tvst

Retina

Assessment of Murine Retinal Acuity Ex Vivo Using Multielectrode Array Recordings

Darwin Babino[1,*], Tyler S. Benster[1,5,*], Laura Laprell[1], and Russell N. Van Gelder[1–4]

[1] Departments of Ophthalmology, University of Washington School of Medicine, Seattle, WA, USA
[2] Biological Structure, University of Washington School of Medicine, Seattle, WA, USA
[3] Laboratory Medicine and Pathology, University of Washington School of Medicine, Seattle, WA, USA
[4] Roger and Angie Karalis Johnson Retina Center, University of Washington School of Medicine, Seattle, WA, USA
[5] Neurosciences Program, Stanford University School of Medicine, Stanford, CA, USA

Correspondence: Russell N. Van Gelder, Boyd K. Bucey Memorial Chair, Professor and Chair, Department of Ophthalmology, University of Washington School of Medicine 325 9th Avenue, Campus Box 359608, Seattle, WA 98104, USA. e-mail: russvg@uw.edu

Received: June 21, 2022
Accepted: September 22, 2022
Published: January 4, 2023

Keywords: retina; multielectrode array; acuity

Citation: Babino D, Benster TS, Laprell L, Van Gelder RN. Assessment of murine retinal acuity ex vivo using multielectrode array recordings. Transl Vis Sci Technol. 2023;12(1):4, https://doi.org/10.1167/tvst.12.1.4

Purpose: Visual acuity, measured by resolution of optotypes on a standard eye chart, is a critical clinical test for function of the visual system in humans. Behavioral tests in animals can be used to estimate visual acuity. However, such tests may be limited in the study of mutants or after synthetic vision restoration techniques. Because the total response of the retina to a visual scene is encoded in spiking patterns of retinal ganglion cells, it should be possible to estimate visual acuity in vitro from the retina by analyzing retinal ganglion cell output in response to test stimuli.

Methods: We created a method, EyeCandy, that combines a visual stimulus-generating engine with analysis of multielectrode array retinal recordings via a machine learning approach to measure murine retinal acuity in vitro. Visual stimuli included static checkerboards, drifting gratings, and letter optotypes.

Results: In retinas from wild-type C57Bl/6 mice, retinal acuity measurement for a drifting grating was 0.4 cycles per degree. In contrast, retinas from adult rd1 mice with outer retinal degeneration showed no detectable acuity. A comparison of acuities among different regions of the retina revealed substantial variation, with the inferior–nasal quadrant having highest RA. Letter classification accuracy of a projected Early Treatment Diabetic Retinopathy eye chart reached 99% accuracy for logMAR 3.0 letters. EyeCandy measured a restored RA of 0.05 and 0.08 cycles per degree for static and dynamic stimuli respectively from the retina of the rd1 mouse treated with the azobenzene photoswitch BENAQ.

Conclusions: Machine learning may be used to estimate retinal acuity.

Translational Relevance: The use of ex vivo retinal acuity measurement may allow determination of effects of mutations, drugs, injury, or other manipulations on retinal visual function.

Introduction

In humans, visual acuity is tested clinically by asking the subject to read high-contrast, standardized optotype letters subtending smaller angles of resolution until the individual can no longer correctly identify individual letters. Visual acuity testing is the cornerstone of the clinical evaluation of visual function and has been used as a primary end point in many pharmacological and surgical clinical trials. Visual acuity testing in animals is inferential and is performed with behavioral tests such as optokinetic reflex testing or learned associative tests such as forced choice maze navigation.[1,2] Optokinetic testing is limited to a specific stimulus, simulating a rotating drum. Acuity to static stimuli cannot be estimated by this technique. The forced choice testing requires repeated training of animals in an associative learning task. Animals vary in their ability to learn this association; in many paradigms, untrainable animals are omitted from trials.[3] All visual acuity testing methods in mice require

intact motor function and normal central nervous system function, which can create significant limitations. For example, mutations resulting in neurologic or muscular deficits along with retinal dysfunction cannot be studied. Studying the effects of agents that influence vision and alertness simultaneously may not be possible. Similarly, studying short-term vision restoration (e.g., with small molecule photoswitches) in which the duration of action of the agent is shorter than the period required for associative learning, may be impossible.

Visual processing is initiated by a highly organized neural network within the retina. Incident light is focused by cornea and lens, and captured by a high density array of photoreceptive cells (rods and cones), which feed information forward through neural retinal processing layers (horizontal, bipolar, and amacrine) via synaptic and gap junction connections, to ultimately drive action potentials in a small and diverse set of retinal ganglion cells (RGCs).[4] All visual information in a scene is thus encoded in the time-series of spikes from this set of neurons. Early investigations of the neural code employed by ganglion cells in transmitting visual information found that statistical analysis of RGC spike trains could predict the presence or absence of light presentation.[5,6] Warland et al.[7] showed that both an artificial neural network trained by back-propagation to match spike trains with stimuli and an optimized linear filter on action potentials were able to decipher visual information from a population of salamander RGCs. More recently, Bayesian decoding has been used to determine visual performance of in vitro retinas tested via multielectrode arrays.[8–10] Such estimates may be considered measurement of "retinal acuity"; the term visual acuity is reserved for performance of the fully intact visual system (including central nervous system function). Previous approaches either established a theoretical upper bound on retinal acuity by extrapolating decoding performance from a few cells to an estimated RGC density or measured a coarse lower bound for retinal acuity significantly below animal VA performance. The advent and rapid advances in machine learning (ML) tools has led to their application in vision research,[11] and raises the possibility that ML could be used to serve as surrogate for central nervous system function in determination of retinal acuity. To date, however, a comprehensive platform that arranges customizable stimuli and data processing for analysis via ML for the purpose of in vitro acuity measurement has not been developed.

Here we describe the EyeCandy platform for the automated assessment of retinal acuity. The platform consists of two major components: a web-based projected image stimulation engine, and a paired analysis package. The stimulation engine is run through a digital projector and optical system to present a defined series of high-resolution visual stimuli to the retina in vitro. Multielectrode array (MEA) recording data of RGC firing are collected and analyzed via a ML approach, in which the computer learns the RGC firing patterns associated with specific presented patterns, and then is tested on novel presentations of similar stimuli. The error rate of the ML is a measure of the information content present in the RGC train and, thus, puts a lower limit to retinal acuity. We demonstrate the usefulness of this system in assessing visual potential of the mouse retina in several paradigms that are challenging for in vivo visual acuity assessment.

Methods

Animals

Adult C57BL/6J (*wt*) and C3H/HeJ (Pde6b^{rd1}/Pde6b^{rd1}, referred to herein as *rd1*) mice between 4 and 12 months old (Jackson laboratory, Bar Harbor, ME) were used in this study. Both male and female mice were incorporated into studies randomly. All animals were treated in accordance with the ARVO Statement for the Use of Animals in Ophthalmic and Vision Research and all animal experiments were conducted following approved protocols by the institutional animal care and use committee at the University of Washington. All parameters were tested across at least four retinas. Figure panels represent experiments on individual retinas.

