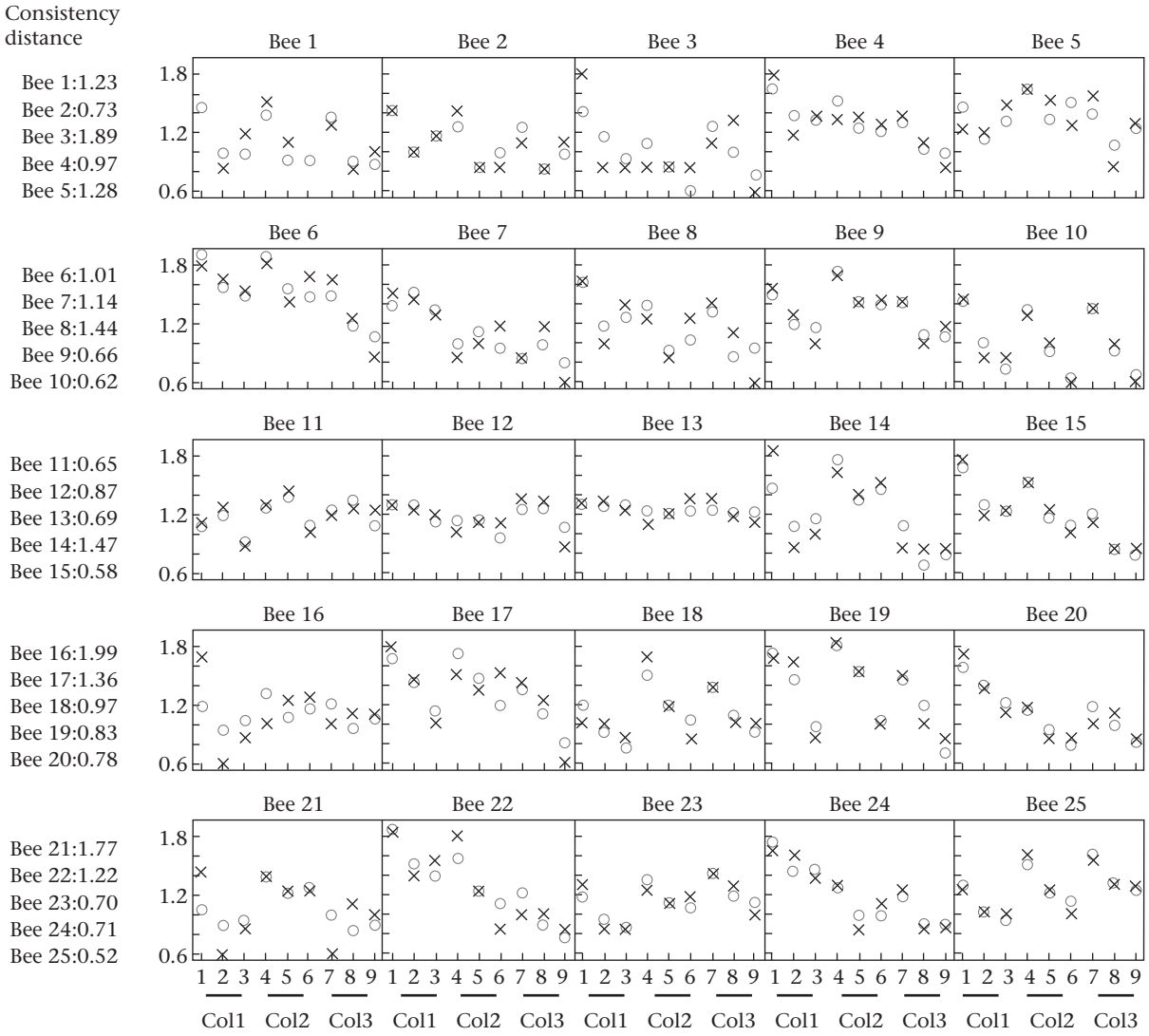

Figure A4. Individual profiles for the 25 bees from the long-term experiment generated using the consistency model. Open circles are fitted values from the model and crosses are actual Box–Cox-transformed data. The consistency distance values represent the difference between the fitted values and Box–Cox-transformed data. Numbers 1–9 on the horizontal axis are the foraging bouts and 'col1', 'col2' and 'col3' are the colours used for each set of three foraging bouts.