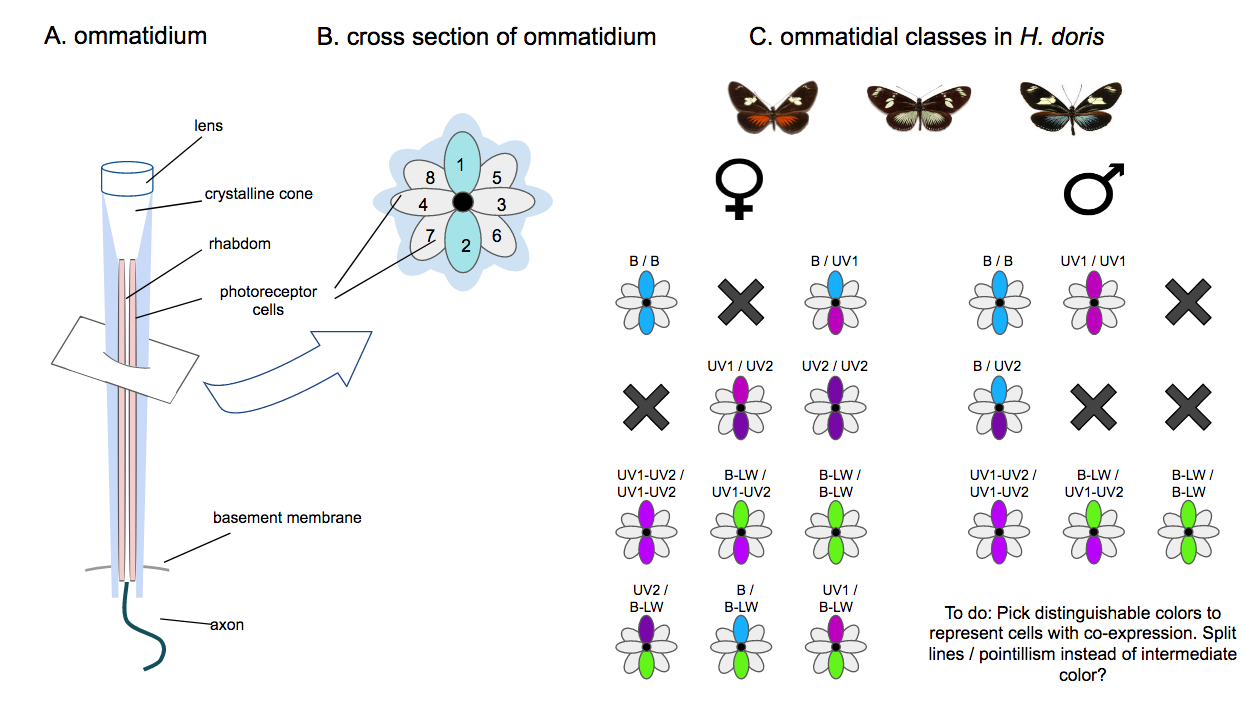


**Figure 1. Spectral sensitivity of *H. erato* photoreceptors and behavior apparatus**

a) Normalized spectral sensitivities of the different photoreceptors in the female *H. erato* eye (solid curved lines), used as a best guess for homologous photoreceptors in *H. doris*. *H. erato* females have photoreceptors maximally sensitive at 355 nm (UV), 390 nm (VS), 470 nm (B), and 555 nm (LW). Data from McCulloch 2016A via single-cell recordings. Dashed lines represent the wavelengths of our behavioral color stimuli, which we chose to be difficult to discriminate without both the UV and VS photoreceptors.

b) Schematic illustration of the two testing conditions: illuminated 380 and 390 nm color filters with variable irradiance and left/right positions.

