

## RESEARCH PAPERS

# Trying to See Red Through Stickleback Photoreceptors: Functional Substitution of Receptor Sensitivities

Mickey P. Rowe\*, Charles L. Baube† & John B. Phillips‡

\* Neuroscience Research Institute, University of California, Santa Barbara, CA, USA

† Department of Biology, Oglethorpe University, Atlanta, GA, USA

‡ Department of Biological Sciences, Virginia Tech, Blacksburg, VA, USA

## Correspondence

Mickey P. Rowe, Neuroscience Research Institute, University of California, Santa Barbara, Santa Barbara, CA 93106-5060, USA.  
E-mail: mrowe@lifesci.ucsb.edu

Received: November 11, 2004

Initial acceptance: January 6, 2005

Final acceptance: May 12, 2005 (S. Forbes)

## Abstract

A key to understanding animal behavior is knowledge of the sensory information animals extract from their environment. For visually motivated tasks, the information animals obtain through their eyes is often assumed to be essentially the same as that perceived by humans. However, known differences in structure and processing among the visual systems of different animals clearly indicate that the world seen by each is different. A well-characterized difference between human and other animal visual systems is the number of types and spectral sensitivities of their photoreceptors. We are developing a technique, functional substitution, that exploits knowledge of these differences to portray for human subjects, colors as they would appear through the photoreceptors of another animal. In a specific application, we ask human subjects to rank hues of male threespine stickleback (*Gasterosteus aculeatus*) throats viewed through stickleback photopigments. We compare these ranks to ranks of the same throat hues viewed through normal human photoreceptors. We find essentially no difference between the two sets of rankings. This suggests that any differences in human and stickleback rankings of such hues would result from differences in post-receptoral neural processing. Using a previously developed model of stickleback neural processing, we established another ranking of the hues which was again essentially the same as the rankings produced by the human subjects. A growing literature indicates that stickleback do rank such hues in the evaluation of males as potential mates or threats. Although our results do not demonstrate that humans and stickleback use the same mechanisms to assess color, our experiments significantly failed to show that stickleback and human rankings of throat hues *should* be different. Nevertheless, a comparison of all these rankings to ranks derived from subjective color scoring by human observers suggests that color scoring may utilize other cues and should thus be used cautiously.

## Introduction

A pair of animals exposed to identical stimuli may perceive those stimuli quite differently. There are several causes for such disparate perceptions. For instance, there may be differences in the anatomical structures directing energy to the sensory receptors,

differences in the types and number of those receptors, and/or differences in the neural processing of the receptors' outputs. A challenge for biologists interested in understanding why animals behave as they do is the impenetrability of each animal's *Umwelt*, or perceptual world (von Uexküll 1957). To discern the causes of behavior, it is sometimes

necessary to determine how the sensory capabilities of study species differ from those of humans, and to interpret the way these differences affect perception of, and ultimately responses to, the surrounding environment (Hughes 1999).

Nowhere is the need for such determinations more apparent than in the study of vision (e.g. Bennett et al. 1994; Vorobyev et al. 1997; Fleishman et al. 1998; Fleishman & Endler 2000; Vorobyev et al. 2001; Kelber et al. 2003). Nevertheless, potential differences in perceptual capabilities are often assumed (or hoped) to be minimal in research on visual signaling, and human perceptual categories are sometimes used as the sole descriptors of stimuli (e.g. Kilner & Davies 1998; McRobert & Bradner 1998; Duckworth et al. 2003; Saino et al. 2003). Given the number of factors that sculpt perceptions, the presumption that human perception is an adequate substitute for that of another animal is dangerous at best. How can we know when human observers can be used fruitfully as surrogates for study subjects while investigating animal biology? One way to address the question is to design experiments that take advantage of known differences between human and nonhuman sensory capabilities and to quantify the effects of those differences.

Improvements in technology used to display visual information have enabled us to simulate particular aspects of the perceptions of other animals. For instance, we can present synthetic stimuli that generate in human subjects the same sensory neuronal activity that occurs in study animals perceiving natural stimuli. Such simulations may improve understanding of how animals process the information in physical stimuli to make decisions resulting in overt behavior. Here we use this method to explore color perception of the threespine stickleback, (*Gasterosteus aculeatus*) a species that is rapidly becoming a model system for the study of evolution in general (Bell & Foster 1994; Schluter 2000; McKinnon & Rundle 2002; Foster & Baker 2004), and of communication systems and behavioral evolution in particular (e.g. Tinbergen 1951; Rowland 1994; Foster 1995; Foster et al. 1998; Olson & Owens 1998; Huntingford 2003).

A feature of the threespine stickleback that has generated particular interest is the male nuptial signal. Although males exhibit mosaic nuptial coloration usually including – in human perceptual terms – a blue eye, red-orange throat, and a blue-green back (Bakker & Mundwiler 1994; Rowland 1994; McLennan 1996), it is the red element of the signal that has drawn the most attention. This signal

has been argued to attract females, to serve as an index of male condition, to inhibit territorial intrusions by adjacent males, and most recently, to function in assortative mating between members of stickleback species pairs (e.g. Rowland 1994 for review; Baube et al. 1995; Rowland et al. 1995a,b; McLennan 1996; Baube 1997; Boughman 2001; Smith et al. 2004). The sensory system of isolated freshwater populations of the threespine stickleback has also been suggested to evolve in response to environmental conditions, specifically as a consequence of signal masking by humic acids. Where humic acids make water opaque, stickleback male throats typically appear black rather than red (Reimchen 1989; McDonald et al. 1995; McKinnon 1995; Scott & Foster 2000), and there is evidence of a parallel shift in the visual system (Boughman 2001).

Within populations, male stickleback can exhibit considerable variation in the extent (area of fish and degree of color saturation) of the nuptial signal. This variation may be used by females as a criterion for mate choice, although the relationship has proven complex (e.g. Milinski & Bakker 1990; Bakker & Mundwiler 1994; Rowland 1994; Bolyard & Rowland 1996; Candolin 1999). Some of the complexity may be due to the way in which the development of the red coloration has been measured. The degree of redness (frequently referred to as 'brightness of red coloration' or just 'brightness' in the stickleback literature) has generally been assessed only through human judgements (e.g. Bakker 1986; McLennan & McPhail 1989; Milinski & Bakker 1990; Bakker & Milinski 1991). More recently, researchers have attempted to utilize objective means to assess coloration using spectrometric measurements of slide photographs (Wedekind et al. 1998) or digital photography (Candolin 1999, 2000; Braithewaite & Barber 2000). These methods rely indirectly upon the human visual system because both color film and the pixels of digital cameras record only the information that a standard human observer would extract upon viewing a captured scene. Consequently the methods may remove any effects of differences among human observers in the perception of stickleback colors, but they do not circumvent the more important potential effects of differences between human and stickleback perceptions of stickleback colors.