MEA Recordings

Euthanasia of all mice was performed using CO_2. Eyes were enucleated quickly and placed into room temperature artificial cerebrospinal fluid (ACSF) under a dim red light. ACSF solution contained 125 mM NaCl, 2.5 mM KCl, 1.25 mM NaH2PO4, 1 mM MgCl2, 2 mM CaCl2, 26 mM $NaHCO_3$, and 20 mM D-glucose and aerated with 95% O_2/5% CO_2. Isolated retinas, one-half sections separated between superior and inferior, were placed whole-mount with retinal quadrant-specific sections, RGC layer facing down, onto a 60-channel MEA spanning approximately 2.56 mm^2 (60 MEA 200/30iR-Ti, Multi Channel Systems, Reutlingen, Germany) and recorded on a MEA 1060-inv-BC system. Retinas were kept at 34°C and continuously perfused with ACSF at 3 mL/min and allowed to settle and recover for 1 hour before recordings. Processing of extracellular spikes was performed using Offline Sorter V3 (Plexon, Dallas,

TX); Butterworth-filtered raw signals were high-pass filtered at 330 Hz and digitized at 20 kHz with a spike threshold setting of 5 SD for each channel. Spike detection occurred using raw signal amplitude thresholds using standard deviations from mean of peak heights histogram with sigmas at –3.80; unsorted waveforms were excluded from further analyses. Spike sorting was achieved using the Automatic Sorting feature with the T-Dist E-M sorting method with a D.O.F multiplier of 15, sorted in two-dimensional space.

Drug Treatments

DNQX (50 µM) and AP5 (25 µM) (Tocris Biosciences, Bristol, UK) in ACSF were administered under continuous perfusion 1 hour before and during visual stimulation. Washout procedures consisted of 1 hour perfusion with ACSF. BENAQ (100 µM) with a final concentration of 10% cyclodextrin (Reagent World, Irvine, CA) was administered for 1 hour under perfusion before light stimulation testing.

Stimulator

A Digital Light Projector (DLP) LightCrafter 4500 (Texas Instruments, Dallas, TX) coupled to a custom optical, two-lens system capable of 6 cycles per degree (cpd) resolution focused light stimulation onto the retina from below (i.e., presenting to the ganglion cell side of the retina, consistent with normal vision). The system provided 1280 × 800 pixels of spatiotemporally patterned stimulation over the area of the MEA with a refresh rate of 60 Hz and control of brightness and simultaneous RGB LED operation. The DLP projector was outfitted with three independently controlled RGB LEDs with a recorded maximum emission at 617, 509 and 455 nm for red, green, and blue, respectively, with an average half-width of ±30 nm. For all *wt* and BENAQ-treated experiments, regardless of isolated or compound use of LEDs, a total photon flux of approximately 5×10^{10}, approximately 3.5×10^{13} and 5.5×10^{16} photons/cm^2*s at 500 nm, respectively, was used as measured for each maximal emission regardless of stimulus color. For white light stimulations, equivalent photon flux per LED was used at photopic light levels. Total quantal catch for the MW- and UV-opsin, respectively was calculated as 1.7 and 5.5×10^{-5} R*/cone/s.[12]

Stimuli

Unless noted otherwise each experimental stimulus lasted 1 second and was preceded by 1 second of darkness. All were followed by a uniform draw of 1.0 to 1.5 second seconds of darkness to limit adaptation. For monitoring the integrity of the retina, each recording began with 1 second of darkness, a 1-second full-field flash, and 1.0 to 1.5 seconds of darkness. These stimuli were repeated every 5 minutes.

In a single recording, we used either (1) a checkerboard protocol, where the positive stimuli displayed a checkerboard pattern where the color of each square swapped at 0.5 seconds for contrast trials and adjusted accordingly for testing stimulus duration on accuracy predictions. For these experiments, stimuli lasted for the tested duration, from 0.1 to 1.0 second, for original pattern and additionally for the negative, nonpattern inversion, or positive, pattern inverted checkerboard. The negative stimuli maintained a static pattern for the entire 1 second, or (2) a drifting square grating protocol, where the positive stimuli moved up and right for 1 second and the negative stimuli moved down and left for 1 second. Both checkerboards and gratings were oriented at 45° to match the diamond pixel arrangement of the digital micromirror device.

In a standard recording, 25 examples of each class for each of 6 resolutions for each of 7 stimulus conditions (i.e., contrast) were displayed for a total of 2100 stimuli presented over approximately 2 hours. Regular periods of rest for up to 30 minutes were included in the protocols to allow chromophore regeneration. All wt retinal recordings were performed with the retinal pigment epithelium intact. The stimuli were presented in random order for each run.

For the Early Treatment Diabetic Retinopathy (ETDRS) letter experiments, a similar protocol was used wherein each stimulus came from 1 of 11 classes (blank, C, D, H, K, N, O, R, S, V, or Z) using the Sloan font. Each letter and blank was presented to the 11 retina for 1 second per trial.

Cell Clustering

Every 5 minutes, we exposed the retina to 1.0 second of darkness, 0.5 seconds of full-field white light, and 2.0 seconds of darkness. We constructed the mean firing rate per unit by binning the spikes across all trials into 1-ms bins, dividing by the trial count, and convolving with a Gaussian kernel (50-ms bandwidth, cutoff at 6-sigma). We take the first two principal components, and then used ordering points to identify the clustering structure[13] according to a cosine distance metric, a minimum steepness (xi) of 0.05, and a minimum of 8 units per cluster. When applied to our highest-acuity recording—a grating contrast experiment—this analysis resulted in 157 of 270 units split into 13 clusters. Although unbiased, the extracted clusters are sensitive to changes in preprocessing. To improve robustness,

we constructed a reachability plot: a graphical representation of ordering points to identify the clustering structure that is analogous to the dendrogram in hierarchical clustering. We use this plot to select a noise threshold, epsilon, that separates "true" clusters from noise by identifying regions of low density, and extracting clusters according to density-based spatial clustering of applications with noise.[14] We empirically found that an epsilon between 0.004 and 0.016 provides consistent results across multiple retinas. We reliably extracted clusters that correspond with off cells (cluster 0), on cells (cluster 3), and on–off cells that are either balanced (cluster 2), have an enhanced off response (cluster 1), or a stronger on response (cluster 4).

Feature Extraction and Cross-Validation

We initially constructed a featurized representation of the RGC responses to each 1-second checkerboard or grating stimuli by counting the number of spikes per unit per 100-ms time bins and flattened these counts into a one-dimension feature vector. The target for each example was either 0 or 1, corresponding with either two identical checkerboards or two alternating checkerboards or drifting grating to the left or drifting grating to the right. For optotype disinclination outputs were 0 to 9, corresponding with each of the ten ETDRS letters. We split the 6 sizes × 7 contrasts × 25 repetitions × 2 classes into 42 groups of 50 samples. Using Monte Carlo cross-validation,[15] we randomly partitioned the 50 samples into 30 training examples that were used to train a radial basis function (RBF) kernel support vector machine (SVM)[16] and 20 withheld examples used to test the accuracy of the SVM. For classification of ETDRS letters, we used a one-against-all strategy for generalization to 11 classes.[17] This partition process was repeated 30 times to construct a sample mean from the accuracy of the SVM on withheld data.

