The Lepidoptera, i.e. the generally diurnal butterflies and nocturnal moths, are probably the most popular and best-known insect order because of their large and conspicuous wings. However, our understanding of the ecological functions of wing patterns is still incomplete. Thermoregulation by sunbasking, crypsis and Batesian and Müllerian mimicry have been shown to be important in a number of species (Vane-Wright and Ackery, 1984; Nijhout, 1991; Vane-Wright and Boppré, 1993). But these functions do not apparently hold universally, particularly not for species showing sexual dimorphism. In his ‘Origin of Species’, Darwin (Darwin, 1859) proposed that female sexual selection was a reason for conspicuous wing colours of males. In a number of species, wing colours and colour patterns have proved to be important in intersexual or intrasexual selection (Silberglied, 1984; Brunton and Majerus, 1995), but this was not true for most species (Silberglied, 1984). In the majority of species the evolution and maintenance of wing patterns cannot be explained.

How are wing colours and wing patterns produced in the Lepidoptera? The wing coloration is caused by the wings’ cover of coloured scales, an apomorphy that also gave the order its scientific name. The colours of the scales have several bases. Physical colours arise from the very elaborate scale structure. Interference colours result from multiple layers, three-dimensional diffraction gratings, Tyndall-blues and plain back scattering of light (Ghiradella, 1991). The pigments in the scales make an even greater contribution to wing colour. Most of these pigments are synthesised de novo during scale development in the pupa. The processes of pattern formation and pigment synthesis are highly regulated by switching on and off genes in ontogenesis (Nijhout, 1991; Brakefield et al., 1998).

A second source of wing pigments is secondary plant compounds taken up from the larval diet (Wilson, 1987; Nijhout, 1991). These pigments need to be stored by the larvae, retained through all instars and then transferred to the developing wings in the pupa. Little information has been published on these processes and on their relevance for the animals.

During the last few years, data on the sequestration of flavonoids by blue butterflies (Lycaenidae) from their larval food plants have become available (Geuder et al., 1997; Schittko et al., 1999; Burghardt, 2000). This work was done...
mainly on the common blue butterfly *Polyommatus icarus*, which is also the focal species of this study. It was shown that some of the flavonoids in the larval diet are taken up, partially metabolised by glycosylation (a process that probably serves as a means of detoxification), and stored. Sequestered flavonoids are retained up to the pupal stage, probably in the fat body. During the late pupal stage, most of the flavonoids are transferred to the wings, more or less simultaneously with the onset of pigment synthesis in the wing scales (Kornmaier, 1999). The sequestration of flavonoids appears to be a very specific process, as is also suggested by flavonoids being kept through all instars and transferred specifically to the wings.

The flavonoids are a major class of secondary plant compounds. All flavonoids are ultraviolet-absorbing, and depending on the type and placement of functional groups and side chains bonded to the central flavan, may possess human-visible colours (Harborne, 1991). The anthocyanins, a large subclass of the flavonoids, are the most important flower pigments (Brouillard and Dangles, 1993). *P. icarus* seems to sequester only flavonols, a subclass of flavonoids, which appear colourless to yellow to the human eye (Burghardt, 2000).

The ecological function of flavonoid sequestration in *P. icarus* is unknown. The specific allocation of these pigments to the wings called for an investigation of their effects on wing coloration. After it became clear that flavonoids do alter wing colours (see below), questions arose as to whether this serves as a colour signal to conspecifics and what information may be transferred by such a signal. We therefore set out to investigate the effects of flavonoids on butterfly wing coloration in detail.

The appraisal of colour in the context of animal communication must be carried out in a way that is independent of the human visual system (Endler, 1990; Cuthill and Bennett, 1993; Bennett et al., 1994). Colours have to be measured physically for the whole wavelength range visible to the species under study. This includes the ultraviolet range from 300 nm to 400 nm, which is usually not perceivable by humans without technical help (Endler, 1990). However, ultraviolet vision is well known in many vertebrates and invertebrates (Menzel, 1979; Jacobs et al., 1991; Jacobs, 1992; Tovée, 1995), and in other species ultraviolet light has been demonstrated to have at least a behavioural influence (Sharma et al., 1999). We therefore determined wing coloration by measuring wing spectral reflectance in a manner that was independent of human vision. Reflectance spectra covered the human-visible and the ultraviolet range (300 nm to 690 nm).

This is the range in which vision in butterflies is made possible by a combination of the light available in nature (lower limit approximately 300 nm) (Henderson and Hodgkiss, 1963; Henderson, 1970) and the spectral sensitivities of butterfly photoreceptors (up to approximately 700 nm) (e.g. Eguchi et al., 1982; Bernard and Remington, 1991).

Subsequent analyses of spectral data must also be performed independently of the human visual system. For most species, little or no data on visual physiology are available, so a statistical approach that is independent of any assumptions on a receiver’s psychology must be chosen (Endler, 1990; Cuthill et al., 1999). Principal component analysis (PCA) with subsequent multivariate analysis of variance (MANOVA) was introduced by Cuthill et al. (Cuthill et al., 1999) as a very sensitive tool for comparing spectra. Here, we will evaluate this method and compare it with an alternative, newly proposed procedure using averaged spectra and their confidence limits.

**Materials and methods**

**Study animal**

The common blue butterfly *Polyommatus icarus* (Rottemburg) has an almost Palaeartic distribution and prefers open habitats. Throughout its large distributional range, the species occurs in sympatry with many of its closely related species. The taxonomy of the genus *Polyommatus* is not yet satisfactorily solved. New species, often strongly resembling *P. icarus*, are added to the genus every year, mainly on the basis of subtle morphological characters such as wing patterns. *P. icarus* is a herbivore restricted to members of the legume family (Fabaceae), but uses a considerable range of herbaceous plant species from this family as larval food source (Martin Cano, 1984). However, because of its limited locomotory ability, a single caterpillar will usually be confined to the food plant where its mother had deposited the egg. Larvae of *P. icarus* preferentially feed on flowers, flower buds and very young fruits rather than on mature foliage (Burghardt and Fiedler, 1996).