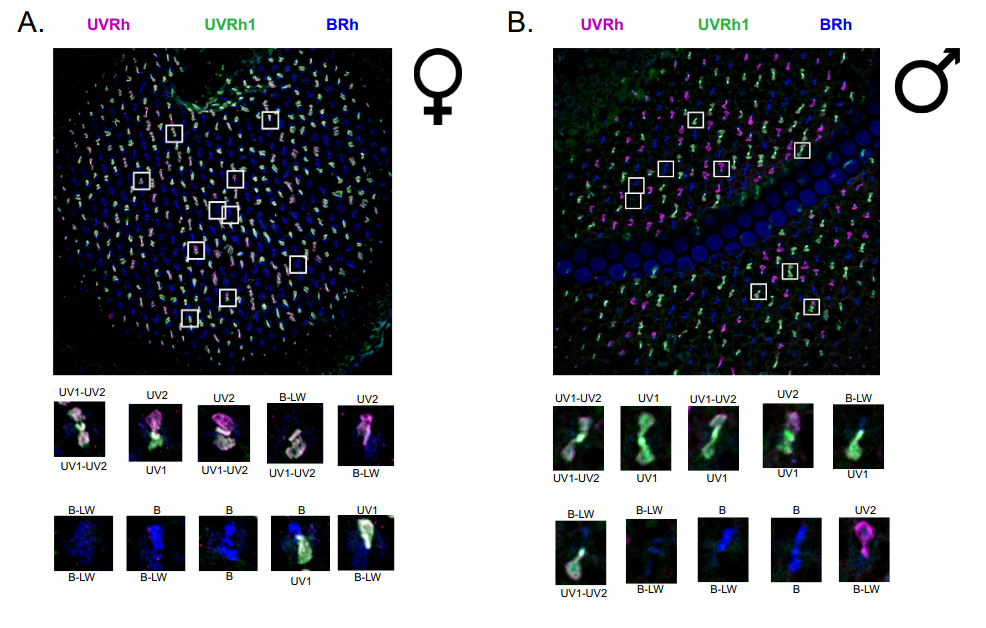
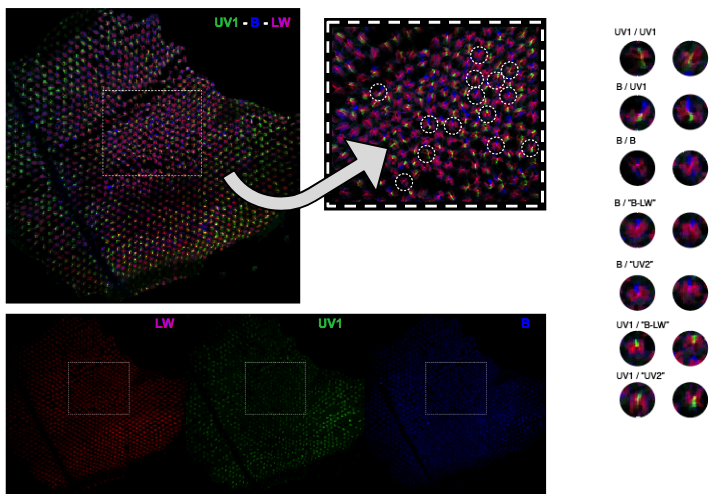
c) The behavior apparatus used for both training and testing, designed following Zaccardi 2006. A food trough at each color stimulus contained either a reward or control solution. We only counted landings inside the choice zone around each color stimulus (dashed box).



**Figure 2. *H. doris* eye anatomy**

**a)** Longitudinal and **b)** transverse views through a single ommatidium. Spectral classes of *Heliconius* photoreceptors 1-8: R3-R8 are always long-wavelength cells. R1 and R2 cells exclusively express short-wavelength opsins, which may be blue, violet, or UV. Across the compound eye ommatidia vary in which photoreceptors they express, so that in the same individual R1 and R2 may be occupied by different opsin combinations.

**c)** Ommatidial classes identified in *H. doris* females and males*.* Grey = LW cell (*LWRh*), blue = blue cell (*BRh*), lilac = violet cell (*UVRh2*), dark purple = UV cell (*UVRh1*).