RBF Kernel SVM Classifier

For each Monte Carlo cross-validation draw, we performed a grid search on two RBF kernel SVM hyperparameters, C, a regularization parameter against the decision function's margin, and gamma, a parameter adjusting the radius of influence of each example 15 SVM models were trained for each draw of training examples, and the best-performing model, as measured by accuracy on training data, was chosen for evaluation on the withheld test data. Thus, for each group we trained number of draws × models per grid search = 450 SVM models. For each full recording, we trained 18,900 unique SVM models.

The Cycles Per Degree (cpd) of the Projected Stimuli

Our setup projects 13.09 pixels per degree in x and 12.54 pixels per degree in y. This was calculated based on a posterior nodal distance of 2 mm.[18] We used the smaller of the two numbers for conservatively converting stimulus size in pixels into cycles per degree or logarithm of the minimum angle of resolution (logMAR).

Acuity Threshold

A neural code is capable of transmitting information at a particular spatial acuity if a receiver is capable of recovering the signal at a low error rate. Therefore, we defined the acuity threshold as the highest value of cycles per degree for which the mean accuracy level outperformed random chance at a 1% or lower significance level. Contrast sensitivity was similarly defined as the largest value of 1/contrast for a given spatial acuity that outperformed random chance at a 1% or lower significance level.

Data Availability

The data used to support this article are available upon request from the corresponding author.

Results

The experimental construct consists of three parts: the stimulator, which projects psychophysical stimuli on the isolated retina in vitro; the MEA system recording system, which samples and collects RGC spike trains in response to each stimulus; and the data processor, which translates the spike trains into data streams to be processed using ML. Both the stimulator and MEA system are commercially available (DLP LightCrafter 4500 and Multi Channel Systems MEA, respectively) (Fig. 1). The stimulator is powered by a webpage served by EyeCandy, which leverages HTML5 canvas to provide cross-platform GPU acceleration for high-performance stimulation. Protocols were written in the EyeCandy Programming Language, enabling the creation of multihour experiments with thousands of stimuli presented in a time-stratified random order to avoid confounds from adaptation or preparation decay. For each experiment, a human- and machine-readable laboratory notebook was automatically generated that included stimulus metadata such as spatial resolution, stimulus class, and condition (e.g., contrast or drift speed). This laboratory notebook and spike-sorted data were passed to Glia, an analysis package

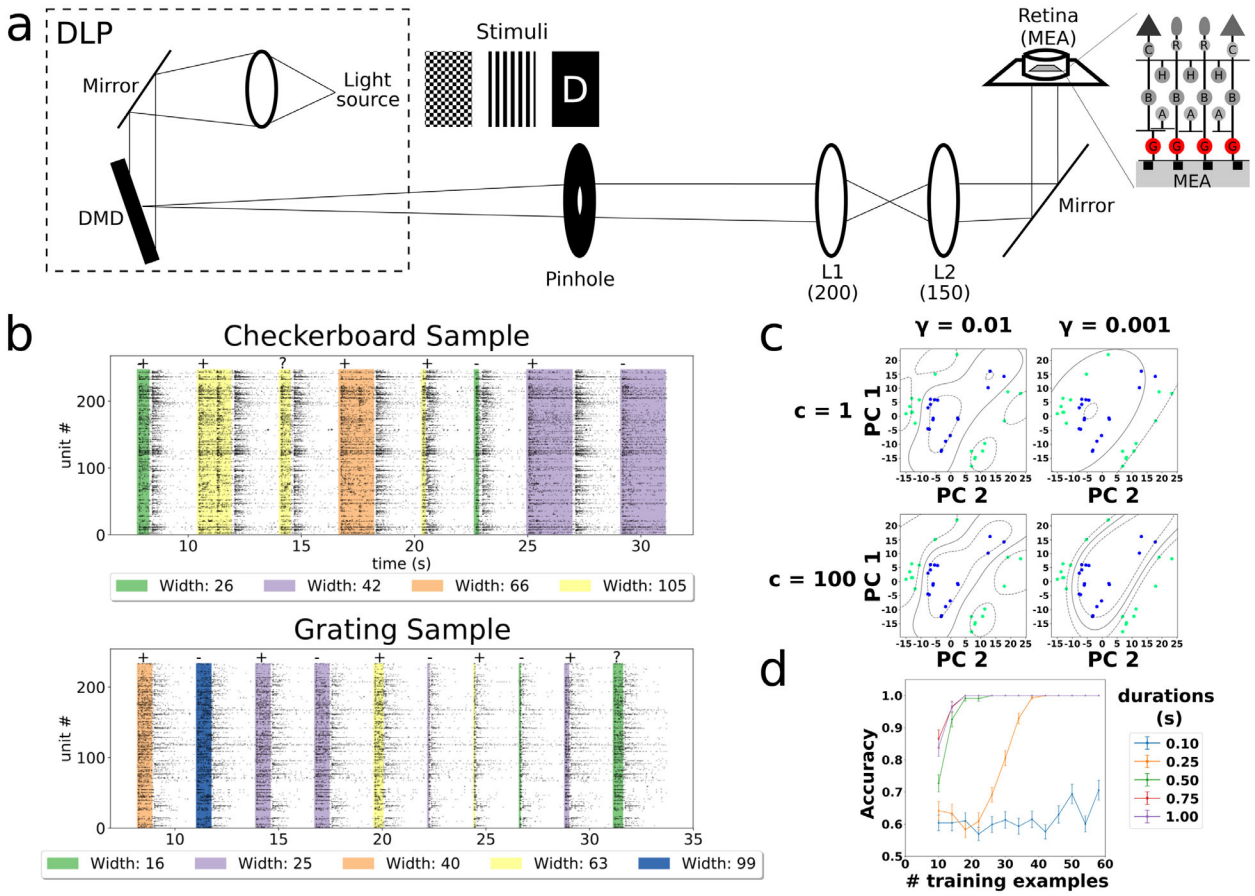


Figure 1. Components of the EyeCandy system. (a) Schematic diagram of tandem DLP-based stimulator and multielectrode array with sample stimuli. Images of the digital micromirror device (DMD) are formed between lenses L1 (f = 200 mm) and L2 (f = 150 mm) and on the MEA-mounted retina. Isolated retinas were placed RGC layer facing down, onto the MEA. C, cones; R, rods; H, horizontal cells; B, bipolar cells; A, amacrine cells; G, ganglion cells. (b) Spike raster plot of multiple trials with varying grating widths (color) and durations (area). Sixty percent of trials for each condition are randomly assigned to training and labeled with direction + or –, and the remaining 40% of trials are assigned to test, denoted by "?". Unit number indicates the number of individual RGCs recorded ($n = 8$ retinas). Classifier performance is estimated by multiple draws (checkerboard and grating stimulus samples) using Monte Carlo cross-validation (MCCV), and uses the classifier accuracy from each draw to calculate the standard error of the mean. (c) Illustration of how the penalty parameter, C, and kernel coefficient, gamma (γ), of the RBF influences a SVM classifier on actual data. To visualize, we projected each data point onto the first two principal components before training an SVM. The classifier performs a grid search over C and gamma to find the best performing SVM on training data. The solid line represents the decision boundary and the dotted line shows the margin. (d) Classification accuracy as a function of number of available training examples across five different stimulus time durations (see Methods).

written in Python, for training a SVM classifier on a time-stratified subselection of data of varying class, but consistent spatial resolution and condition. A separate classifier was trained for each combination of spatial resolution and condition. We defined retinal acuity as the highest- acuity set of stimuli for which we obtain a P value of less than .01 statistically significant classification accuracy on held-out test data.