Colorimeters and cameras can produce objective determinations of whether or not one stickleback throat is 'redder' than another. However, these redness values are not necessarily any more relevant to stickleback perceptions than are more direct human

assessments (please note the distinction between photometers or colorimeters, and spectrometers or spectroradiometers; colorimeter and photometer designs incorporate properties of human visual systems). In the threespine stickleback literature, human assessments have been justified by a supposed similarity between the absorption spectra of human photopigments and the absorption spectra of stickleback photopigments (Frischknecht 1993; McKinnon 1995; Künzler & Bakker 2001). However, as shown in Fig. 1, there are potentially significant differences between human and stickleback pigments. Here we explore the effect of these pigment differences on the relative rankings of 'redness' of male stickleback throats.

## Methods

Stimuli were generated using functional substitution, a method previously described (Rowe & Jacobs 2004) and diagrammed in Fig. 2. More detail is provided in the Appendix. The method was modified from that of Vorobyev et al. (1997, 2001) who also displayed images depicting information other animals extract from biologically relevant visual stimuli. Similar methods have also been used to explore the visual worlds of horses (Carroll et al. 2001) and color-deficient humans (Pokorny & Smith 1977; Brettel et al. 1997). In the present study, stickleback throat colors were portrayed two different ways. In one form of presentation, we used human cone pigment absorption spectra (Smith & Pokorny 1975) to derive the settings for the computer monitor (steps 2–3 in Fig. 2) that displayed the colors. For the other presentation, stickleback cone pigment absorption

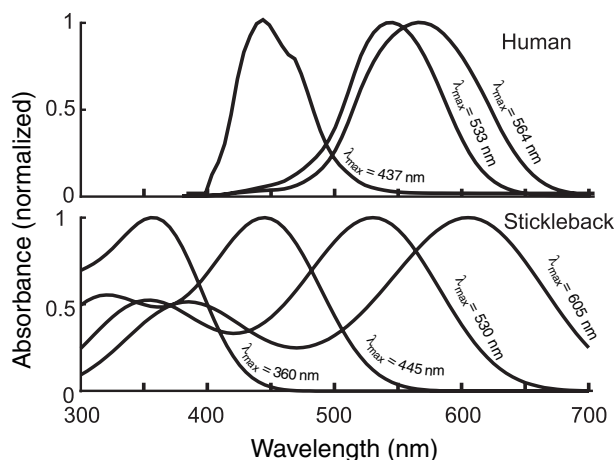
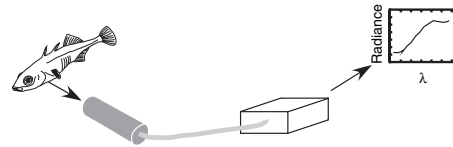
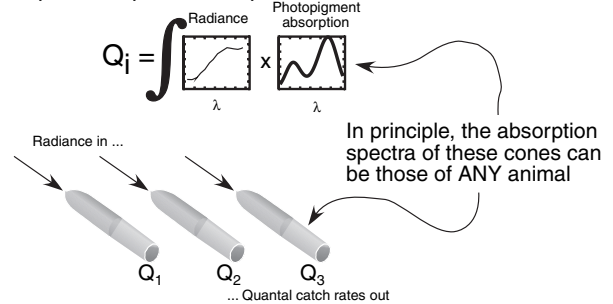


Fig. 1: Human and stickleback cone spectral sensitivities

### Step 1: Measure (or compute) light



### Step 2: Compute cone quantal catches



### Step 3: Compute monitor settings for human cones

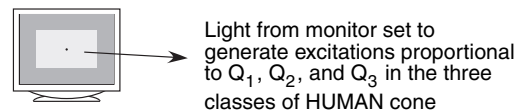


Fig. 2: Application of functional substitution of photoreceptor spectral sensitivities

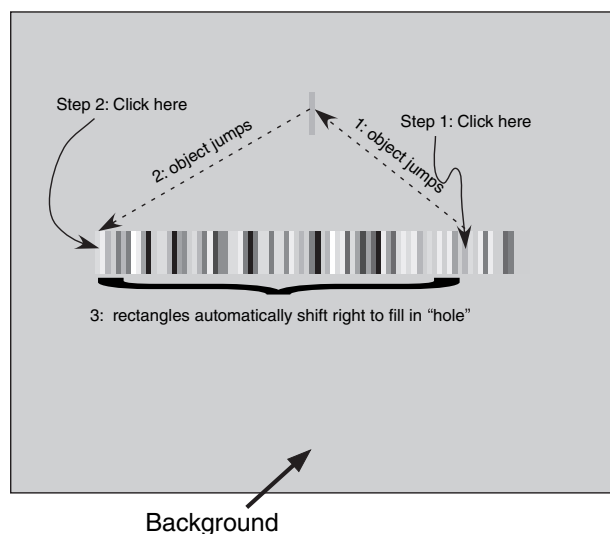
spectra (Rowe et al. 2004) were used to derive the monitor's settings. In the first form of presentation, colors were portrayed as they should appear to a human observing the throats directly. In the second form, colors were portrayed as they should appear to a human with stickleback visual pigments functionally substituted for their own.

Stickleback have four cone pigments, ultra-violet sensitive, short wavelength sensitive, middle wavelength sensitive, and long wavelength sensitive (Rowe et al. 2004). We here designate the four pigments UV, sS, sM, and sL. Humans have only three cone pigments, and we designate these hS, hM, and hL. Because of the mismatch in the number of cone pigments, it is not possible to substitute the entire complement of stickleback pigments for the available human pigments. However, our research suggests that the UV pigment, provides little if any information beyond that provided by sS, sM, and sL (Rowe et al. 2004). That is, given what we know of the variation in stickleback throat reflectances and what we can infer of the processing of color by stickleback, UV cone outputs do not help the animals discriminate male throat colors. Differences in throat

color among males are fully characterized by the outputs of the sS, sM, and sL cones. Therefore, we feel justified in ignoring UV cones in these experiments.