*P. icarus* butterflies exhibit a strong sexual dimorphism, with males having an iridescent bright blue to violet upperside and females having dark brown uppersides with a submarginal row of more or less pronounced orange spots (Fig. 1). The undersides of both sexes are much less dimorphic, with submarginal orange spots and a number of black spots with white surrounds set on a background that is greyish in males and more brownish in females (Fig. 2). In this study, we looked only at the wing colours of female butterflies.

All individuals used for spectroradiometry and as dummies in the behavioural experiments were from the F1 or F2 generations of females caught in the field in northern Bavaria, Germany. The mate-choice experiments were carried out in the field or with our breeding stock. The latter consisted of a permanently maintained population to which field-caught animals and their offspring were added regularly to avoid inbreeding.

Caterpillars were reared on an artificial diet based on peas and wheatgerm, with some dried alfalfa (*Medicago sativa*) leaves added as a feeding stimulus. This diet contains only traces of flavonoids, and none of these flavonoids is sequestered by *P. icarus* larvae (Burghardt, 2000). To this mixture, the flavonoid quercetin (quercetin dihydrate, purum; Roth, Karlsruhe, Germany) (2.5 % dry mass) was added to yield a second diet that differed only in flavonoid content. Larvae fed this quercetin-spiked diet sequester large amounts of this flavonol, most of which is stored as quercetin-3-O-galactoside (Burghardt, 2000).
Fig. 1. The upper sides of imagos of *Polyommatus icarus*. (A) Males exhibit an iridescent bright blue to violet coloration due to structural elaboration of wing scales. (B) Females have dark brown uppersides with a row of more or less pronounced submarginal orange spots. Females may have an additional iridescent blue coloration, as in males, ‘dusted’ over a smaller or larger part of the wings. Scale bars, 5 mm.

Fig. 2. Undersides of female imagos of *Polyommatus icarus*. (A,B) Colour photographs, (C,D) ultraviolet photographs. (A) Individual reared on an artificial diet from which no flavonoids were sequestered. (B) Individual reared on the same artificial diet as in A, but with the flavonoid quercetin (2.5% dry mass) added. Only minor differences in wing coloration in the visible range are found between the diets. (C) Same individual as in A. The orange and black spots are highly ultraviolet-absorbing, the background is intermediate in ultraviolet reflectance and the white spots reflect most of the ultraviolet light. (D) Same individual as in B. The wing undersides absorb ultraviolet light effectively, with the white spots being almost indistinguishable from the background coloration. A gradient towards higher ultraviolet reflectance is visible from anterior to posterior in the forewings. A grey scale of known reflectance was included in every photograph (Mean spectral reflectance in blocks from 300 nm (dark) to 400 nm (light): 0.1%, 9%, 18%, 32%, 62%, 78%). Scale bars, 10 mm.
Reflectance spectroradiometry

The spectral reflectance of wing colours was measured using an Instaspec II diode-array photometer equipped with an MS125 spectrograph (400 lines mm\(^{-1}\) grating) (Oriel, Stratford). The measurement range was 300 nm to 690 nm, with a resolution of approximately 0.5 nm. Reflectance was measured relative to a Spectralon 99% white reflectance standard (Labsphere, Congleton). An Oriel 77184 sighting optics with fibre optics, extension tubes and an Ultra-Fluar 10/0.20 quartz microscope objective (Zeiss, Jena) allowed simultaneous viewing and measurement of wing patterns. All measurements were taken from a direction normal to the wing plane. Measured spot diameters ranged from 0.11 mm to 0.21 mm, and the numerical apertures (N.A.) of the measuring beam lay between 0.14 and 0.15. Illuminating light was delivered from an XBO 75W/2 OFR xenon arc lamp (Osram, Munich) with an Oriel 68806 stabilised power supply, Oriel fibre optics and focusing beam probe. Illuminating light was focused onto the wings (N.A.=0.07–0.09) at 45° to the wing plane. Illumination was always from the distal direction. To ensure exactly the same wing orientation for all measurements, the wings were positioned with their hind margins parallel to the projection of the illuminating beam into the wing plane. The wings were mounted on black velvet to eliminate stray light. The wings were handled very carefully with pointed forceps at the most proximal end only, where no measurements were taken.

In all subsequent analyses, only the spectral reflectance data for the hindwings are presented. A clear gradient of ultraviolet reflectance was present in the forewings, while no such spatial variation was apparent in the hindwings within a given food treatment (see Results). For each individual, a randomly chosen hindwing was used in the measurements. For each of the human-visible colour categories (black, white, orange and brown for undersides; brown and orange for uppersides), a number of measurements were taken from each wing (usually 5 for black, orange and brown and 10 for white and background). Individual measurements of a given colour category were sampled from all areas of a wing. For every measurement, the exact position (i.e. pattern element or position relative to veins) on the wing was recorded.

Comparing spectral data

Principal component analysis (PCA) (for an introduction to the method see Kim and Mueller, 1978a; Kim and Mueller, 1978b; Chatfield and Collins, 1980) has become increasingly popular for comparing spectra (Young, 1986; Bakker and Arnold, 1993; Cronin et al., 1997; Cuthill et al., 1999; Grill and Rush, 2000). PCA reduces the amount of data to a small number of variables for each spectrum. These new variables can then be used for statistical analyses, such as MANOVA (Kim and Mueller, 1978a; Kim and Mueller, 1978b; Chatfield and Collins, 1980; Cuthill et al., 1999). This method has the advantage of being independent of any assumptions about visual physiology. It simply makes a statistical comparison of spectra. We will explain this method briefly. (For more details on the usage of principal component analysis with spectra, see Endler, 1990; Cuthill et al., 1999).