Estimation of Ex Vivo Retinal Acuity to Static and Dynamic Stimuli

A variety of visual behavior tests have been developed for rodents with either static or dynamic stimuli.[1–3] We initially tested two widely used paradigms using black-and-white stimuli for a binary classification task: an alternating checkerboard (a static measure since there is no directional motion of the stimulus), and a kinetic drifting grating (Fig. 1b). For checkerboard acuity, we used RBF kernel SVMs to distinguish whether the retina was presented with either (a) two sequential identical checkerboards or (b) two sequential alternating checkerboards. For drifting gratings, we trained the classifier to distinguish between two opposing directions of drift. For both paradigms, three degrees of freedom were modulated: spatial resolution, contrast, and stimulus duration. For drifting gratings, a fourth parameter—drift

speed—was also varied. Unless noted, a typical test for each paradigm per retina consisted of 6 sizes × 7 contrasts × 25 repetitions × 2 classes (+, –), resulting in 42 groups of 50 samples. A separate classifier was created for each group.

We evaluated the performance of several standard classification approaches: logistic regression, multilayer perceptron, convolutional neural networks, nearest neighbor, random forests, and SVM. Multilayer perceptron, random forests, and SVM consistently had the highest performance. We selected SVM as our model of choice; during Monte Carlo cross-validation, we found that SVM had lower variance in classification accuracy than random forests or multilayer perceptron models. Although deep learning approaches achieve state-of-art performance in numerous benchmark tasks, SVM is better suited for problems with low sample counts and where fast and robust training is required.[19]

Of the 50 samples, 30 training examples (15 of each class) were used to train a particular SVM and 20 withheld examples (10 of each class) were used to test the classification accuracy of the classifier. Analyses using varying fractions of training examples and withheld test examples suggested that the 60% training set provided near-saturation discrimination for high contrast stimuli. Using higher proportions of training versus held-out datasets resulted in better performance for low-contrast stimuli, but at the cost of substantially higher standard error of measurement (Supplementary Fig. S4).

Using the training data, we performed a grid search over the RBF's C and gamma parameters , influencing the margin of the decision zone and the influence of each example, respectively, and used the best performing classifier to evaluate performance on the test data (Fig. 1c). We empirically estimated the minimal number of training examples needed for near-optimal classification accuracy for a high acuity stimulus (Fig. 1d) to maximize the number of unique stimuli per recording. The mean and standard error of classification accuracy were estimated via 30 draws of a Monte Carlo cross-validation, a standard ML technique that repeatedly draws a random training and test split without replacement. We chose the 99th percentile of the binomial distribution as our threshold for statistically significant classification accuracy, which is 15/20 = 75% for 20 test stimuli with P(+) = P(−) = 0.5.

All tests were performed on freshly harvested C57BL/6 mouse retinas with removal of the pigment epithelium and cultured in oxygenated ACSF on the MEA. We initially established the effect of stimulus duration on classification accuracy, because this parameter plays a crucial role in the number of stimuli that can be presented in one experiment. Testing the durations of checkerboard presentation from 0.1 second to 1.0 second, we found that a stimulus of at least 0.25 seconds was required to elicit a retinal response that could be classified with statistically significant prediction of retinal acuity (Fig. 2a, top right). Increased duration of stimulus presentation correlated with an increase in measured retinal acuity. Maximal measured acuity for full contrast checkerboard in the wild-type mouse retina was 0.25 cpd (Fig. 2a, top left). To address retinal acuity to kinetic stimuli, we developed a binary, drifting grating stimulus test that incorporates testing for a range of spatial frequencies, durations, contrasts, and drift speeds. Considering that the most abundant RGC cells in the mouse retina, W3 RGCs, have the highest density and smallest receptive field with specialized visual responses to movement, we expected that dynamic stimuli would improve retinal acuity testing.[20] The maximal retinal acuity measurement, 0.4 cpd, was accomplished at the lowest speed recorded, 8°/s, at 100% contrast (Fig. 2b, top right). In comparison with static checkerboard testing, retinas were sensitive to drifting gratings at 0.1-second durations of presentation at a constant speed of 16°/s. Most durations tested had a retinal acuity limit of 0.2 cpd (Fig. 2a, top right). Retinas blind from outer retinal degeneration (*rd1*) did not reach threshold for acuity with either static or dynamic stimulus (lower subpanels in Figs. 2a and 2b). Given that *rd1* retinas have a spontaneous hyperactivity with rhythmic bursting of RGCs, these results indicate that the system does not associate acuity with random spiking.[21,22] This negative control also eliminates direct photoelectric stimulation of MEA as a potential confound for acuity measurements.

By altering contrast between adjacent blocks and stripes in the checkerboard and drifting grating paradigms, contrast sensitivity functions could be derived. Contrast sensitivity (Figs. 2c, 2d) showed maximal sensitivity at 0.15 and 0.22 cpd for static and dynamic measurements, respectively. The curves display a consistent decrease in contrast sensitivity with increasing spatial frequency. This result is comparable with in vivo testing with a two-alternative, forced-choice method.[23]

To begin to analyze the contributions of individual retinal cell types to the overall retinal acuity, we first performed unsupervised clustering analysis of individual sorted cell responses to the 0.5-second bright flash stimulus.[24] This analysis yielded five distinct clusters, with response parameters roughly corresponding to ON- (clusters 3 and 4), OFF- (clusters 0 and 1), and ON–OFF RGCs (cluster 2) (Supplementary Fig. S1a, S1b). Cluster 2 cells were most numerous in this