**Figure 3. Opsin expression in *H. doris* photoreceptors**

Sections of *H. doris* **a) female** and **b) male**eyesimmunostained for opsin expression. Sections are stained with different triple labels, **UV1-UV2-B** (above) or **UV1-B-LW** (below). In both triple labels, blue = blue cell (BRh) and green = UV cell (UVRh1), while magenta = violet cell (UVRh2) if **UV1-UV2-B** or magenta = longwavelength cell (LWRh) if **UV1-B-LW**. Sections stained **UV1-UV2-B** do not show the 6 LW cells in R3-R8, while sections stained **UV1-B-LW** do not show UV2 cells in R1-R2. Individual cells with multiple stains suggest co-expression of multiple opsins.

Show side-by-side co-expression example.

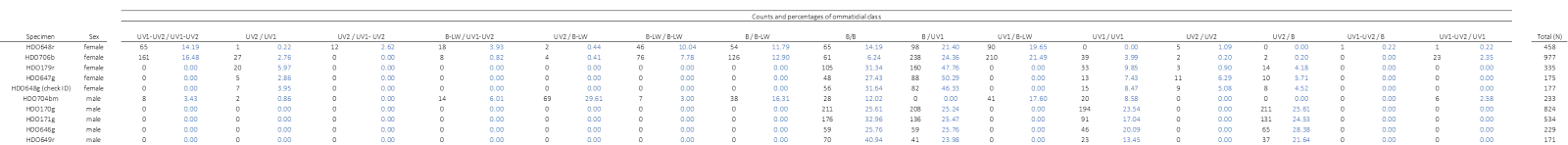
Triple labeling experiments

* Female: HDO710fr, (UV1-UV2-B)=(green-magenta-blue), (UV1-B-LW)=(green-blue-magenta)
* Male: HDO...m… (UV1-UV2-B)=(green-magenta-blue), (UV1-B-LW)=(green-blue-magenta)

To do: Give more visual weight to cells with co-expression. Box them?

Add panels that show the results of both sets of triple labeling experiments.