We start with spectra, which are measured values for a number of wavelength intervals. These wavelength intervals can be seen as variables that are highly intercorrelated because most biological objects exhibit broadband absorption or reflection of light. The idea of PCA is to transform these independent variables into new, orthogonal variables, the principal components (PCs) using linear combinations. The new variables are the values used to transform the original data points into PC scores.

Before PCA, we reduced the wavelength resolution of our spectra, which originally consisted of 795 points each, by calculating mean values for 10 nm intervals. Means represented our data better than medians, and there were no outlying datapoints to pose a problem. The original and reduced data were almost identical, as judged by visual inspection of the spectra. We decided to separate the analysis of mean reflectance of spectra from that of spectral shape, as proposed by Endler (Endler, 1990) and applied for the first time by Cuthill et al. (Cuthill et al., 1999). To achieve this, we subtracted from every spectrum its mean reflectance and then used these standardised spectra for PCA (StatSoft, 1999a).

Mean reflectance, as a measure of ‘brightness’, was analysed by analysis of variance (ANOVA). The first three PCs of standardised spectra were analysed separately, as a measure of spectral shape, by MANOVA. ANOVA and MANOVA (StatSoft, 1999a) were both repeated-measures designs with larval diet as a between-group factor and spectra as replicate measurements from individual butterflies. PCA and statistical analyses were performed separately for every colour category.

An alternative method of comparing groups of spectra is to use the means of the spectra and the confidence intervals of these means. Calculation of these statistics is simple and straightforward. For a given group of \( n \) spectra, each made up of \( i \) data points (measured values at wavelength \( \lambda_i \)), one calculates the arithmetic mean and standard error for every value of \( \lambda_i \). Multiplication of the tabulated value of Student’s \( t \) for \( n-1 \) degrees of freedom at a chosen probability level, e.g. \( P=95\% \), by the standard errors yields the confidence intervals of the mean spectrum (Sokal and Rohlf, 1981; Sachs, 1997). If the confidence intervals of two groups of spectra do not overlap, one can reject the null hypothesis that the parametric means of both groups are the same at a probability level \( P \) (if \( n>10 \) for both groups). One can compare for overlap of confidence intervals in specific wavelength ranges, and at the same time, relate to absolute reflectance values in these ranges.
To demonstrate the use of confidence intervals of spectra in derived measures of colour, we calculated the quantum catch of hypothetical photoreceptors for mean spectra and their confidence limits (Bernard and Remington, 1991). Mean spectra and confidence limits transform into mean values of quantum catch with confidence intervals for the wing colours. We used the nomogram templates given by Stavenga et al. (Stavenga et al., 1993). Values for the maximum sensitivity of the photoreceptors were taken from the work of Bernard and Remington (Bernard and Remington, 1991) on two Lycaena species because these are the only such data published for lycaenid butterflies. The illuminating light was the generalised CIE daylight spectrum D65 (Schanda, 1997) transformed to quantum flux and normalised to unity.

Courtship and mate choice in Polyommatus icarus

At favourable ambient temperatures, male *P. icarus* patrol the habitat in search for females. After locating a female, a male approaches her, usually exhibits a distinct fluttering behaviour next to the female, alights, sometimes still fluttering, and tries to copulate (see Pellmyr, 1982). Female polyommatin blue usually mate only once (Häuser, 1990; Drummond, 1984). Males may mate several times, but require at least a day for the production of a new spermatophore (for related species, see Drummond, 1984). Male approaches to females occur quite frequently but are only very rarely successful. The courtship sequence is almost never completed. Usually the male decides to stop at some point in the sequence before trying to copulate and goes on searching for another female. Alternatively, the female refuses copulation when the male tries to grasp her with his genitalia. Rape is impossible because in the Polyommatus section the female has to actively evert a sclerotised gonoporus for copulation (Häuser, 1993). Therefore, both sexes exhibit mate choice.

We observed male reactions to female dummies made from dead dried females. Freshly emerged females were allowed to stretch and harden their wings, then killed by freezing, placed on insect pins and dried at room temperature. Dummies were prepared from females that had been raised either on an artificial diet from which no flavonoids are sequestered or on the same artificial diet but with quercetin added. Dummies were prepared with their wings in postures found in the field: the wings closed above the body, the wings opened at an angle of approximately $90^\circ$ or the wings spread at almost $180^\circ$. Prior to the behavioural experiments, the dummies were dried for 2 days at 40–50°C to evaporate any volatile substances.

Mate-choice experiments in the laboratory

Experiments were performed in a flight cage (2.1 m × 1.1 m × 1.0 m high) in an environmental chamber at 30°C. The cage was lit by fluorescent lamps (at >10kHz) with visible and ultraviolet light (eight Osram Biolux L58W/72-965 and two Osram L36W/73). Lamp tubes were positioned 20 cm above the cage ceiling made of a transparent plastic screen. In this cage the butterflies exhibited essentially all the behaviour patterns observed in the field, and they courted and mated readily.

In the centre of the cage, nine wooden poles (diameter 10 mm, length 23 cm) were placed upright in a square with one pole in the centre and three poles along each side of the square (or three rows of three sticks each; both poles within a row and rows 10 cm apart). The dummies were pinned onto these poles during the experiments. Three replicates of each of the two dummy types (differing in either food treatment, size or alteration of the external appearance; see below) were used in each test. The remaining three positions were left empty. The dummies were positioned so that the nearest neighbours of every dummy were always either (i) one of the other dummy types or (ii) empty poles. Positions were changed every 30 min to avoid positional bias. A complete experiment included nine such rounds of 30 min each, resulting in a total observation time of 270 min per experiment. During an experiment, a single dummy was placed in all nine positions and all positions were occupied for the same time by the different dummy types. The size of the dummies was controlled, with the mean wing lengths of dummy types within an experiment differing by no more than 0.5 mm (except where the effect of dummy size was tested).

