Note: As of July 29, 2025, this document is updated to replace an earlier version. At the time of this update the first two participants (both control subjects) had completed 2-3 / 10 sessions. The primary changes are:

- Change the inclusion criteria for participants with migraine. This change was motivated by our difficulty finding participants who met our initial criteria.
- Change to a procedure in the psychophysical test that begins with a staircase before then using a Bayesian adaptive procedure for selecting stimuli.

****

UPenn Research Team (alphabetical order): Geoffrey K Aguirre, Ruby Bouhassira, Kristina Kitsul, Samantha Montoya, Alan Stocker

Research Project: Flicker photophobia as an experience of inefficient coding

Purpose and Approach

We will use a two-interval forced choice task to measure temporal frequency discrimination of flickering hemifields. The participant will view stimuli through eyepieces showing two side-by-side uniform hemifields, one for each eye. Participants will fuse the images using central fixation marks. On each trial, participants will view two sets of flickering stimuli and their task will be to report the interval with the mismatched pair. On each interval (3 second duration with a 0.75 second interstimulus interval), participants will view two flickering stimuli (one on each side). One of the intervals will show the same reference frequency in both eyes (out of phase). In the other interval, one side will show the reference frequency, and the other side will show the test frequency, controlled by either a staircase with a 3-down 1-up rule or QUEST+ (with a maximum of $\pm$ 6.75 dB from the reference). The mismatched interval will be randomized across trials, as will the side (left or right) containing the test stimulus. The participant will indicate which interval contained the different set of stimuli with the up/down arrows on a keyboard or the front mounted up/down buttons of a standard game controller. Data will be collected at two contrast levels crossed with two modulation directions (light flux and L–M). In half the trials, the test frequency will be lower than the reference, and the other half will be higher. We will test 5 log-spaced frequencies between 3 and 20 Hz. The goal is to investigate differences in frequency discrimination between L-M and LF and between controls and people with migraine and photophobia.

Subjects

We will collect a complete dataset from 15 healthy, headache free participants, and 15 participants with migraine and photophobia ages 18-40. Participants will meet these criteria:

<u>Controls:</u>

- Classified by the POEM (Kaiser et al., 2019; and as updated in the POEM analysis code as of the date of this pre-registration) to be “headache-free”.
- Score of 4 or below on CHYPS (1st quartile from Price et al., 2025).

<u>Migraine:</u>

- Classified by the POEM (Kaiser et al., 2019; and as updated in the POEM analysis code as of the date of this pre-registration) to be “migraine with aura”. (Any aura type accepted)
- Score of 17 or above on CHYPS (3rd quartile from Price et al., 2025).

Exclusion criteria for all subjects include:

• An answer of “yes” to the question “Do you have ongoing symptoms from a recent concussion?”

• An answer of “yes” to the question “Are you currently receiving treatment for a disease of the eye?”

• Abnormal color vision as judged by the Cambridge Color Test [>= 10 protan (red), >= 10 deutan (green), or >= 15 tritan (blue)].

• A best corrected visual acuity (BCVA) below 20/25 for either eye.

• A history of generalized epilepsy.

Participants will perform a practice task prior to starting data collection. Poor performance on this practice task may lead the participant to being excluded from the study (see below under *Subject*).

It is possible that we will identify other exclusion criteria as we screen subjects for the study. If participants are excluded for reasons identified prior to their participation in data collection, we will simply add these reasons to the list of exclusion criteria and not consider this a deviation from the protocol.

Participants will be permitted to listen to music or podcasts/audiobooks played from their smart phone speaker during the testing session or through headphones, but they need to be able to hear the feedback from the testing computer.

Participant Preparation

At the start of a data collection session, participants will be acclimated to the experimental room. The room illumination will be set so that the luminance of the walls of the room at eye level is roughly equivalent to the luminance of the stimulus field. The operator will review the experimental procedure with the participant and adjust the apparatus and chair for subject comfort.