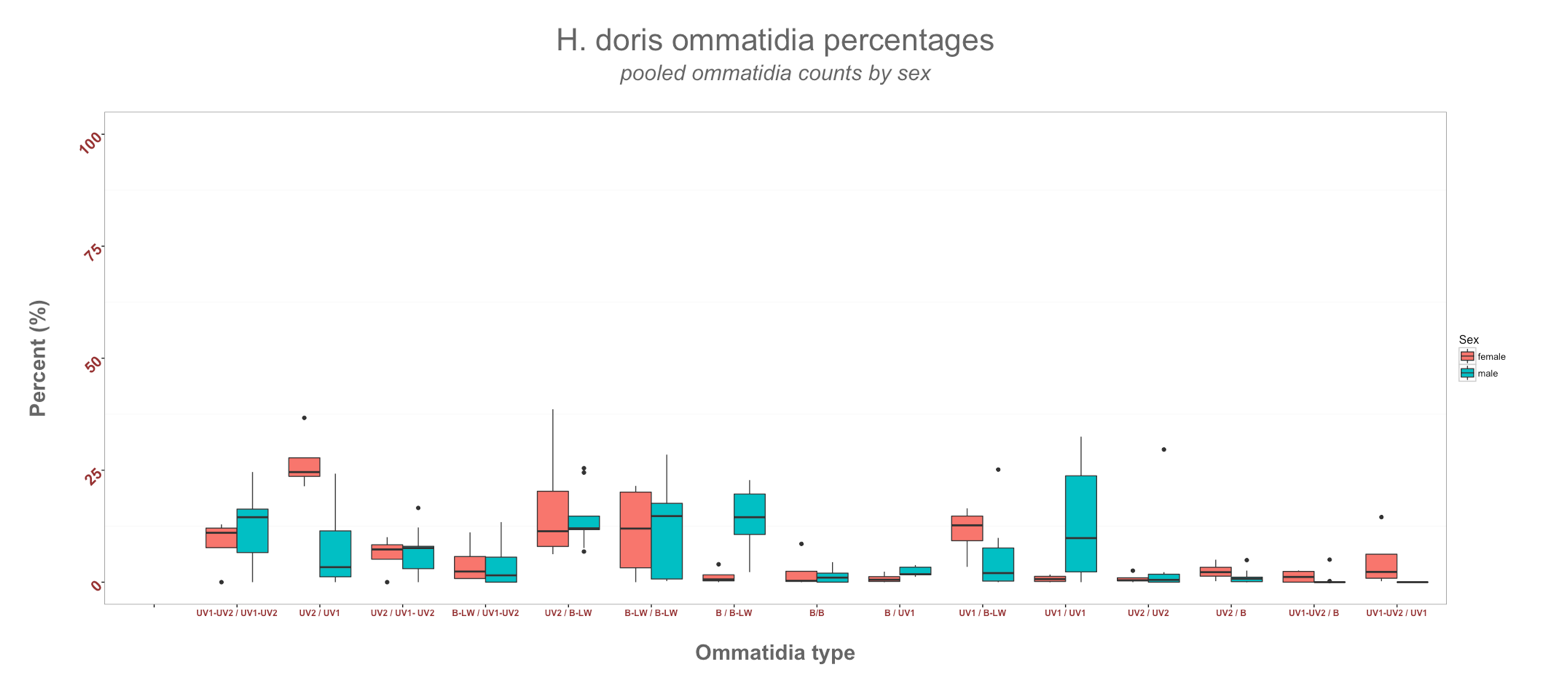
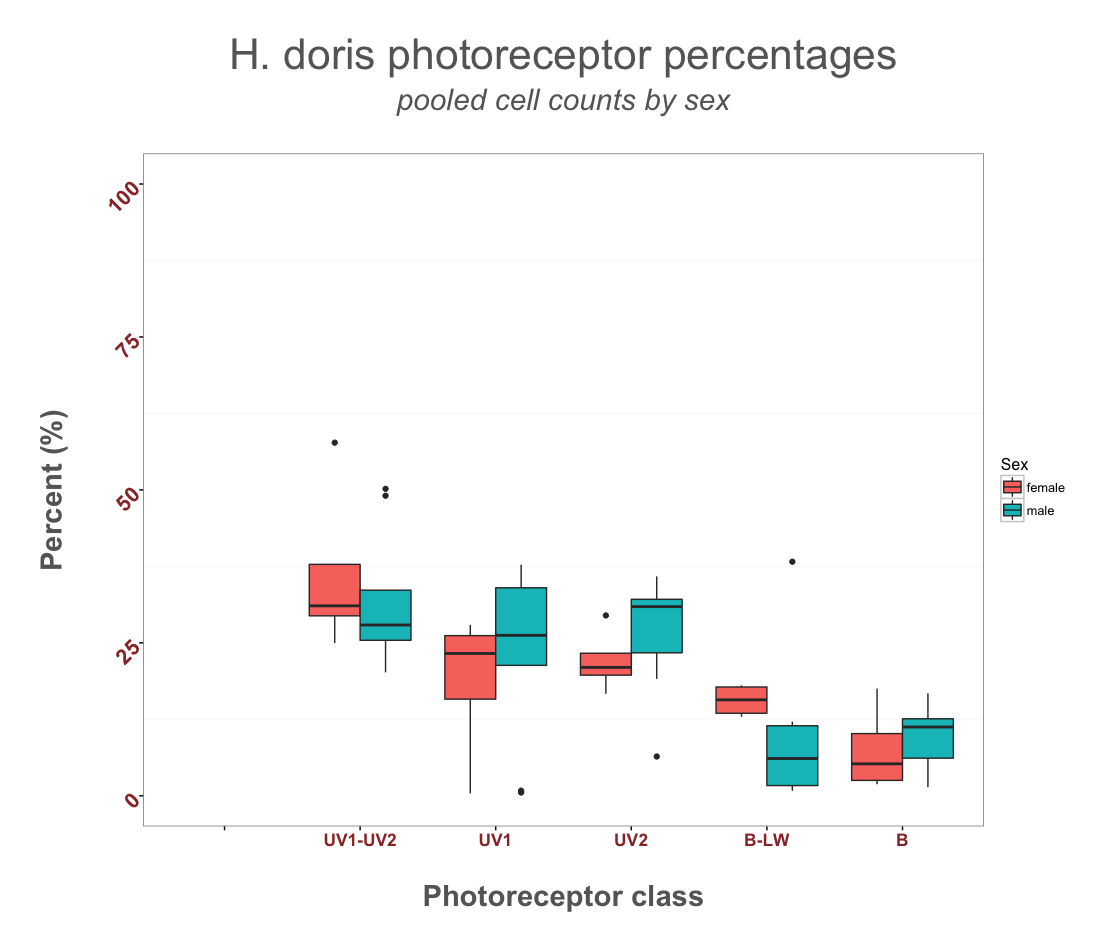
Use specimen 706bf (section RA) to show female B-LW coexpression.