The experiments were performed with our breeding population from the flight cage. However, most of the females in the cage were removed just before an experiment and immediately replaced afterwards because they tended to sit and bask for prolonged periods on the poles and thus to ‘block’ them (see below). 50–80 butterflies were in the cage during the experiments, but not all of these were active. The numbers of active butterflies were more or less constant during an experiment but differed between experiments.

During the experiments, the reactions of the males to the dummies were counted and recorded on a dictaphone. The observer sat outside the flight cage and watched the animals through a white mosquito net from a darker part of the room. The presence of the observer had no noticeable effect on the behaviour of the butterflies. The reactions of the males fell into several clearly distinguishable categories. If a male changed his flight path and approached a female dummy or steered directly towards it, this was scored as ‘approach’. If a male exhibited the distinct fluttering behaviour next to the dummy, usually circling around it, this was counted as ‘fluttering’. ‘Landings’ and ‘copulation attempts’ were the other categories distinguished. When a butterfly landed on a pole and remained there, this position was considered as occupied because the reactions of other males to this pole might have been caused by the presence of the living butterfly rather than the dummy. Reactions to occupied poles were therefore not scored, and the duration of occupancy recorded. Similarly, reactions to a pole were not scored when another individual was flying in the vicinity.

In a series of six experiments (three experimental designs, each replicated twice), we tested male reactions to dummies made from flavonoid-free *versus* flavonoid-rich females. In each experiment, only dummies of the same posture type (wings opened at 90°, wings fully spread, or wings closed) were presented. With the wings opened at 90°, both ventral
and dorsal wing surfaces were visible. In experiments with the wings fully spread or closed, each dummy was placed on a round cardboard disk (diameter 3.5 cm) mounted on top of the pole. With the wings fully spread, the undersides of the wings were mostly hidden by the white, ultraviolet-absorbing cardboard. In experiments with the wings closed, only the undersides of the wings were visible.

A seventh experiment tested the influence of dummy size. A group of small (mean wing length 11.5±0.2 mm) versus large (mean wing length 15.4±0.3 mm; means ± s.d., N=6) flavonoid-rich dummies was presented (wings spread at 90°).

In an eighth experiment, we used dummies with altered external appearance. Dummies were made as usual from flavonoid-free females with the wings opened at 90°. The undersides of the dummies were then sprayed with a solution of the quercetin-3-O-diglycoside rutin (purum; Roth, Karlsruhe, Germany; 2 mg ml⁻¹ in ethanol). Rutin is frequently found in adult P. icarus (Burghardt, 2000). The absorption spectrum (300–690 nm) of rutin is virtually identical to that of other quercetin-3-O-glycosides, particularly to quercetin-3-O-galactoside, which is the dominant flavonoid in butterflies reared on the quercetin-spiked diet (Burghardt, 2000; H. Knüttel, unpublished results). The externally applied rutin reduces ultraviolet reflectance of the wings, and therefore mimics the light-absorbing effect of dietary flavonoids on wing coloration. At the same time, any other potential physiological effects of dietary flavonoids which might alter the constitution of the dummies were precluded. The rutin solution was cautiously delivered from a vaporiser and dried off almost immediately, leaving a layer of tiny rutin crystals on the wing scales. This was repeated until reflectance spectroradiometry indicated colours very similar to those of naturally flavonoid-rich butterflies. The second group of dummies in this experiment was sprayed with ethanol only.

The use of Pearson’s chi-squared tests in the statistical analysis of frequency-type behavioural data, although being chosen quite frequently by many researchers, is highly debated (e.g. Hurlbert, 1984; Lombardi and Hurlbert, 1996; Wilson and Dugatkin, 1996). Therefore, we decided to use an alternative approach (Kramer and Schmidhammer, 1992). Each of the experiments described above was divided into subsets. For each subset, the proportions of reactions to the two dummy types corrected for actual observation time (see below) were calculated, thereby collapsing all reactions of the males in a subset into a single data point for each dummy type. Subsequently, using a two-sided t-test for dependent data sets, we tested the null hypothesis that these proportions did not differ between the two dummy types in an experiment, i.e. that the reactions of the male population in the flight cage were not biased with respect to dummy type. In the experiments that were run twice, the proportions of both experiments were collated and subjected to a single t-test.

In the experiments which tested the effect of dummy size, or of external application of rutin, each subset for statistical analysis comprised one round of 30 min (resulting in nine estimates of relative proportions of reactions). In the other experiments (and their replicates), the frequency of reactions within a round of 30 min was sometimes rather low, so we pooled three consecutive rounds for a subset (resulting in six estimates per test situation).

The numbers of reactions and duration of occupancy by butterflies were summed for each subset and for each of the two dummy types tested in an experiment. Actual observation times for every dummy type were calculated by subtracting the duration of occupancy from the total observation time during a subset (3 dummies × 30 min = 90 min, or 3 dummies × 90 min = 270 min). To account for variable actual observation times, the numbers of reactions scored during a subset were standardised to a total observation time of 90 min, or 270 min, to yield corrected numbers of reactions. From these values, we calculated the proportions of reactions to the two dummy types in each subset.

Mate-choice experiments in the field

We tested the effects of the flavonoid content of female dummies on their attractiveness for wild free-flying males. Experiments were performed as described by Burghardt et al. (Burghardt et al., 2000) with dummies made from females reared on the artificial diets detailed above.