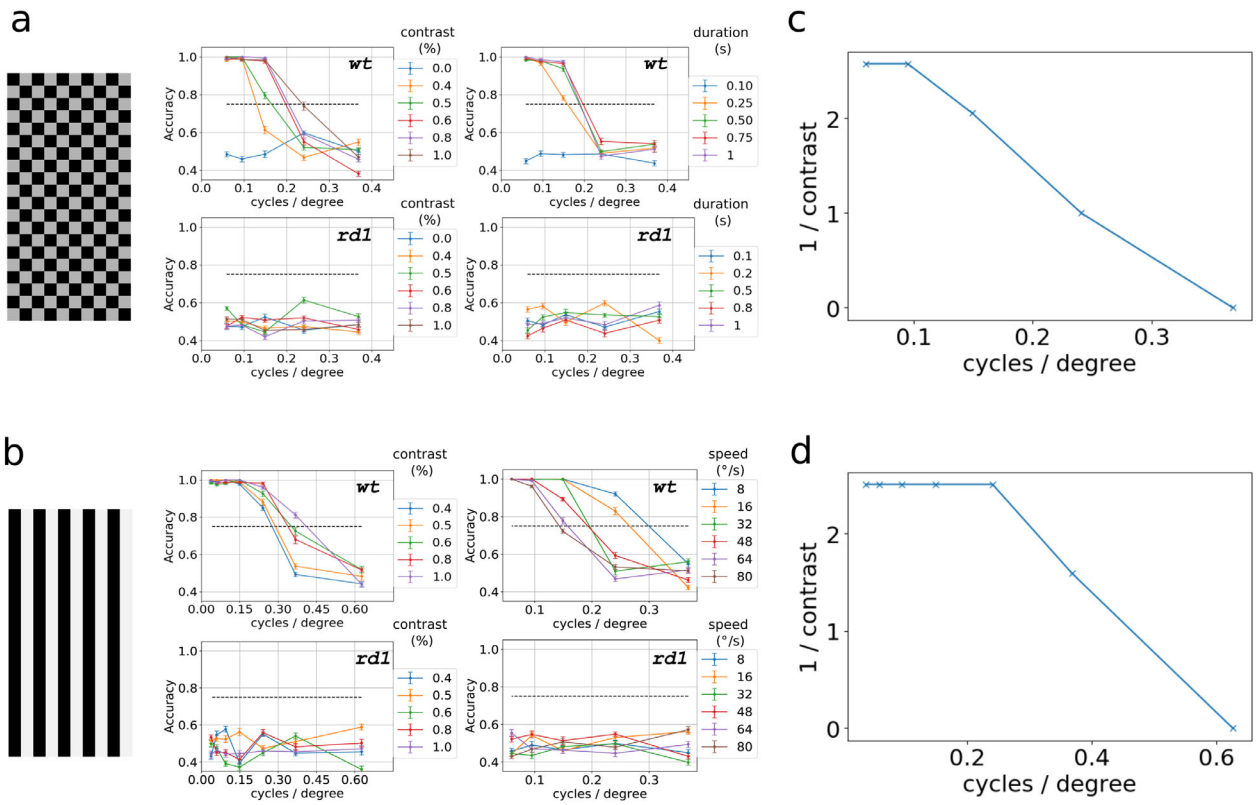


Figure 2. Static and dynamic measure of retinal acuity. (a) SVM classification performance in identifying checkerboard stimulus protocol (see Methods) presented on each trial as a function of spatial frequency for varying contrasts (*left*) and presentation duration (*right*) on *wt* (*top*) and control blind animals, *rd1* (*bottom*). (b) SVM classification performance in identifying grating stimulus protocol (see Methods) presented on each trial as a function of spatial frequency for varying speeds (*left*) and contrasts (*right*) on *wt* (*top*) and control blind animals, *rd1* (*bottom*). Contrast is reported in percentages, speeds as degrees per second, and durations as seconds. The dashed black line denotes the performance expected by chance at 1%. (c and d) Summary of wild-type contrast thresholds for data in Figures 2a and 2b, respectively, are shown as a function of spatial frequency for static and dynamic stimuli. Error bars are ± standard error of the mean. *wt*, C57BL/6J mice; *rd1*; C3H/HeJ (Pde6b^{rd1} / Pde6b^{rd1}) mice. All data points were recorded from one retina and found consistent across $n = 10$ retinas.

analysis. We next performed multiple random samplings of each clustered cell type and attempted to train the SVM program with only those cells (i.e., could the system discriminate motion direction from 1 cell, 2, cells, 5 cells, etc., of a particular cluster). For comparison, we also sampled from all populations by their proportion in the retina (Supplementary Fig. S1c). With the exception of rare cluster 2 (ON–OFF) cells, single cells showed no direction selectivity and SVM could not discern the direction of movement from single cells of any class (Supplementary Fig. S1e). The performance of the ensemble retina improved with increasing numbers of units analyzed, showing nearly perfect classification for high-contrast gratings at a low spatial frequency after analysis of only 10 units. However, even after analysis of 15 cells, the overall performance was far below the performance generated from full retina with 60 units sampled (0.37 cpd) (Supplementary Fig. S4), demonstrating that SVM used aggregate information from the full retinal dataset in task discrimination. As expected, classification for lower contrast stimuli required input from more cells (e.g., SVM was able to correctly discern direction of movement for 98% of 0.15 cpd stimuli at 100% contrast using input from 15 cells vs 80% classification for 50% contrast stripes). When given input only from a single clustered cell type, the SVM measured markedly different retinal acuity depending on cluster (Supplementary Fig. S1c). Although the system could readily classify 0.24 cpd stimulus at 100% contrast from 10 cells from cluster 3 (ON), the system could not reliably distinguish direction of movement from an equivalent number of cluster 0 or cluster 1 cells (predominant -OFF populations). Interestingly, performance for the cluster 3 cells in isolation outperformed performance of the retina as a whole for a given number of cells (i.e., 10 cluster 3 cells showed better retinal acuity measurement than 10 cells distributed across classes). Whether this represents more useful information derived from the concentrated cluster 3

cell population or a dilution of useful information for this specific task from other cell populations was not tested.

Murine Retinal Acuity Shows Regional Variation

The retinas of many species show localized areas of increased acuity such as the fovea in primates and the visual streak in canines.[25] Localized regions of higher acuity have not been described for the rodent retina to date. Supplementary Figure S2 shows a representative analysis of retinal acuity made from a single wild-type mouse retina, divided into quadrants, using the static checkerboard stimulus. ML classification performance on checkerboard stimuli for varying contrasts were dependent on spatial frequency and had a direct correlation with increasing contrast. Overall, the highest retinal acuity was measured from the inferior–nasal quadrant with the inferior–temporal retina very nearly matching those results with few exceptions. At full contrast, the inferior–nasal retinal acuity threshold was 0.25 cpd. The superior–nasal retina outperformed its temporal counterpart with higher accuracy per spatial frequency and contrast measurements with maximal limit at 100% contrast of approximately 0.15 cpd. Analyses were consistent across three retinas; therefore, subsequent experiments were performed on the inferior–nasal sections.

Reversible Pharmacological Inhibition of Retinal Acuity

Having established that EyeCandy could reliably measure retinal acuity of a retina in vitro from RGC recordings, we tested the sensitivity of the system to disruptions in the retinal visual stream. To block the glutamatergic pathway from photoreceptors to RGCs, retinas were treated with glutamatergic inhibitors DNQX and D-AP-5. Classification of retinal acuity in retinas continuously perfused with the inhibitory cocktail failed to reach statistical significance for either static or dynamic stimuli (Supplementary Fig. S3a and S3b, middle). Perfusion washout for 1 hour with ACSF restored significant retinal acuity classification with a small loss in spatial frequency sensitivity when compared to pre-treated retinas (Supplementary Fig. S3a and S3b, left and right panels).