**Table 1. Ommatidial counts in compound eyes of male and female *H. doris***

Counts and percentages of different ommatidial classes as categorized by their SW cells in individual *H. doris.* The last letter of the specimen ID indicates each butterfly’s color morph. Ommatidial classes were counted from stained retinas. In butterflies only the two shortwavelength cells express different opsins (blue, violet, or UV) across ommatidia, therefore these cells define the ommatidial classes.

Add images of ommatidia?



**Figure 4. Sexes compared by photoreceptor class and ommatidial type**

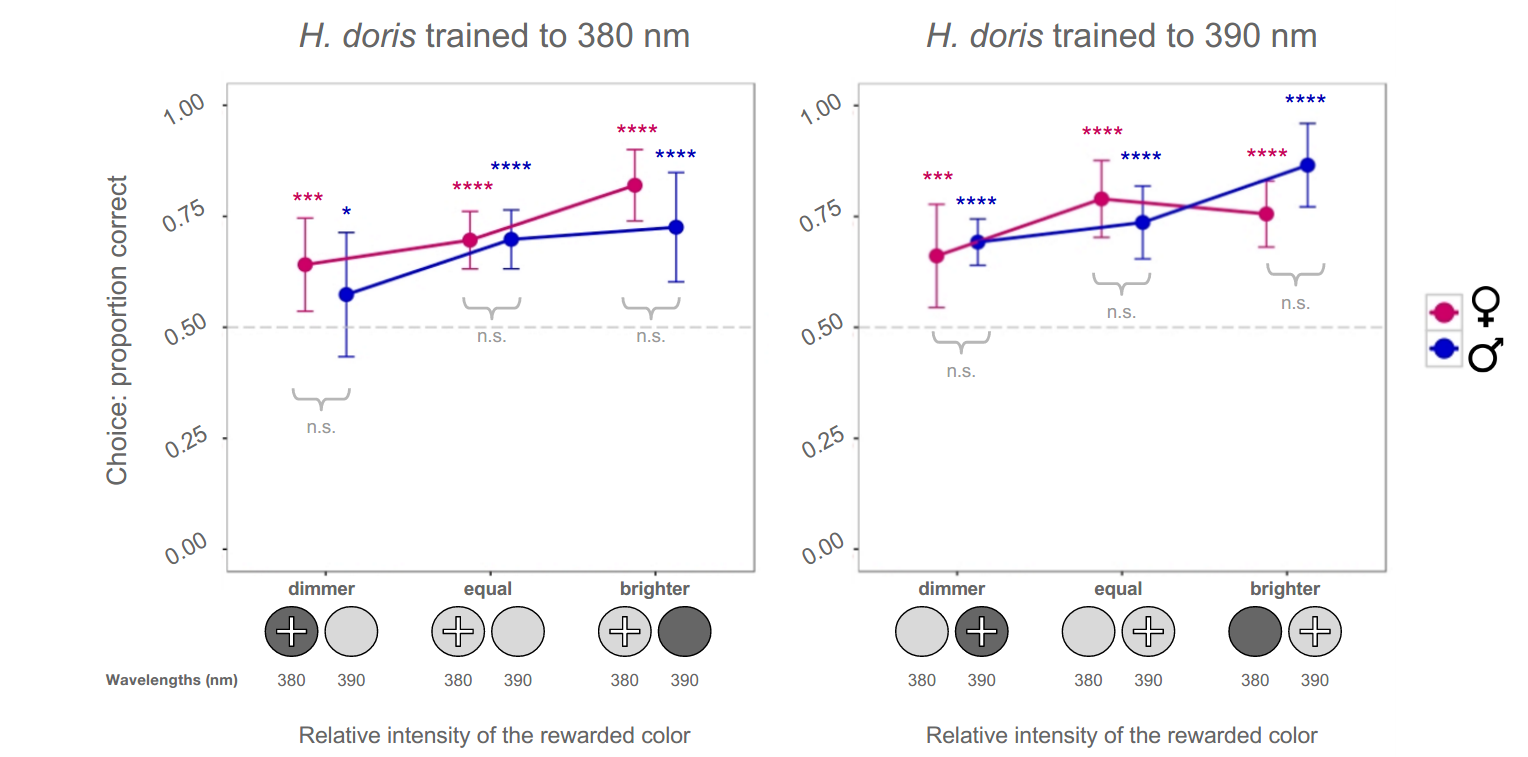
Boxplots of sex differences in photoreceptor class (A) and ommatidial type (B).