Results

Wing colours in flavonoid-free and flavonoid-rich butterflies

To the human eye, the uppersides of the wings of flavonoid-rich and flavonoid-free butterflies were indistinguishable. On the undersides of the wings of flavonoid-rich butterflies, the white and background coloration often appeared to be slightly more creamy or yellowish (Fig. 2A,B). This effect was not very pronounced, however, and usually only found when comparing larger series of such specimens. In the ultraviolet range, in contrast, distinct differences became obvious (Fig. 2C,D). Flavonoid-free individuals reflected much more ultraviolet light than flavonoid-rich individuals. In flavonoid-rich individuals, the white patterning sometimes disappeared almost completely, thereby altering the overall appearance greatly. The difference between the orange and black spots and the background coloration was mostly lost. A gradient of increasing ultraviolet reflectance from anterior to posterior occurred in the forewings. Living animals will appear even more disparate in the ultraviolet range than shown in our photographs because the posterior parts of the undersides of the forewings are usually hidden behind the hindwings.

These findings are corroborated by the reflectance spectra (Fig. 3). The spectra of the dark brown uppersides of the wings increased monotonically with wavelength (Fig. 3F). The orange spots showed very low and flat reflectance spectra up to about 500 nm with a sharp increase in reflectance at higher wavelengths and a short drop at approximately 660 nm (Fig. 3D,E). The mean spectra for both food treatments and their confidence intervals were almost the same except for the orange spots from the undersides of the wings (Fig. 3D–F). In these, reflectance was lower in the ultraviolet range
(300–400 nm) in flavonoid-rich animals (Fig. 3D). This difference was significant, as indicated by non-overlapping confidence intervals, but very small in absolute value (maximum approximately 1%). The reflectance spectra of white spots (Fig. 3A) distinctly differed between food treatments. The spectral reflectance of the white spots of flavonoid-free animals was more or less flat in the visible range (400–690 nm) and decreased slowly at shorter wavelengths. In contrast, the reflectance of the white spots of flavonoid-rich animals was very low in the ultraviolet region but increased drastically from approximately 400 nm upwards to reach the values of flavonoid-free animals at approximately 450 nm. At even longer wavelengths, it increased slightly, to exceed the values of the flavonoid-free animals above approximately 620 nm. The spectra of the background coloration (Fig. 3B) monotonically increased, those of flavonoid-free animals being flatter. The background spectra of flavonoid-rich animals remained very low in the ultraviolet region but exceed those of flavonoid-free animals above 420 nm. A similar pattern was found for the spectra of the black spots (Fig. 3C), but with much lower absolute reflectance values.

Mean reflectance as a measure of the ‘brightness’ of wing colours did not differ between food treatments for the orange spots and brown upperside coloration (Table 1). One would have expected mean reflectance to be lower in flavonoid-rich animals because of the absorption of light by flavonoids. Surprisingly this was true only for the white spots, which had a greatly reduced reflectance below approximately 450 nm, particularly in the ultraviolet region. The opposite was true for the black spots and the background coloration because the lower ultraviolet reflectance was more than compensated by a higher reflectance at longer wavelengths. For the black spots, the difference, although significant, was very small in absolute value (0.4%).

The correlation matrices of all principal component analyses were ill-conditioned, i.e. in these correlation matrices some variables were highly intercorrelated and therefore redundant (StatSoft, 1999b). This is not surprising because all the reflectance spectra were very smooth. To be able to invert an ill-conditioned correlation matrix, a step necessary for PCA, the statistics program used (StatSoft, 1999a) reduces slightly the correlation of the variables until it is able to invert the correlation matrix. This poses no problem, because the pattern of the PCs will remain essentially the same, although the resulting estimates are not exact (StatSoft, 1999b).

Eigenvalues for the first few PCs were very high, i.e. they
explained most of the between-spectra variation in spectral shape. All PC1s extracted more than 67% of the variation, and the first three PCs together always accounted for more than 95% of the variation (Table 2).

The coefficients of PC1 were usually highest at short and long wavelengths and lowest, and often negative, in the intermediate wavelength range (Fig. 4). PC1 therefore provides some differential measure of short plus long versus intermediate wavelengths. Coefficients of PC1 and PC2 frequently varied in opposing directions as a function of wavelength. However, at very short or long wavelengths, sometimes the values were almost the same (Fig. 4A,B). The absolute values of the coefficients of PC3 were always larger than those of PC1 and PC2, sometimes by an order of magnitude (Fig. 4D–F). We were unable consistently to relate shape or variation of reflectance spectra to PC coefficients. Hence, the physical meaning of the PCs remains unclear.

Table 1. Mean spectral reflectance of the wing colours of female Polyommatus icarus reared on a flavonoid-free (F–) or a flavonoid-rich (F+) diet

<table>
<thead>
<tr>
<th>Wing side and colour</th>
<th>Mean reflectance (%)</th>
<th>Means, F–/F+</th>
<th>F</th>
<th>d.f.</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Underside</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>43.7/36.1</td>
<td>51.91</td>
<td>1, 53</td>
<td>0.0001</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>3.9/4.3</td>
<td>6.13</td>
<td>1, 53</td>
<td>0.0165</td>
<td></td>
</tr>
<tr>
<td>Orange</td>
<td>14.8/14.8</td>
<td>0.03</td>
<td>1, 53</td>
<td>0.8735</td>
<td></td>
</tr>
<tr>
<td>Background</td>
<td>13.0/16.3</td>
<td>34.50</td>
<td>1, 52</td>
<td>0.0001</td>
<td></td>
</tr>
<tr>
<td>Upperside</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brown</td>
<td>5.7/5.4</td>
<td>0.87</td>
<td>1, 42</td>
<td>0.3561</td>
<td></td>
</tr>
<tr>
<td>Orange</td>
<td>14.2/14.0</td>
<td>0.15</td>
<td>1, 31</td>
<td>0.7011</td>
<td></td>
</tr>
</tbody>
</table>

F, statistics of repeated-measures ANOVA; d.f., degrees of freedom.