ETDRS Eye Chart Letter Classification

To establish stimuli relevant to clinical human visual assessment, we developed a protocol that incorporated the ETDR Study (ETDRS) eye chart to further test retinal acuity.[26] This optotype chart is routinely used in human clinical trials to measure visual acuity. ETDRS letters, white on black background to decrease photo bleaching, were projected randomly with varying logMARs onto the MEA-mounted retina and RGC population responses were used to classify letter recall accuracy using one-versus-all ML paradigm. Along with letter presentations for 0.5 seconds, the protocol included full field flashes and periods of darkness to avoid adaptation (Fig. 3a). To visualize how acuity effects the similarity of population response to various letters, we used t- distributed stochastic neighbor embedding (t-SNE)[27] to determine the extent to which trials of the same letter cluster together (Fig. 3b). As the logMAR increases above 3.0, clusters of letters separate cleanly, with C and O and N and H remaining in proximity as expected by their high degree of similarity in pixel space. The SVM classifier outperforms what would be possible based on the two-dimensional t-SNE analysis by operating in a high dimensional space with labeled examples to draw optimal hyperplanes. As would be expected for images occupying a greater portion of retina, performance improved substantially with number of units analyzed and showed continued improvement that had not saturated for smaller optotypes at 60 units (Supplementary Fig. S4). The performance of the SVM can be evaluated by examining the confusion matrix C on held-out test data (Supplementary Fig. S5a) with entries, where c_{ij} is the number of stimuli known to be letter i that are predicted to be letter j. Increasing letter size (i.e., logMAR) increased total accuracy, with an almost perfect classification of 99% achieved at a logMAR of 3.0. The minor misclassifications at a logMAR 3.0 of occurred with a single error with N classified as H and a combined total of five errors between O and C. Similar results were obtained across several retinas (Supplementary Fig. S5b). Although this test may not represent an accurate account of a mouse's ability to identify letters in vivo, it demonstrates two important features of the test system: (1) the adaptability to present different types of psychophysical tests and (2) the ability to extract more complex information content from a population of RGCs beyond simple linear shapes.

Measurement of Visual Acuity of Blind Retina After Photopharmacological Treatment

Many vision restoration techniques under development make use of the remaining inner retinal circuitry that persists after photoreceptor degeneration.[28] These methods include implantation of electrode arrays or

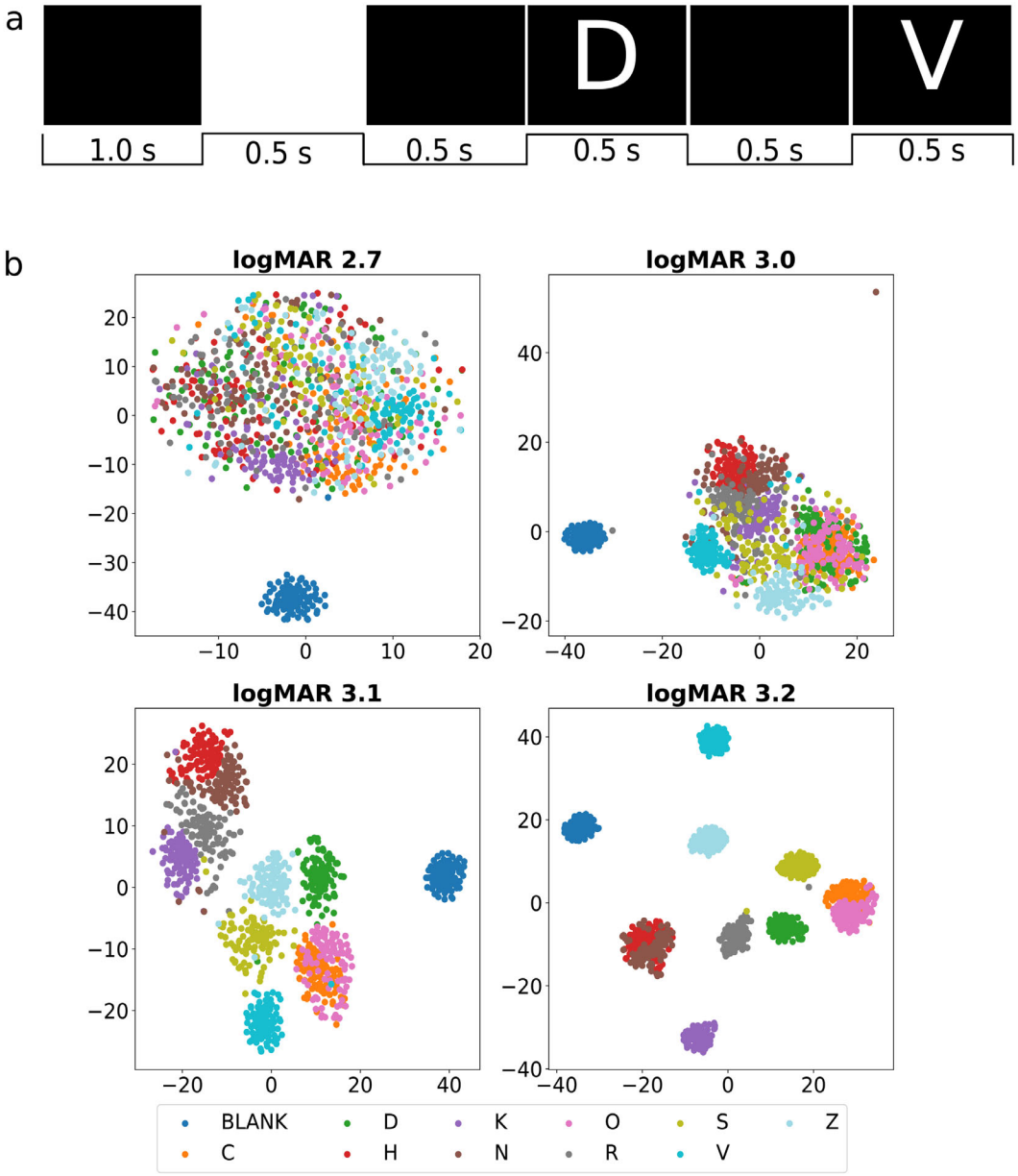
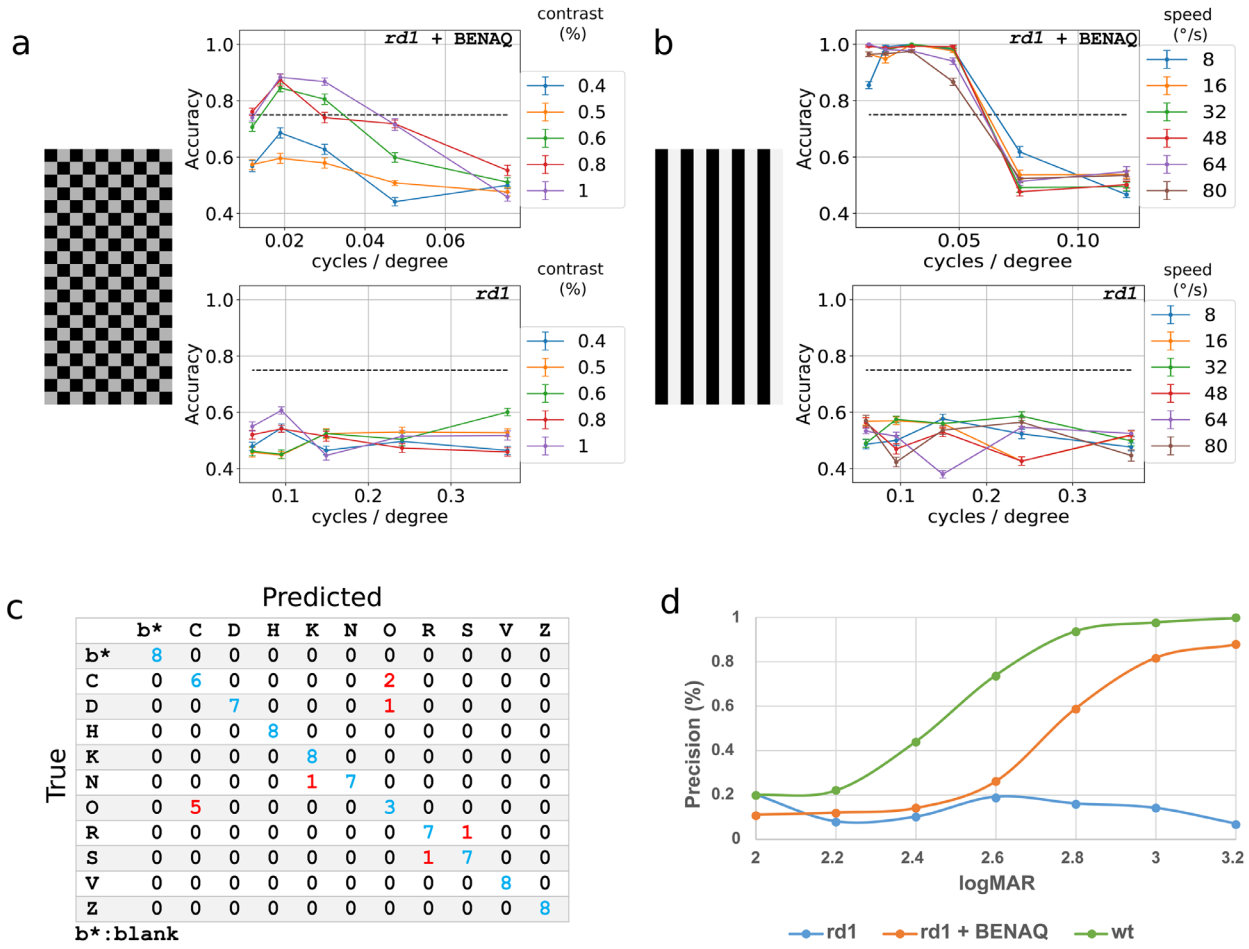