(A) The percentage of R1 and R2 photoreceptors that express a particular short-wavelength opsin in males (N=4) and females (N=3). Both sexes express opsins UVRh1, UVRh2, and BRh, but to differing degrees.

Females who exercise at least 3 hours weekly (N=6, red) vs. females who don’t (N=6, blue). We used an unpaired student’s t-test to compare mean bpm at rest and at peak exercise.

(B) The percentage of each ommatidial types in the compound eye in the same set of males and females. Ommatidial class is classified by R1 and R2 opsin combination. Male eyes lack the UV2/UV2 and UV1/UV2 ommatidia types present in females. We used a 2-sample t-test to compare sexes for the mean percentage of each photoreceptor and ommatidial class. For the t-test, the number of photoreceptors sampled was 3280 for males and 1374 for females, and the number of ommatidia sampled was 1758 for males and 687 for females. \*p<0.05, ∗∗p < 0.01; ∗∗∗p < 0.001; \*\*\*\*p<0.0001 between the two groups.

To do: Add images of ommatidia, photoreceptors.

****

**Figure 5. Behavioral color choice**

a) The proportion of correct choices made by individual *H. doris* trained and tested with 380 nm as the rewarded color and 390 nm as the unrewarded (N = 10 F, 12 M), as a function of the relative intensity of the rewarded color. Points are mean choice within a sex. The x-axis represents changes in illumination, where we varied the relative intensity of the colors to affect the obviousness of the correct choice. We used 3 intensity ratios of the rewarded color to the unrewarded color: Dimmer = 1:15, Equal = 1:1, or Brighter = 15:1. Butterflies chose the rewarded color across all 3 intensity trials, even Dimmer, despite the 15-fold intensity mismatch between training and test.

The dashed line represents the null of 50%, where binominal choice is no better than chance. Error bars represent the 95% binomial confidence interval. \*p<0.05, ∗∗p < 0.01; ∗∗∗p < 0.001; \*\*\*\*p<0.0001; ns, no significance between sexes.

b) *H. doris* behavior choice as above, but for butterflies trained and tested with 390 nm as the rewarded color and 380 nm as the unrewarded (N = 10 F, 11 M).