Fig. 4. First three principal component (PC) coefficients from analyses of the shape of reflectance spectra (see Fig. 3) of the hindwings of female Polyommatus icarus. Solid lines, PC1; dashed lines, PC2; dotted lines, PC3. (A) White, underside. (B) Background coloration, underside. (C) Black, underside. (D) Orange, underside. (E) Orange, upperside. (F) Brown, upperside. For further details see Fig. 3.
Flavonoids, butterfly wing colours and mate choice

In spite of this difficulty, we analysed PC scores for differences between food treatments. MANOVA of the first three PCs indicated no significant differences in the orange spots and brown coloration of wing uppersides (Table 2). In contrast, the spectral shape of all colours of the undersides of the wings were highly significantly different between food treatments. For the orange spots, this difference was due to PC3, which accounted for only 3.24 % of the variation. For the other colours, a difference was found in PC1, which explained most of the variation, and at least in one other PC.

We compared confidence limits of the calculated quantum catch for hypothetical photoreceptors (Fig. 5). As expected from the analyses presented above, values for the ultraviolet receptor (P360) were significantly different for all colours of the undersides of the wings, as were values for the blue receptor (P437) for white spots. In the background coloration of the undersides of the wings, the blue receptor exhibited no significant difference but the green (P500) and red (P568) receptors did. The values for the red receptor were significantly different for the black spots, too.

**Mate-choice experiments in the laboratory**

In the mate choice experiments in the laboratory, responses of males to female dummies were frequently observed. Male ‘approaches’ occurred most often, whereas ‘fluttering’ and copulation attempts happened but were very rare. We therefore pooled all male reactions to dummies of a given type (i.e. food treatment, size or external alteration of wing colours). There was no size preference for either larger or smaller dummies (proportion for large dummies 0.51±0.04, mean ± S.D.; N=9; Table 2).

### Table 2. Results of repeated-measures MANOVA and subsequent ANOVA of the first three principal component scores for the spectral shape of wing colours comparing female Polyommatus icarus reared on a flavonoid-free or a flavonoid-rich diet

<table>
<thead>
<tr>
<th>Wing side and colour</th>
<th>MANOVA PC1–PC3&lt;sup&gt;a&lt;/sup&gt;</th>
<th>PC1&lt;sup&gt;b&lt;/sup&gt;</th>
<th>PC2&lt;sup&gt;b&lt;/sup&gt;</th>
<th>PC3&lt;sup&gt;b&lt;/sup&gt;</th>
<th>PC1–PC3&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Wilks’ λ</td>
<td>P</td>
<td>d.f.&lt;sup&gt;d&lt;/sup&gt;</td>
<td>F</td>
<td>P</td>
</tr>
<tr>
<td><strong>Underside</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>0.433</td>
<td>0.0001</td>
<td>3, 51</td>
<td>301.62</td>
<td>0.0001</td>
</tr>
<tr>
<td>Black</td>
<td>0.486</td>
<td>0.0001</td>
<td>3, 51</td>
<td>46.67</td>
<td>0.0001</td>
</tr>
<tr>
<td>Orange</td>
<td>0.698</td>
<td>0.0003</td>
<td>3, 51</td>
<td>2.24</td>
<td>0.1401</td>
</tr>
<tr>
<td>Background</td>
<td>0.279</td>
<td>0.0001</td>
<td>3, 50</td>
<td>107.00</td>
<td>0.0001</td>
</tr>
<tr>
<td><strong>Upperside</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brown</td>
<td>0.941</td>
<td>0.4797</td>
<td>3, 40</td>
<td>0.17</td>
<td>0.6758</td>
</tr>
<tr>
<td>Orange</td>
<td>0.994</td>
<td>0.9804</td>
<td>3, 28</td>
<td>0.18</td>
<td>0.6732</td>
</tr>
</tbody>
</table>

<sup>a</sup>Statistical results for overall design with first three principal components (PC). Null hypothesis: no difference between food treatments.

<sup>b</sup>Statistical results for ANOVA on individual PC (single effect larval diet).

<sup>c</sup>Variation is the percentage of total variation explained by this PC.

<sup>d</sup>d.f., degrees of freedom.

---

Fig. 5. Calculated relative quantum catch with 95% confidence intervals for hypothetical lycaenid photoreceptors based on the averaged reflectance spectra of wing colours and their 95% confidence intervals. (A) Flavonoid-free females. (B) Flavonoid-rich females. Data for both categories can be compared directly because they were normalised to their common maximum. The ordinate is the quantum catch (arbitrary units). The abscissa shows wing colours. P360–P568, rhodopsin nomograms with maximum sensitivity at 360 nm (ultraviolet), 437 nm (blue), 500 nm (green) and 568 nm (red), based on data from Lycaena (Bernard and Remington, 1991). An asterisk indicates non-overlapping confidence intervals and, therefore, statistically significant differences in quantum catch between food treatments.
In all experiments where a choice was possible between flavonoid-rich and flavonoid-free dummies, more male responses were scored for flavonoid-rich female dummies (Table 3). This preference was found to be statistically significant irrespective of the posture type of the dummies. The proportions of the reactions to the flavonoid-rich dummies differed to some degree between the posture types, however. When only the undersides of the wings were visible, 60% of the reactions were scored for the flavonoid-rich dummies. This proportion was 57% for the dummies with both wing sides visible, and only 53% for the dummies that had the undersides of the wings mostly hidden by a flower dummy.

We sprayed flavonoid-free dummies with a solution of the flavonoid rutin to alter the coloration of the undersides of the wings so that they resembled naturally flavonoid-rich dummies. These again elicited highly significantly more reactions from males than the flavonoid-free control group (Table 3).

**Mate-choice experiments in the field**

In the field, we presented female dummies to wild, free-flying males. Of the males passing a dummy within a radius of 15 cm, 51.1% showed a noticeable response, while 49.9% ignored the dummies (both flavonoid-rich and flavonoid-free). The fraction of by-passing males did not differ between flavonoid-rich and flavonoid-free dummies (Fig. 6). Hence, males were not more attracted to either dummy category. There was, however, a very clear difference between the two dummy types in the mode of reaction of responsive males. At flavonoid-rich dummies, 65.3% of the male responses ended in the more intense fluttering behaviour, but this happened at only 41.6% of the flavonoid-free dummies. Therefore, males were more interested in flavonoid-rich dummies, spending more time and energy there (highly significant after Bonferroni–Hochberg correction for two dependent tests on the same data set).