Figure 3. In vitro ETDRS eyechart testing. (a) Stimulation protocol of ETDRS eyechart. (b) Two-dimensional visualization of population responses during individual trials using t-distributed stochastic neighbor embedding ($n = 6$ retinas).

photodiodes that electrically stimulate the remaining inner retina (for review see[29]) and optogenetic approaches in which light-sensitive proteins are transgenically expressed in surviving cells.[30–32] Previously, we and others have shown that light sensitivity can be restored in *rd1* mouse models with synthetic azobenzene photoswitch compounds.[33] These compounds are light-activated, voltage-gated potassium channel blockers which confer light sensitivity onto RGCs in the absence of photoreceptors. We subjected blind *rd1* retinas treated with the photoswitch BENAQ to static and kinetic visual tests. The highest static retinal acuity, 0.05 cpd, was measured at full contrast (Fig. 4a, top). An improvement in retinal acuity to 0.08 cpd was measured with the kinetic grating test (Fig. 4b, top) with the slowest speed tested (8°/s) at 100% contrast. Figure 4b demonstrates that blind retinas treated with BENAQ are motion sensitive with a consistent trend of increasing retinal acuity with decreasing grating speed. In the more complex ETDRS test, an overall 88% precision in letter recognition was restored in BENAQ-treated retinas at the lowest retinal acuity tested, logMAR 3.2. Figure 4c (left, control; right, BENAQ treated) shows a single run



True \ Predicted	b*	C	D	H	K	N	O	R	S	V	Z
b*	8	0	0	0	0	0	0	0	0	0	0
C	0	6	0	0	0	0	2	0	0	0	0
D	0	0	7	0	0	0	1	0	0	0	0
H	0	0	0	8	0	0	0	0	0	0	0
K	0	0	0	0	8	0	0	0	0	0	0
N	0	0	0	0	1	7	0	0	0	0	0
O	0	5	0	0	0	0	3	0	0	0	0
R	0	0	0	0	0	0	0	7	1	0	0
S	0	0	0	0	0	0	0	1	7	0	0
V	0	0	0	0	0	0	0	0	0	8	0
Z	0	0	0	0	0	0	0	0	0	0	8

Figure 4. Retinal acuity of BENAQ-treated *rd1* retina. (a) SVM classification performance in identifying checkerboard and (b) grating protocols presented on each trial as a function of spatial frequency for varying contrasts and speeds on BENAQ-treated *rd1* retinas (*top*) and untreated (*bottom*). (c) Confusion matrix classifying 10 letters of the ETDRS chart each projected individually for logMAR 3.2 on *rd1* retina treated with BENAQ. Blue letters indicate correctly predicted trials per letter and red indicates incorrect letter predictions. d, Percent precision of SVM classification in letter recall as a function of logMAR for all sizes tested in control (*wt* retina, *green*), *rd1* blind retina (*blue*) and *rd1* BENAQ-treated retina (*orange*). Error bars are ± standard error of the mean. *wt*, C57BL/6J mice; *rd1*; C3H/HeJ (*Pde6b*^{rd1} / *Pde6b*^{rd1}) mice (n = 4 retinas).

with 8 test trials per 10 letters presented. The only significant errors occurred across classification with the letters O and C; all other letters achieved over 88% precision. Overall, retinal acuity for BENAQ-treated *rd1* retinas showed approximately 0.5 log decreased acuity compared with wild-type animals (Fig. 4d).

Discussion

We describe an automated system for estimating a lower bound of retinal acuity of the murine retina in vitro. This system uses ML to associate trains of action potentials emanating from RGCs with specific stimuli. The system assumes that, if a ML algorithm can accurately classify a particular stimulus from RGC spike trains, then the resolving power of the retina must have been sufficient for this task. In wild-type mice, a maximal retinal acuity of 0.4 cpd was measured using a drifting grating presented with white light on the inferior–nasal quadrant of the retina. In studies using the visual water maze task in wild-type C57/Bl6 animals, the range of reported acuities varies between 0.3 and 0.5 cpd.[1,2] Although it is possible that either the mouse brain or the computer classification system could extract more useful information for classification than the other, finding that acuity measurements made by the computerized system are very close to those measured in mice by behavioral testing suggests comparable ability to extract visual

information. Because our system only uses a small fraction of all RGCs in the retina (<10%[34]), the measured retinal acuity can be viewed as an estimated lower bound on the information transmitted by the total retina to the brain. Rompani et al.[35] showed that in most cases only a small number of RGC's relay information to a single mouse lateral geniculate nucleus. Recent findings suggest that as few as 14 RGCs are involved in the process of complete summation underlying human luminance detection.[36] Here we show that our system is sensitive to several factors important to vertebrate vision with an average sampling of approximately 250 RGCs.