[**To do:**

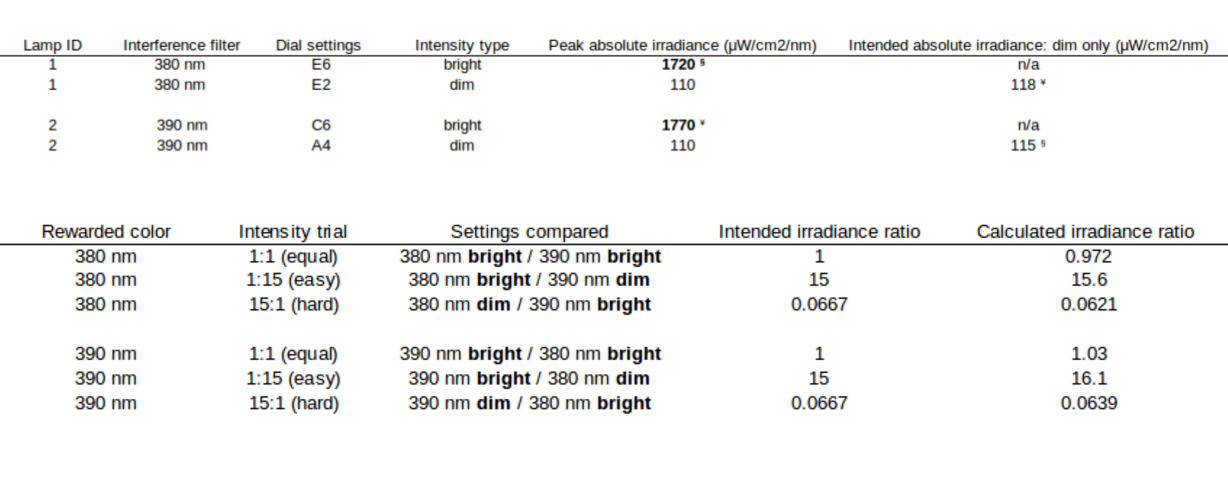
**Run a Friedman test.** This summarizes the data more succinctly than the Wilcoxon test. It detects differences in treatments across multiple test attempts and is applicable to complete block designs. The procedure involves ranking each row (or block) together, then considering the values of ranks by columns.

**Run a Kruskal–Wallis test by ranks.** This tests whether samples originate from the same distribution. It is used for comparing two or more independent samples of equal or different sample sizes. It *extends the Mann–Whitney U test* when there are more than two groups.]

[**Link to behavior video**](https://youtu.be/VciaXpP6Gno) (Video is hidden and unavailable without a direct link.)

**Supplementary video 1. *H. doris* behavior training**

A partially trained butterfly chooses the unrewarded color. She is re-released for a second attempt. She correctly chooses the rewarded color and is allowed to feed from the concealed trough. Butterflies were only permitted a single choice per exposure in the testing phase of the experiment.

****

**Supplementary table 1. Irradiance ratios and lamp settings used in the behavior experiment**

a) Emissions measurements of the two color stimuli at the settings used in the behavior experiment. We compared irradiance ratios of the stimuli at all 36 possible lamp settings on the Schott LCD before choosing the most appropriate settings. The bright settings made the irradiance of the color stimuli as similar as possible (bold). The dim settings approximated 1/15th the irradiance of the bright setting of the opposite color stimulus (see matching symbols).

b) The irradiance ratios produced by the lamp settings used in the experiment, indicating how well the chosen settings achieved the intended ratios. The irradiance ratios of the trained to the untrained color stimulus approximated 1:1, 1:15, and 15:1 as closely possible.

Light intensity vs. irradiance notes

We measured the transmitted light which passed through each filter. This was necessary because we cannot assume filters lit by the same amount of transmitted light have equal irradiance levels, because the amount of light passing through each filter also depends on properties such as filter thickness and coating. We picked our light settings based on perceived brightness, which depends on the level of **irradiance** (the amount of light actually striking the eye) rather than **light intensity** (the brightness of the bulb). Irradiance depends on the intensity and distance of the light source.

**Irradiance** encompasses units of area (square meter, m2) and time (min), thus giving the amount of light energy striking a 2-dimensional surface over a period of time. In contrast, **light intensity** commonly refers to the brightness of a light source.