Our original stimuli were patches of throat color (the ventral surface of the fish between the opercula) from a population of stickleback that breed in tidal pools on the eastern shore of Long Island, NY, USA. The synthetic stimuli were derived from measurements of the reflectance of throats from 86 male fish and the illumination incident on the pools from which the animals were captured (Rowe et al. 2004). The simulated colors were presented to humans with normal color vision as assessed by Ishihara pseudoisochromatic plates, Hardy-Rand-Rittler polychromatic plates, and Rayleigh anomaloscopy. We generated stimuli with the aid of MATLAB® (the MathWorks, Inc., Natick, MA, USA) and the Psychophysics Toolbox (Brainard 1997; Pelli 1997) running on a Power Macintosh G3 (Apple Computer, Inc., Cupertino, CA, USA). Stimuli were presented on either a Diamond Pro 710 monitor (Mitsubishi Electronics America, Inc., Cypress, CA, USA) driven by a 10-bit video card (Radius, Inc., Sunnyvale, CA, USA), or a ViewSonic P95f+ (ViewSonic, Walnut, CA, USA) driven by a Radeon 9200 (ATI, Markham, Ontario, Canada). The monitor was calibrated with a SpectraScan® PR®-650 spectroradiometer (Photo Research, Inc., Chatsworth, CA, USA).

The human subjects' task was to rank the simulated colors according to their hue. Figure 3 contains a



**Fig. 3:** Psychophysical ranking procedure. This scaled diagram shows the spatial arrangement of the stimuli and steps taken to rearrange them during the performance of the psychophysical tasks

diagram indicating how the colors were sorted to generate the rankings (additional detail is provided in the Appendix). To assess repeatability, subjects sorted the patches two to five times for each of the two conditions (human photopigments and stickleback photopigments). Rankings of individual colors were compared across subjects and across the two conditions as well as within those categories. In both conditions, the background color, the color against which the array of patches was displayed, was derived from measurements of the average background radiance measured in the breeding pools (Rowe et al. 2004).

The fish measured to generate the colors were also subjectively rated according to criteria established by Rowland (1984, 1989) and Baube (1997). In brief, color scores were based on a scale of 1 (slight pale coloration around the mouth lining) to 5 (extensive, intense-red coloration ranging laterally and ventrally) in increments of 0.5. Color scores were determined independently by three to four observers for each fish.

We have predicted that stickleback should evaluate male throat colors neurophysiologically by comparing the outputs of sL and sS cones (Rowe et al. 2004). This prediction is based on our modeling of the processing of the information that stickleback photoreceptors extract from throat colors. Other comparisons are possible, but this comparison yields the largest difference in signals originating from stickleback throat colors. Here we computed the difference in excitations of a stickleback's sL and sS cones while the animal is viewing the 86 fish throats. We thereby generated a stickleback's predicted physiological ranking of the throat patches. The rankings made with the psychophysical tasks were compared with the rankings generated by color scores and with the rankings predicted from the presumed physiology to assess concordance among the various methods. Details on the method of comparison between psychophysical and physiological rankings are provided in the Appendix. Finally, we repeated the physiological ranking computations for comparisons of other visual pigments, both human and stickleback. These comparisons clarify how the subjects performed the task and how the pigment substitutions impacted performance.

## Results

A total of four reflectances were discarded because their representations were outside of monitor gamut for either or both versions of the psychophysical

experiment (see Appendix). The psychophysical rankings of the remaining 82 reflectances were quite consistent across trials irrespective of whether rankings were compared for an individual observer, across observers, or across pigment conditions. These results are summarized in Table 1. The correlations suggest that despite their differences, human and fish visual pigments are similar enough that with the same post-receptoral processing they provide the same relative rankings of redness of small patches of simulated throat color.

Based on their color scores, the measured fish represented a typical range of throat color variation for individuals from this Long Island population (Fig. 4). Individual judgements of fish color were strongly correlated across observers (Kendall's coefficient of concordance,  $W > 0.7$ ,  $p < 0.025$ ; Siegel & Castellan 1988). Human-perceived variation in stickleback throat color follows a roughly normal distribution. Color score ranking did not correlate well with psychophysical ranking of the redness of simulated throat patches even when the ranking was performed with quantal absorptions calculated for human cones (Fig. 5b; linear correlation coefficient,  $r = 0.50$ ).

However, the psychophysical data were concordant with our predicted physiological rankings determined by computing the difference in excitations of the sL and sS cones (Fig. 6). The normalized residual 'error' between the computed physiological signal and the means as computed from the psychophysical data are small and show no clear pattern (Fig. 6c). Comparison of different physiological models are rather illuminating. As shown in Table 2, it is clear that subjects' evaluations of the hues were derived

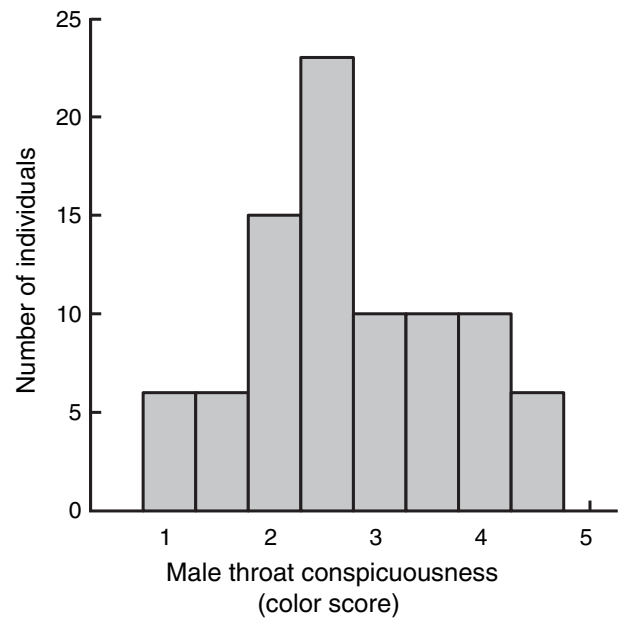


Fig. 4: Color score distribution of fish used for reflectance measurements

from comparisons of the outputs of their *own* L and M cones. The differences in the responses of hL and hM are highly correlated with the differences between the responses of sL and sS in all conditions for this particular set of throat colors.