**Table 3. Results of mate choice experiments**

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Corrected number of reactions1</th>
<th>Proportion2</th>
<th>Test statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N3</td>
<td>r²</td>
<td>P</td>
</tr>
<tr>
<td>Wings at 90 °</td>
<td>322.3</td>
<td>0.57±0.06</td>
<td>6 2.63 0.0467</td>
</tr>
<tr>
<td>Wings fully open</td>
<td>2867.5</td>
<td>0.53±0.01</td>
<td>6 6.16 0.0016</td>
</tr>
<tr>
<td>Wings closed</td>
<td>3316.1</td>
<td>0.60±0.05</td>
<td>6 5.07 0.0039</td>
</tr>
<tr>
<td>Wings sprayed</td>
<td>3260.1</td>
<td>0.58±0.04</td>
<td>9 5.76 0.0004</td>
</tr>
</tbody>
</table>

For details, see text.

1 Sum of the corrected numbers of reactions of the subsets.

2 Mean of the proportions of the subsets (values are means ± s.d.). The value of the proportion for flavonoid-free dummies is given by 1 minus this proportion.

3 Number of subsets.

4 Statistics of two-sided t-test for dependent data sets.

$t=1.04; P=0.33$). In all experiments where a choice was possible between flavonoid-rich and flavonoid-free dummies, more male responses were scored for flavonoid-rich female dummies (Table 3). This preference was found to be statistically significant irrespective of the posture type of the dummies. The proportions of the reactions to the flavonoid-rich dummies differed to some degree between the posture types, however. When only the undersides of the wings were visible, 60% of the reactions were scored for the flavonoid-rich dummies. This proportion was 57% for the dummies with both wing sides visible, and only 53% for the dummies that had the undersides of the wings mostly hidden by a flower dummy.

**Discussion**

Thus far, sequestration of flavonoids by butterflies has been studied largely from a chemical perspective. Our study reveals that flavonoids have a profound influence on the wing coloration of the common blue *Polyommatus icarus*, particularly in the ultraviolet range. Flavonoids sequestered from the larval diet enhance ultraviolet absorption of the undersides of the wings, although no difference was found in the uppersides of the wings. The diets used in this study differed only in flavonoid content, while they were identical with respect to all other compounds. One would have expected the wing colours of flavonoid-rich animals to have a lower mean reflectance because of the absorption of light by flavonoids. However, this was true only in the white spots, while the opposite was the case for the black spots and background coloration (Table 1). In the background coloration and the black spots of flavonoid-rich butterflies, reflectance was lower in the ultraviolet range, as expected, but was higher in the visible range (Fig. 3B,C). Therefore, a factor other than ultraviolet light absorption by flavonoids may also contribute to the difference in wing colours between the food treatments. The wings of flavonoid-free butterflies appear more greyish, as expressed in the flatter reflectance curve for the background coloration. The enhanced production of black eumelanin, which has a flat absorption spectrum (Nijhout, 1991), in the wing scales during scale development may be responsible for this finding, suggesting some coupling between *de novo* synthesis of autochthonous pigments and the deposition of allochthonous pigments in colour production during scale development. A ‘lack’ of allochthonous flavonoids, sensed in some way, might enhance eumelanin production. If a certain amount of ultraviolet absorption, especially in the background coloration, were desirable for a butterfly more eumelanin would compensate for the reduced light absorption caused by the lack of flavonoids. Unlike flavonoids, black eumelanin will greatly alter wing reflectance in the visible range, too. We did not investigate or test this idea.

Although shown only for females in this study, flavonoids...
influence wing colours in both sexes of this species and for a broad range of natural diets (H. Knüttel, unpublished results). Hence, there is a large intraspecific variability in wing coloration due to individual ontogenesis in this butterfly species. Flowers and buds are the preferred diet of *P. icarus* and more flavonoids are generally sequestered from these parts of the plant (Burghardt et al., 1997; Burghardt et al., 2001). Larvae grow faster and larger and beneficial myrmecophilous interactions with ants are more strongly developed when they are fed protein-rich flowers rather than foliage (Burghardt and Fiedler, 1996). The amount of flavonoids sequestered may, under natural conditions, be correlated with the quality of the food, which may in turn relate to higher quality as a potential mate. In butterflies, both the body size and the protein content derived from larval nutrition typically indicate high female fecundity (Honek, 1993; Wheeler, 1996). In *P. icarus*, flavonoid content and concentration is generally positively related to body mass, further strengthening the idea of a link between flavonoid sequestration during ontogenesis and fitness-related characters (Burghardt et al., 2001).

In laboratory experiments, we found that males indeed preferred flavonoid-rich female dummies. It seems very probable that this preference is mediated visually and not by chemoreception. Flavonoids, and their glycosides in particular, are not volatile (Harborne, 1991), and we attempted to evaporate any other volatile substances from our dummies prior to the behavioural experiments. We found differences in wing colour only in the undersides of the wings. However, a male preference for flavonoid-rich dummies was found irrespective of which sides of the wings were visible. In the dummies with the wings fully spread, the undersides of the wings were not completely hidden, and body coloration, which is also influenced by flavonoids (Fig. 2), may have served as an additional cue to the males. In these dummies, the proportion of male reactions to the flavonoid-rich dummies was lower than in the dummies of the other posture types, in which the undersides of the wings were readily visible. Moreover, visual signals from the undersides of the wings seem to play the major role in mate recognition from a distance in many lycaenid butterflies (Douwes, 1976; Wago et al., 1976; Pellmyr, 1982). Therefore, it seems likely that visual signals from the wing undersides are mainly responsible for the males’ preference for flavonoid-rich female dummies.