Apparatus, stimulus, data collection

All of the stimuli will be generated with a digital light synthesis engine (CombiLED) under computer control. The stimuli will be presented through custom-made eyepieces with a semicircular, uniform hemi-field of 27.5° diameter. The participant will view the stimulus with their natural pupil. Visual stimuli are tailored to an individual observer’s age, taking their predicted lens density and predicted pupil size into account (which in turn is based upon age, luminance of the stimulus background, and stimulus area). The background level of the stimulus is adjusted by passing the light through a 0.5 log neutral density filter. The achieved luminance of the stimulus background will vary slightly by subject, but will be roughly ~3000-4000 cd/m^2. The chromatic content of the background may also vary slightly between observers as this shift is needed to achieve

the desired contrast levels. Ultimately, the nominal, targeted contrast upon the post-receptoral mechanisms will be set to 2.5% low and 7.5% high on L–M, and 10% low and 30% high on LMS.

Participants will complete 10 sessions across 4-5 visits; more than one session may take place on a single day as long as there is a break of more than 15 minutes between the first and second sessions and an hour break before a third session. No more than three sessions will be completed in a single day. The first three sessions will use a staircase procedure (starts at 1dB, 3-down 1-up rule), and the next seven will use QUEST+ (with a maximum of $\pm$ 6.75 dB from the reference). This is because during data collection with control participants, QUEST+ was selecting test frequencies that were either higher than critical fusion frequency or too slow to perceive. This resulted in participant responses plateauing at the maximum dB value. Beginning with a staircase procedure will ensure that the distance of the test frequency from the reference will gradually increase until it reaches threshold.

Each session is composed of 10 blocks with 20 trials. Each block measures one post-receptoral direction (L–M or light flux), with the direction alternating across blocks. Each block presents 20, two-alternative forced choice trials, each consisting of 2 intervals. In one of the intervals, both sides show the reference frequency (out of phase) and the other shows the reference frequency and a test frequency. The participant is asked to choose the interval with the mismatched pair of flickering lights. The reference flicker frequency within a trial is drawn from a uniform (log) distribution of values between 3 and 20 Hz. The initial frequency of the stimulus presented after the delay is drawn from a uniform distribution of ±6.75 dB around the reference frequency. Participants receive feedback after each trial, with either a positive “beep” sound if they chose the correct interval, or a negative beep is they chose the Incorrect interval. The participant is paid a bonus of 5 cents for each correct response they make during a session.

Data analysis and primary hypothesis testing

Data will be analyzed within-subject, grouped across sessions and blocks into each of the four stimulus conditions (high and low contrast level, crossed with L–M and light flux post-receptoral directions). We will test for differences in the discrimination thresholds (in dB) vs. the reference flicker frequency (on a log axis). We will test if thresholds differ by more than the 95% confidence interval between high and low contrast levels, between the L–M and light flux directions, and between participant groups.

Data Exclusion Criteria and handling errors in data collection

Device

If the testing apparatus fails temporarily (due to e.g., software crash) and it is possible to resume testing, then we will discard the single block during which the failure occurred, repeat this block, and then continue to complete the session. If the failure occurs between sessions, then the study can continue as originally planned after recovering from the failure. There may be failures of the testing apparatus that cannot be resolved within a few minutes. If so, we will discard all the data collected during this session, and we will invite the subject to return to re-collect data for the session.

Subject

If a subject leaves the study prior to the completion of 10 sessions of data collection, and they have completed fewer than 8 sessions, we will discontinue their participation in the study and exclude all of their collected data. If they have completed 8 sessions, we will keep their data.

If a participant needs to briefly stop data collection during a session, we will treat that circumstance in the same manner as a brief equipment failure that interrupts collection of data during a block. If a participant needs to stop data collection during a session and is unable to continue we will discard all data from that session and attempt to re-collect the session on a subsequent day. If the subject discontinues a second session we will discontinue their participation in the study.

Participants will complete practice trials that are between 5 and 6 dB at 3 and 20 Hz. If they get more than 3 wrong (4 or more) out of 10 on either LF or L-M, we will ask them to repeat the practice task. If they still get more than 3 wrong, they are excluded.

The operator may observe subject behavior during the session that prompts concern that the collected data will be of low quality (e.g., excessive movement and fidgeting or sleepiness). The operator will discuss these observations with the study team, and data for a subject may be excluded from the study on this basis, as long as the decision is made prior to examination of the psychophysical data. While we will collect eye movement data using electro-oculography, we will not use these measurements as a basis of excluding participants.

If participants are discontinued, we will recruit additional subjects in an effort to meet our prespecified study numbers.