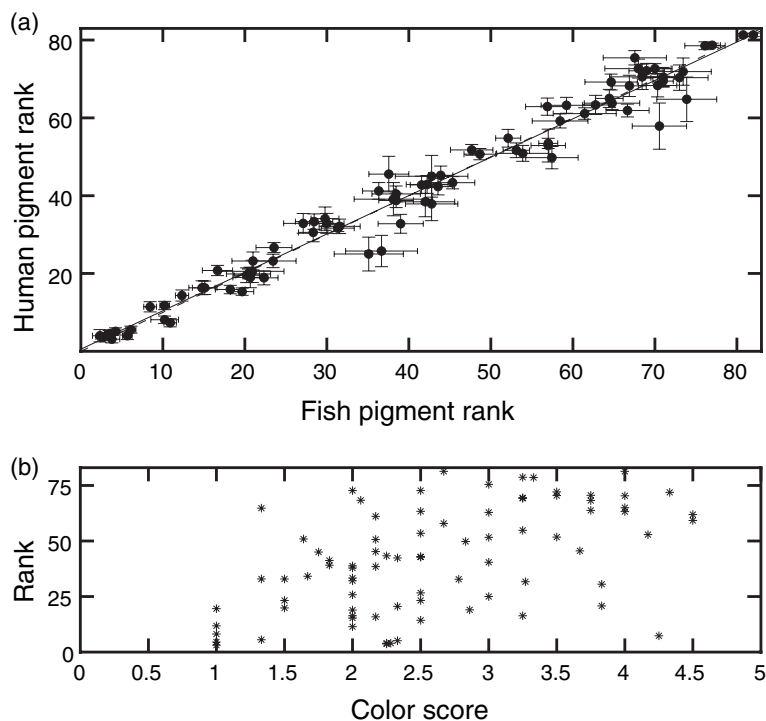
## Discussion

Our goal was to test the idea that differences between human and stickleback photopigments would result

Table 1: Correlations within and across subjects in psychophysical tasks

		Subject 1		Subject 2		Subject 3		All Subjects	
		Human (n = 5)	Fish (n = 5)	Human (n = 2)	Fish (n = 2)	Human (n = 2)	Fish (n = 2)	Human (n = 9)	Fish (n = 9)
Subject 1	Human	0.963	0.978	0.957	0.960	0.953	0.952	0.993	0.977
	Fish		0.944	0.957	0.972	0.965	0.977	0.983	0.997
Subject 2	Human			0.978	0.974	0.969	0.950	0.982	0.967
	Fish				0.966	0.955	0.961	0.975	0.984
Subject 3	Human					0.970	0.979	0.980	0.974
	Fish						0.966	0.970	0.987
All Subjects	Human							0.946	0.987
	Fish								0.945

Numbers along diagonal are the average Spearman's  $\rho$ -value (rank correlation coefficient) for all possible pairwise comparisons within a given data set (derived from Kendall's  $W$  computation; Howell 1997). All other elements are derived by averaging within-category ranks and computing Spearman's  $\rho$  across category. For subjects two and three along the diagonal, the results of these two computations are numerically identical because with  $n = 2$  there is only one comparison to be made.



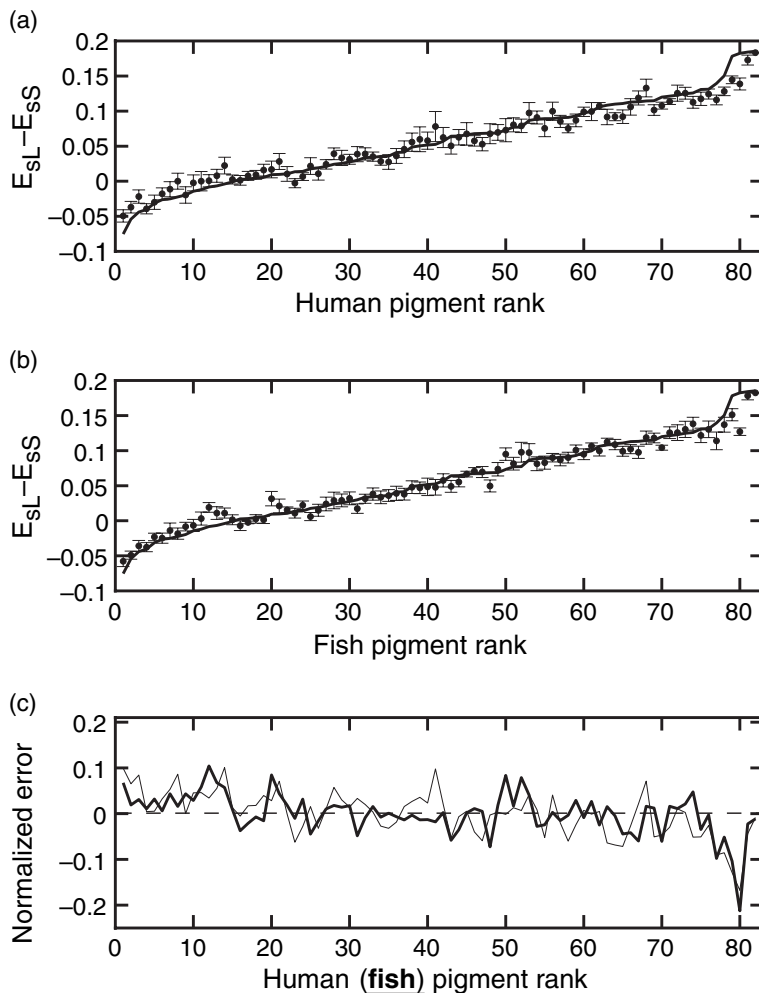
**Fig. 5:** Psychophysical rankings. In (a), mean and SE of ranks for the two types of psychophysical experiment are plotted against each other. Congruence between the two sets of rankings is so complete that at this scale the dashed line indicating a perfect correlation is obscured by the solid line indicating a linear fit to the data. In (b), mean ranks are plotted against the color score assigned to each fish

in differences between human and stickleback judgements of the red component of stickleback nuptial signals. Results presented in Table 2 suggest that such differences do occur but are more subtle than we had expected. They also provide valuable insight into the technique we used to ask the question. It is well accepted that human judgements of 'redness' are derived primarily from a comparison of the outputs of human L and M cones. If that was the criterion our subjects used to rank hues in this experiment, then the entries in bold should have been the largest rank correlation coefficients in their respective rows. And they are. This finding gives confidence that the experiment worked as intended. A second point of interest is that the correlations are higher when subjects viewed the hues through stickleback photopigments. This indicates (and subjects' verbal reports concur) that the task was easier under that condition. Consequently we can infer that it would be easier to rank order stickleback throat hues if we had stickleback photopigments instead of our own. Numerical comparisons of the cone responses (i.e. the numbers used to generate the stimuli for the psychophysical task) indicate that the hues should be more saturated for stickleback photopigments than for human photopigments.