Still, it is possible that flavonoids in the diet may trigger the production of, or serve as precursors of other unknown substances that are sensed and preferred by the males. So, in another experiment, we used artificially altered dummies that were originally flavonoid-free but had the undersides of their wings sprayed with the flavonoid rutin frequently found in *P. icarus*. In these dummies, any potential physiological change in chemical composition due to the sequestration of flavonoids during ontogenesis was precluded. Rutin-sprayed, ultraviolet-absorbing dummies were again highly preferred, eliminating the possibility of olfactory cues being the stimuli mediating male preference.

We then tested male preference in the field to avoid possible limits imposed by indoor behavioural experiments. In spite of the presence of living females in the natural habitat, the males frequently responded to the dummies. Neither flavonoid-rich nor flavonoid-free dummies attracted more males, but males spent more time and energy closely inspecting the flavonoid-rich dummies. Exactly the same results were obtained when using flavonoid-rich or flavonoid-poor dummies reared on natural host-plants (Burghardt et al., 2000).

From these results, we conclude that, once having detected a female, males of *P. icarus* are more strongly attracted to flavonoid-rich females. This preference is mediated visually and is probably due mostly to the colours of the undersides of the wings. As flavonoid-induced colour differences are most pronounced in the ultraviolet range, this seems to be the spectral range most important to this behaviour. The ultraviolet coloration of undersides of the wings is to a large degree environmentally determined, and intraspecific variation is enormous. It is not therefore a species-specific characteristic, as has been assumed previously and used taxonomically (see Knüttel and Fiedler, 2001; for *Gonepteryx* spp. see also Brunton and Majerus, 1995).

Our study is the first case known to us in which secondary plant compounds have been shown to alter the external appearance of an insect and where this signal is used in visually mediated mate choice. In contrast, this phenomenon is well known in vertebrates, in which carotenoids are of special importance in this context (e.g. Grether et al., 1999; Blount et al., 2000).

One of the goals of this study was to evaluate the statistical analysis of spectra by PCA and MANOVA as introduced by Cuthill et al. (Cuthill et al., 1999) and to compare this approach with the use of mean spectra and their confidence intervals, used here for the first time. It soon became apparent that PCA of spectra in combination with MANOVA is a very sensitive tool in grouping spectra, as Cuthill et al. (Cuthill et al., 1999) found in an analysis of iridescent bird coloration. This method has the advantage of being independent of any assumptions about the visual system of the receiver of a signal. A serious drawback of PCA is the lack of any straightforward explanation of the meaning of the principal components and that this meaning is a different one for every set of data analysed. It is not possible to compare PCs resulting from different studies (Endler, 1990; Cuthill et al., 1999). The addition of a few spectra to a set of data may completely change the result of a PCA (Endler, 1990).

We do, however, consider it valuable not only to be able to show that there is a difference between groups of spectra, but also to be able to locate the spectral origin of this difference precisely. The PC coefficients are the values used to transform the original data into the PCs. From these, it is possible to attempt to identify the meaning of a PC. However, the ‘spectra’ of PC coefficients often are quite complex in shape, even for very smooth and straight (reflectance) spectra (Fig. 3, Fig. 4), so this task is not trivial. We did not accomplish very much with our data. The situation becomes even more complicated when one tries to take into account the factor loadings, which
are the correlations between the original variables and the PCs. Here, the picture becomes even worse. One not only really does not know what a PC actually means, but there is no mathematical justification stating how many of the PCs, and therefore how much of the variation in the data, should be included in the subsequent statistical analysis. However, the MANOVA results may depend on this choice.

There are only rules of thumb as to how many PCs to include in later analyses. Either PCs are considered with an eigenvalue larger than 1 or that account for more than 5% of total variation (Kaiser, 1960; Chatfield and Collins, 1980; StatSoft, 1999b). The decision is arbitrary and depends on the experimenter. Biologically relevant information may, however, be found in a PC that accounts for a very small amount of the variation and is therefore not considered for statistical analysis (Cuthill et al., 1999). We decided a priori to always include the first three PCs in our analyses. If we had chosen to include only PCs accounting for more than 5% of the variation, we would have had to leave out the PC3 in all analyses except for the black spots. Thus, we would have missed a significant difference between food treatments in the orange spots of the undersides of the wings. Although very small in absolute value, this difference was found and located in the spectrum simply by looking at the confidence limits of the mean spectra of these spots.

The wavelength resolution of the spectra and, therefore, the number of variables used for PCA need to be considered in this context. The eigenvalue of a PC is the fraction of total variation it explains multiplied by the number of variables (StatSoft, 1999b). The effect of the decision to include PCs with an eigenvalue larger than 1 then depends heavily on the wavelength resolution chosen.

As a result of all these complications involved in PCA and MANOVA, we found the use of mean spectra and their confidence intervals very valuable. The calculation is easily available his ultraviolet photography equipment to us. We are grateful to Klaus Lunau and two unknown referees for helpful comments on an earlier version of the manuscript and to Stuart Church and the copy editor for substantial linguistic improvements. We thank Bernd Kormmaier, Wolfgang Weith and Annick Servant, who were of great help in breeding the butterflies. H.K. gratefully acknowledges financial support by the Konrad-Adenerauer-Stiftung.

We thank Louise Lieberknecht who undertook some of the behavioural experiments. Klaus Lunau generously made available his ultraviolet photography equipment to us. We are grateful to Klaus Lunau and two unknown referees for helpful comments on an earlier version of the manuscript and to Stuart Church and the copy editor for substantial linguistic improvement. We thank Bernd Kormmaier, Wolfgang Weith and Annick Servant, who were of great help in breeding the butterflies. H.K. gratefully acknowledges financial support by the Konrad-Adenerauer-Stiftung.

**References**