A surprising result of our investigation was the discrepancy between the results of two different

methods for utilizing human vision to judge the animals (Fig. 5b). The reason for this discrepancy is currently unclear. In seeking an explanation, we highlight an important issue with respect to judging animal colorfulness. In one method, color scoring, the spatial extent of the colored region was taken into consideration. In the other method, ranking representations of the redness of single patches, spatial variation of color within an individual was ignored. If this is the main source of the discrepancy, then multiple spectrometric measurements must be made of each fish to capture more of the variation among fish. However, we do not currently know how much of this variation is sampled by fish as they evaluate each other.

An alternative explanation for the discrepancy revolves around another difference between the methods. In the psychophysical tasks reported here, subjects made simultaneous relative rankings of colors as opposed to sequential absolute judgements. Absolute judgements, as were performed for color scoring the fish, might generally be considered more difficult to make. The concordance among different observers scoring the fish, however, suggests that the best explanation for the discrepancy lies elsewhere. Agreement among color scorers suggests that different scorers responded similarly to differences between fish. The discordance between color scoring



**Fig. 6:** Comparing psychophysics to physiology. Circles and error bars are the mean physiological signal  $\pm$ SE when the ranks are sorted according to the psychophysical task with stimuli rendered for human visual pigments (a) or stickleback S, M, and L visual pigments (b) (see Appendix for additional details on computations). Lines in (a) and (b) are computed physiological signal in ascending order. To the extent that the two rankings are concordant, the data points should be on the lines. In (c), differences between data sets of (a) (thin line) and (b) (thick line) are shown. Physiological signals were normalized prior to computation of differences to facilitate comparison across data sets

**Table 2:** Spearman's rank correlation coefficients between different models of physiological processing and psychophysical rankings

		Physiological model			
		Human hL-hM	hL-hS	Stickleback sL-sM	sL-sS
Psychophysical experiment	Human	<b>0.975</b>	0.789	0.971	0.884
	Stickleback	0.989	0.841	<b>0.990</b>	0.926

Bold values indicate the item which should have the largest value in each row if psychophysical rankings were most closely correlated to the difference in output of the subject's own L and M cones.

and hue sorting suggests that the between fish differences noted by color scorers were not perfectly correlated with differences in the spectral reflectance of the throat patch sampled from each fish. However, it should be born in mind that there is also noise associated with spectrometric measurements as well as their interpretation in the hue-sorting task. None of the ranks proposed here, from color scoring,

from psychophysics, or from our physiological models can be considered an absolute standard against which all other rankings should be compared. Human color scorers, having experienced a range of appearances of fish, may use several dimensions as they assign fish to categories. Such experience improves consistency among judgements. Experienced raters use multiple sources of information with various degrees of correlation and thus make reliable absolute judgements (Wickens & Hollands 2000).

It should also be noted that we cannot guarantee that fish reflectances did not change between color scoring and spectrometric measurements. The two types of data were acquired at roughly the same time in order to minimize the effects of color changes. Such changes do present another possible explanation for the discrepancy, however.

In any case, to understand stickleback behavior, we must develop suitable methods to quantify color-



fulness. A major difficulty with the reliance upon color scoring is that we do not know whether or not the cues used by human color scorers are the same as the cues used by stickleback examining each other. A first step toward addressing this issue is asking whether stickleback and human observers might agree on the relative rankings of the palette of nuptial colors. There is no way to determine a priori the extent to which human and stickleback judgements of coloration should be similar. Commercially available digital cameras (e.g. Candolin 1999) cannot be used to resolve this issue. Cameras designed to reproduce images for humans do only that; they do not reproduce images as they would appear to other animals. Spectral measurements of the light reflected from fish circumvents this problem but also introduces a new one. It is not currently practical to measure reflectance over the entire body of an animal.

So what strategies should we employ in our efforts to quantify animal coloration? Ultimately we need to understand what information the animals extract with their visual systems and how that information is processed to lead to overt behaviors. As a practical issue, we cannot record all of the information available to the animals during their evaluations of each other. Therefore we must use judgement to determine what data to collect and then how to process it. The vast majority of equipment built for collecting information from light is designed for humans. Therefore, one path that needs to be explored is the adequacy of such technology for acquiring data relevant to the perceptions of other animals. There are two ways to address this issue, both of which we have begun to investigate here for the case of stickleback nuptial signals. First we can ask how our own evaluations of fish would change if our perceptions were mediated by stickleback sensory anatomy and physiology (Fig. 5a). Essentially we can make a chimaera, an observer that is functionally part fish and part human. The closer we can get to providing human observers with the sensory information used by the stickleback, the closer we can come to understanding the choices stickleback make. Second, we can use anatomical and physiological data to model the information processing carried out by the stickleback nervous system, and ask how judgements based on the results of such processing compare to the results of processing by the human nervous system (Fig. 6).

Based upon our current results, we recommend that researchers use a combination of spectrometry and digital or photo imaging to quantify stickleback color. Spectrometric measurements are favored because they come closest to the ideal of represent-

ing all data available to the fish. Until and unless it becomes possible to quickly acquire such data for all points on a fish's body, however, spectral reflectance data should be supplemented with images acquired under a controlled setting. In both cases, measurements should be standardized. For instance, the reflectance data used for this study were all acquired from the ventral surface of the fish inferior to the middle of the opercula. As the technology improves and the speed with which such measurements can be made increases, more locations should be sampled. New or additional locations should be anatomically well defined.

Spectral data have the advantage that they can be reanalyzed as more is learned about stickleback visual processing. Subjective color scoring of fish does not provide this benefit. Digital imaging falls somewhere between these extremes. The results presented in this paper suggest that for the red component of stickleback nuptial signals, commercial digital cameras can provide records useful for predicting how stickleback would rank the males under study. That is, the differences between human and stickleback cone pigment absorption spectra do not severely impact relative judgement by humans of stickleback throat color (Fig. 5a), and these judgements are strongly correlated with our simple physiological model's prediction of how color would be ranked by stickleback (Fig. 6). Consequently, we can predict from our data set that human and stickleback judgements of male 'redness' can be considered similar. We caution that this conclusion cannot be extended beyond the red component of stickleback color, and should only be extended warily beyond the data set used for this analysis. That is, the spectral reflectance functions of the throats of stickleback from other populations may differ from those of the animals in our study population in ways that are apparent to stickleback but not necessarily to humans. We would like to emphasize that our conclusions for human evaluations of male stickleback throat coloration cannot be generalized beyond this species and this component of male nuptial coloration. Further work is needed both to compare human and stickleback responses to other components of male stickleback nuptial coloration (e.g. iris color) and to compare relative rankings of color signals in other species, to determine if general patterns emerge. For example, are human and nonhuman assessments of hue likely to be more congruent for reflectances that primarily vary in particular regions of the spectrum? The overall goal is to explore ways to test the adequacy of human perception as a



surrogate for the perception of nonhuman animals; both to facilitate research on the function and evolution of color signals, and to gain new insight into the perceptual consequences of differences in visual system design. We hope that we have highlighted the potential problems as well as provided guidance in approaches to solving those problems.

### Dedication

We would like to dedicate this paper to the memory of Bill Rowland who inspired in all of us our first interest in stickleback.

### Acknowledgements

We thank Susan Foster for helpful comments on the manuscript, and Jerry Jacobs and David Brainard for discussion and support during its writing. The manuscript was significantly improved by the comments of two anonymous reviewers. We also thank our subjects for their participation. This work was supported by Sigma Xi, the Research Society, the small-grant programs of the Indiana University National Science Foundation Research Training Group for the Integrative Study of Animal Behavior (CL Baube), the Research Training Group in Animal Behavior (NSF Grant no. BIR 9413220), NIH grants EY10016 (awarded to D.H. Brainard), EY002052 (awarded to G.H. Jacobs) and F32 EY7077 (MP Rowe).

### Literature Cited

- Bakker, T. C. M. 1986: Aggressiveness in sticklebacks (*Gasterosteus aculeatus* L.): a behaviour-genetic study. *Behaviour* **98**, 1–144.
- Bakker, T. C. M. & Milinski, M. 1991: Sequential female choice and the previous male effect in sticklebacks. *Behav. Ecol. Sociobiol.* **29**, 205–210.
- Bakker, T. C. M. & Mundwiler, B. 1994: Female mate choice and male red coloration in a natural threespine stickleback (*Gasterosteus aculeatus*) population. *Behav. Ecol.* **5**, 74–80.
- Baube, C. L. 1997: Manipulations of signaling environment affect male competitive success in threespine sticklebacks, *Gasterosteus aculeatus*. *Anim. Behav.* **53**, 819–833.
- Baube, C. L., Rowland, W. J. & Fowler, J. B. 1995: The mechanisms of colour-based mate choice in female threespine sticklebacks – hue, contrast, and configurational cues. *Behaviour* **132**, 979–996.
- Bell, M. A. & Foster, S. A. 1994: Introduction to the evolutionary biology of the threespine stickleback. In: *The Evolutionary Biology of the Threespine Stickleback* (Bell, M. A. & Foster, S. A., eds.) Oxford Univ. Press, New York, pp. 1–27.
- Bennett, A. T. D., Cuthill, I. C. & Norris, K. J. 1994: Sexual selection and the mismeasure of color. *Am. Nat.* **144**, 848–860.
- Bolyard, K. J. & Rowland, W. J. 1996: Context-dependent response to red coloration in stickleback. *Anim. Behav.* **52**, 923–927.
- Boughman, J. W. 2001: Divergent sexual selection enhances reproductive isolation in sticklebacks. *Nature* **411**, 944–948.
- Brainard, D. H. 1997: The psychophysics toolbox. *Spat. Vis.* **10**, 433–436.
- Braithewaite, V. A. & Barber, I. 2000: Limitations to colour-based sexual preferences in three-spined sticklebacks (*Gasterosteus aculeatus*). *Behav. Ecol. Sociobiol.* **47**, 413–416.
- Brettel, H., Viénot, F. & Mollon, J. D. 1997: Computerized simulation of color appearance for dichromats. *J. Opt. Soc. Am. A* **14**, 2647–2655.
- Candolin, U. 1999: The relationship between signal quality and physical condition: is sexual signalling honest in the three-spined stickleback? *Anim. Behav.* **58**, 1261–1267.
- Candolin, U. 2000: Increased signalling effort when survival prospects decrease: male-male competition ensures honesty. *Anim. Behav.* **60**, 417–422.
- Carroll, J., Murphy, C., Neitz, M., Ver Hoeve, J. & Neitz, J. 2001: Photopigment basis for dichromatic color vision in the horse. *J. Vis.* **1**, 80–87, <http://journalofvision.org/1/2/2/>, DOI 10.1167/1.2.2.
- Duckworth, R. A., Badyaev, A. V. & Parlow, A. F. 2003: Elaborately ornamented males avoid costly parental care in the house finch (*Carpodacus mexicanus*): a proximate perspective. *Behav. Ecol. Sociobiol.* **55**, 176–183.
- Fleishman, L. J. & Endler, J. A. 2000: Some comments on visual perception and the use of video playback in animal behavior studies. *Acta Ethol.* **3**, 15–27.
- Fleishman, L. J., McClintock, W. J., D'earth, R. B., Brainard, D. H. & Endler, J. A. 1998: Colour perception and the use of video playback experiments in animal behaviour. *Anim. Behav.* **56**, 1035–1040.
- Foster, S. A. 1995: Understanding the evolution of behavior in threespine stickleback: The value of geographic variation. *Behaviour* **132**, 1107–1129.
- Foster, S. A. & Baker, J. A. 2004: Evolution in parallel: new insights from a classic system. *Trends Ecol. Evol.* **19**, 456–459.
- Foster, S. A., Scott, R. J. & Cresko, W. A. 1998: Parallel hierarchical variation and speciation. *Phil. Trans. R. Soc. Lond. B Biol. Sci.* **353**, 207–218.
- Frischknecht, M. 1993: The breeding colouration of male three-spined sticklebacks (*Gasterosteus aculeatus*) as an

- indicator of energy investment in vigour. *Evol. Ecol.* **7**, 439–450.
- Howell, D. C. 1997: *Statistical Methods for Psychology*, 4th edn. Duxbury Press, Belmont, California.
- Hughes, H. C. 1999: *Sensory Exotica: A World Beyond Human Experience*. MIT Press, Cambridge, Massachusetts.
- Huntingford, F. 2003: A history of *Animal Behaviour* by a partial, ignorant and prejudiced ethologist. *Anim. Behav.* **66**, 409–415.
- Kelber, A., Vorobyev, M. & Osorio, D. 2003: Animal colour vision – behavioural tests and physiological concepts. *Biol. Rev. Camb. Philos. Soc.* **78**, 81–118.
- Kilner, R. & Davies, N. B. 1998: Nestling mouth colour: ecological correlates of a begging signal. *Anim. Behav.* **56**, 705–712.
- Künzler, R. & Bakker, T. C. M. 2001: Female preferences for single and combined traits in computer animated stickleback males. *Behav. Ecol.* **12**, 681–685.
- McDonald, C. G., Reimchen, T. E. & Hawryshyn, C. W. 1995: Nuptial color loss and signal masking in *Gasterosteus aculeatus*: an analysis using video imaging. *Behaviour* **132**, 963–977.
- McKinnon, J. S. 1995: Video mate preferences of female three-spined sticklebacks from populations with divergent male coloration. *Anim. Behav.* **50**, 1645–1655.
- McKinnon, J. S. & Rundle, H. D. 2002: Speciation in Nature: the threespine stickleback model system. *Trends Ecol. Evol.* **17**, 480–488.
- McLennan, D. A. 1996: Integrating phylogenetic and experimental analyses – the evolution of male and female nuptial coloration in the stickleback fishes (*Gasterosteidae*). *Syst. Biol.* **45**, 261–277.
- McLennan, D. A. & McPhail, J. D. 1989: Experimental investigations of the evolutionary significance of sexually dimorphic nuptial coloration in *Gasterosteus aculeatus* (L.): temporal changes in the structure of the male mosaic signal. *Can. J. Zool.* **67**, 1767–1777.
- McRobert, S. P. & Bradner, J. 1998: The influence of body coloration on shoaling preferences in fish. *Anim. Behav.* **56**, 611–615.
- Milinski, M. & Bakker, T. C. M. 1990: Female stickleback use male coloration in mate choice and hence avoid parasitized males. *Nature* **344**, 330–333.
- Olson, V. A. & Owens, I. P. F. 1998: Costly sexual signals: are carotenoids rare, risky, or required? *Trends Ecol. Evol.* **13**, 510–514.
- Pelli, D. G. 1997: The video toolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.* **10**, 437–442.
- Pokorny, J. & Smith, V. C. 1977: Evaluation of single-pigment shift model of anomalous trichromacy. *J. Opt. Soc. Am.* **67**, 1196–1209.
- Reimchen, T. E. 1989: Loss of nuptial color in threespine sticklebacks (*Gasterosteus aculeatus*). *Evolution* **43**, 450–460.
- Rowe, M. P. & Jacobs, G. H. 2004: Cone pigment polymorphism in New World monkeys: are all pigments created equal? *Vis. Neurosci.* **21**, 217–222.
- Rowe, M. P., Baube, C. L., Loew, E. R. & Phillips, J. B. 2004: Optimal mechanisms for finding and selecting mates: how threespine stickleback (*Gasterosteus aculeatus*) should encode male throat colors. *J. Comp. Physiol. A* **190**, 241–256.
- Rowland, W. J. 1984: The relationships among nuptial coloration, aggression, and courtship of male three-spined sticklebacks, *Gasterosteus aculeatus*. *Can. J. Zool.* **62**, 999–1004.
- Rowland, W. J. 1989: The effects of body size, aggression and nuptial coloration on competition for territories in male threespine sticklebacks *Gasterosteus aculeatus*. *Anim. Behav.* **37**, 282–289.
- Rowland, W. J. 1994: Proximate determinants of stickleback behaviour: an evolutionary perspective. In: *The Evolutionary Biology of the Threespine Stickleback*, (Bell, M. A. & Foster, S. A., eds.). Oxford Univ. Press, New York, pp. 297–344.
- Rowland, W. J., Bolyard K. J., Jenkins J. J. & Fowler, J. 1995a: Video playback experiments on stickleback mate choice: female motivation and attentiveness to male colour cues. *Anim. Behav.* **49**, 1559–1567.
- Rowland, W. J., Bolyard, K. J. & Halpern, A. D. 1995b: The dual effect of stickleback nuptial coloration on rivals: manipulation of a graded signal using video playback. *Anim. Behav.* **50**, 267–272.
- Saino, N., Ambrosini, R., Martinelli, R., Ninni, P. & Møller, A. P. 2003: Gape coloration reliably reflects immunocompetence of barn swallow (*Hirundo rustica*) nestlings. *Behav. Ecol.* **14**, 16–22.
- Schluter, D. 2000: *The Ecology of Adaptive Radiation*. Oxford Univ. Press, New York.
- Scott, R. J. & Foster, S. A. 2000: Field data do not support a textbook example of convergent character displacement. *Proc. R. Soc. Lond. B Biol. Sci.* **267**, 607–612.
- Siegel, S. & Castellan, N. J. 1988: *Nonparametric Statistics for the Behavioural Sciences*, 2nd edn. McGraw-Hill, New York.
- Smith, V. C. & Pokorny, J. 1975: Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm. *Vision. Res.* **15**, 161–171.
- Smith, C., Barber, I., Wootton, R. J. & Chittka, L. 2004: A receiver bias in the origin of three-spined stickleback mate choice. *Proc. R. Soc. Lond. B Biol. Sci.* **271**, 949–955.
- Stavenga, D. G., Smits, R. P. & Hoenders, B. J. 1993: Simple exponential functions describing the absorbance

- bands of visual pigment spectra. *Vision. Res.* **33**, 1011–1017.
- Tinbergen, N. 1951: *The Study of Instinct*. Oxford Univ. Press, Oxford.
- von Uexküll, J. 1957: A stroll through the world of animals and men: a picture book of invisible worlds. In: *Instinctive Behavior: the Development of a Modern Concept* (Schiller, C. H., translator and ed.). International Univ. Press, New York. pp. 5–80.
- Vorobyev, M., Gumbert, A., Kunze, J., Giurfa, M. & Menzel, R. 1997: Flowers through insect eyes. *Isr. J. Plant Sci.* **45**, 93–101.
- Vorobyev, M., Marshall, J., Osorio, D., de Ibarra, N. H. & Menzel, R. 2001: Colourful objects through animal eyes. *Color Res. Appl.* **26**(Suppl.), S214–S217.
- Wedekind, C., Meyer, P., Frischknecht, M., Niggli, U. A. & Pfander, H. 1998: Different carotenoids and potential information content of red coloration of male three-spined stickleback. *J. Chem. Ecol.* **24**, 787–801.
- Wickens, C. D. & Hollands, J. G. 2000: *Engineering Psychology and Human Performance*, 3rd edn. Prentice Hall, Upper Saddle River, New Jersey.

## Appendix

### Computations for Colors of Synthetic Stimuli

Photoreceptor quantal catch rates ( $Q_i$ 's) were computed according to:

$$Q_i = \int F(\lambda) S_i(\lambda) d\lambda \quad (1)$$

where  $F(\lambda)$  is the quantal flux striking the photoreceptor at wavelength  $\lambda$ ,  $S_i(\lambda)$  is the photoreceptor's absorption spectrum, and the subscript,  $i$ , indicates which photoreceptor is under consideration (e.g.  $i = sS$  for the stickleback short wavelength sensitive receptor).

Photoreceptor-incident quantal fluxes were derived from:

$$F(\lambda) = I(\lambda) R(\lambda) T(\lambda) \quad (2)$$

where  $I$  is the irradiance striking the fish's throat,  $R$  is the fish throat's reflectance, and  $T$  is the percentage of light transmitted from that throat to the observer's photoreceptors.  $I$  and  $R$  were measured (Rowe et al. 2004). We assumed the major transmission losses were at the cornea and lens. We know of no data on these transmissions in stickleback eyes, but we have no reason to expect that they are wavelength dependent. Therefore,  $T(\lambda)$  was assumed to be a constant scale factor that could be ignored in the generation of stimuli designed to simulate the view through stickleback photoreceptors.  $T(\lambda)$  is

implicit in the values used for human cone spectral sensitivities (Smith & Pokorny 1975; Color and Vision Database, <http://www.cvrl.org>) which are essentially the product of  $S(\lambda)$  and  $T(\lambda)$  and can be utilized as such after substitution of Eq. 2 into Eq. 1.

For reflectances,  $R$ , 86 male fish were acquired, maintained, and measured for spectral reflectance as described previously (the CE-395 data set of Rowe et al. 2004). Irradiance,  $I$ , was the normalized average irradiance illuminating the tide pools where these fish breed (Rowe et al. 2004). After computing all relevant  $Q$ 's, monitor settings were computed for the generation of synthetic stimuli.

All computer monitors have a finite gamut. The electron guns of a computer-controlled CRT cannot produce less light than they produce when set at black level (the floor) or more light than they produce when set at maximum intensity (the ceiling). Absolute radiances from a monitor are generally much smaller than radiances of reflected sunlight (hence sunlight striking the face of a monitor renders its images invisible). Consequently, monitors cannot reproduce the absolute intensities of daylight scenes. We attempted to approach such intensities as closely as possible while ensuring all displayed stimuli were appropriately scaled to one another. Each simulated throat patch was designed to produce the same ratios of  $Q_S/Q_L$  and  $Q_M/Q_L$  as would be produced if the real throat patch were observed in the field. Because there are three electron guns ( $R$ ,  $G$ , and  $B$ ) and only two ratios specified here, these ratios do not completely constrain the monitor settings. We enforced the additional constraints that stimuli should be properly scaled relative to each other, and that the most saturated stimulus for each condition should be produced at the edge of the monitor's gamut. For each throat reflectance spectrum we determined monitor settings such that at least one of the electron guns was set at maximum while the other two were set to produce appropriate ratios of the  $Q$ 's. At these settings we could compute absolute  $Q$ 's produced by the monitor and relate them via a scale factor to the absolute  $Q$ 's that would be produced by light reflected from the corresponding throat patch if viewed in the field. The minimum (across fish) of these scale factors is the largest that would appropriately scale all of the stimuli relative to each other without any being compromised by the ceilings of the monitor's guns.

Using the above procedure, four of the 86 stimuli were compromised by the floor of one or more of the electron guns; the only way to produce the computed absolute  $Q$ 's would be if one of the guns

produced a negative amount of light. These four 'out of gamut' stimuli were not used in farther computations – a necessary compromise required by the use of a computer monitor. A monitor's dynamic range is not large enough to simultaneously satisfy the criteria that all stimuli be appropriately scaled to each other while producing the same relative Q's as would be produced by viewing the throats in the field.

### The Psychophysical Task

A large rectangle subtending visual angles of roughly 60° horizontal by 50° vertical was displayed at the background settings (derived from the average of the background radiance spectra measured in the field; Rowe et al. 2004). In the center of this rectangle we simultaneously displayed small rectangles each subtending approximately 4.5° vertically and 0.5° degrees horizontally. Each small rectangle simulated the color of one throat patch. The small rectangles were displayed side by side in a single horizontal row separated from each other by a narrow vertical band in which the monitor settings were the same as for the background (Fig. 3). Starting from an initially randomized order, the subjects rearranged the rectangles from most to least 'red' by using the computer's mouse to select an individual rectangle and then to indicate a new location in the lineup to which the selected rectangle should be moved. The rectangles were rearranged with this procedure until the subject was satisfied with the order.

### Comparing Psychophysics and Physiology

In our physiological model of the processing of color by stickleback, the important metric of throat color is the difference in the excitation of the long wavelength sensitive cone ( $E_{sL}$ ) and that of the short wavelength sensitive cone ( $E_{sS}$ ). These excitations are derived from the quantal flux computations according to:

$$E_i = \frac{Q_{is}}{(Q_{is} + Q_{ib})} \quad (3)$$

where  $Q_{is}$  is the quantal flux absorbed by photoreceptors of type  $i$  (e.g.  $sS$ ) stimulated by light from a stickleback throat, and  $Q_{ib}$  is the quantal flux absorbed by photoreceptors of type  $i$  when stimulated by background light (the same radiance used to generate the background in the psychophysical experiments).

To compare the psychophysical data to the physiological model, we first determined the identity of the fish given each rank (data were pooled across subjects for each of the two photopigment conditions). For instance, with Q's calculated for human pigments, the fish ranked number one (i.e. the least red) was fish ID no. 40 in four trials, fish ID no. 2 in four trials, and fish ID no. 39 in one trial. The differences between  $E_{sL}$  and  $E_{sS}$  when viewing the fish so identified were then used to compute a mean and standard error for each rank. Continuing with the example, the quantity,  $E_{sL} - E_{sS}$  is  $-0.0753$  for fish ID no. 40,  $-0.0224$  for fish ID no. 2, and  $-0.0541$  for fish ID no. 39. The average value for the difference across the nine trials is thus  $[4 \times (-0.0753) + 4 \times (-0.0224) + (-0.0541)]/9 = -0.0494$ . This value can then be compared with the value of  $E_{sL} - E_{sS}$  for the fish of the corresponding rank (one) by ordering the values of  $E_{sL} - E_{sS}$ . Fish ID no. 40 had the lowest value for this difference, and hence its value ( $-0.0753$ ) was physiologically ranked number one. The curves in Fig. 6a,b show the values of  $E_{sL} - E_{sS}$  in ascending order; the solid circles show the mean values of  $E_{sL} - E_{sS}$  of the fish given each rank in the psychophysical task. Error bars are  $\pm SE$ . The advantage of presenting the data this way rather than merely comparing ranks is that it shows how the conformity between the two types of rank are or are not affected by the absolute difference between  $E_{sL}$  and  $E_{sS}$ .