Using EyeCandy, we were able to investigate retinal acuity for three experimental conditions that would be challenging for behavioral testing. First, we were able to establish the relative acuity of different sections of the mouse retina, finding the highest acuity in the inferonasal quadrant. These results are consistent with reports of increased photoreceptor density described in the inferior mouse retina[37]; however, because the measured acuity was still well below the theoretical limits predicted by cone spacing in the mouse retina, other mechanisms (such as regionally varying circuitry within the retina[38]) may be responsible. Second, we were able to measure the acute change in acuity after pharmacological administration of glutamatergic blockers and establish reversibility of this intervention. Systemic administration of these agents would induce potentially lethal neurologic dysfunction. The EyeCandy platform may thus have utility for toxicity testing of retinal function in vitro. Finally, we were able to demonstrate the rapid return of light-dependent retinal function associated with acute treatment of the retina with the photoswitch potassium channel blocker azobenzene, BENAQ, and were able to demonstrate restoration of acuity to both static and dynamic stimuli to approximately 0.5 log less than wild type. Because the half-life of some photoswitches may be on the order of hours,[33] learning-based acuity methods (such as forced choice mazes) may not be used successfully with these agents.

Numerous previous studies have analyzed the physiologic responses of single ganglion cells to visual stimuli in vitro and generated estimates of spatial acuity of different ganglion cell classes.[39,40] The current technique complements these methods. First, by analyzing the aggregate information content from a full multielectrode array, a global retinal acuity is estimated. This acuity seems to exceed the acuity attributable to any individual cell or cell type (Supplementary Figs. S1, S4), and thus specifically answers the question as to whether the aggregate retinal response encodes sufficient information content for a discriminative task. In future studies, the redaction of very specific RGC types with known spatial and temporal tuning can be performed to understand the necessity of particular classes of RGCs to acuity to specific stimuli,[41] whereas the determination of acuity achieved with only these cell types can determine the sufficiency of ganglion cell subtypes in generating acuity. Second, the system can be used for analysis of more complex stimuli, such as ETDRS optotypes, that exceed the receptive fields of individual ganglion cells. Finally, future iterations of this system may be able to perform ML-guided reconstruction of images from retinal output (in effect, using ML to solve the retinal encoding problem), a task that would be impossible from single cell recordings.

Numerous opportunities for improvement of the system exist. Although we have shown that the number of training and test sets used in the current protocol is adequate to recapitulate retinal acuity comparable with whole animal visual acuity, it is likely that larger training sets may allow increased extraction of information and improved performance for low-contrast stimuli or for restored responses with photoswitches or other methods for retinal reanimation. We focused on the mouse retina as a well-studied model with readily available genetic tools.

However, it must be recognized that the mouse retina has inherently low acuity compared with primate retina. The EyeCandy system should be directly adaptable to other species in the future. We may also have limitations (particularly in primate retina) in the relatively low sampling density of the MEA used in this study (with 60 electrodes). Newer generation MEAs have two to three orders of magnitude greater electrode coverage and may allow extraction of more useful information for higher estimates of retinal acuity.[42] It is also not clear that all RGC cell types are equivalently sampled (particularly displaced ganglion cells); alternative arrays with three-dimensional penetrating electrodes may address this challenge.[43] The EyeCandy projection system is capable of projecting more complex stimuli, allowing for testing of more physiologically relevant probes such as real-world scenes. Finally, although we have used ML for the discrimination of a small set of image possibilities, this technique could also be used to reconstruct images from aggregate RGC firing output and, in this way, be useful in analyzing the visual code from retina to brain.

The demonstration of measurable acuity from photoswitch-treated *rd1* retinas suggests that the EyeCandy system may be used to estimate and compare the effects of vision restoration methodologies, such as stem cell photoreceptor replacement,[44]

small molecule photoswitches, optogenetic gene therapy,[45] and opto-electronic stimulation[46] on potential acuity ex vivo. In the current work, we show proof of concept for the photoswitch BENAQ,[47] but several other photoswitches with varying cell targets in the retina remain to be tested.[48] Additionally, the treatment of *rd* retinas with agents such as meclofenamic acid that decrease spontaneous oscillations[49] or with inhibitors of retinoic acid that reduce ganglion cell hyperactivity after retinal degeneration[50] may increase the retinal acuity of retinas treated with vision restoration methods.

The current system is semiautomated, requiring the user to harvest and mount the retina on the MEA, run the EyeCandy stimulus package, spike sort the resulting MEA recordings, and analyze these with Glia. No experience with programming is required by the operator. The system makes use of free and open source software including Scikit-Learn library for Python. Currently, the system has an optical limitation of presenting 6 cpd for a grating stimulus, which is far beyond the visual limitations of the mouse retina but within the anticipated acuity limit of primate retinas. Further optimizations of the system can be achieved by upgrading the optics, using a projector with higher resolution, and using high-density MEAs. We anticipate this system will have broad applications to problems in vision science, particularly in translational studies. Questions such as the relative effects of disease on overall acuity and individual cell subtypes, or the relative efficacy of vision restoration technologies on different cell classes and their contribution to overall acuity may be addressed in studies in the near future.

Acknowledgments

The authors thank J. Neitz, M. Manookin, A.Y. Lee, R. Sabesan, and A. Sandt for their assistance.

We thank Photoswitch Biosciences for providing the BENAQ compound.

Supported by NIH R24EY023937, P30EY001730, an unrestricted grant from Research to Prevent Blindness, the Murdoch Charitable Trust, and the Mark J. Daily, MD Research Fund.

Disclosure: **D. Babino**, None; **T.S. Benster**, None; **L. Laprell**, None; **R.N. Van Gelder**, None

* DB and TSB contributed equally to this work.

References

1. Pinto LH, Enroth-Cugell C. Tests of the mouse visual system. *Mamm Genome*. 2000;11:531–536.
2. Prusky GT, West PW, Douglas RM. Behavioral assessment of visual acuity in mice and rats. *Vision Res*. 2000;40:2201–2209.
3. Wong AA, Brown RE. Visual detection, pattern discrimination and visual acuity in 14 strains of mice. *Genes Brain Behav*. 2006;5:389–403.
4. Masland RH. The neuronal organization of the retina. *Neuron*. 2012;76:266–280.
5. Fitzhugh A. A statistical analyzer for optic nerve messages. *J Gen Physiol*. 1958;41:675–692.
6. Fitzhugh R. The statistical detection of threshold signals in the retina. *J Gen Physiol*. 1957;40:925–948.
7. Warland DK, Reinagel P, Meister M. Decoding visual information from a population of retinal ganglion cells. *J Neurophysiol*. 1997;78:2336–2350.
8. Barrett JM, Hilgen G, Sernagor E. Dampening spontaneous activity improves the light sensitivity and spatial acuity of optogenetic retinal prosthetic responses. *Sci Rep*. 2016;6:33565.
9. Jacobs AL, Fridman G, Douglas RM, et al. Ruling out and ruling in neural codes. *Proc Natl Acad Sci USA*. 2009;106:5936–5941.
10. Quian Quiroga R, Panzeri S. Extracting information from neuronal populations: information theory and decoding approaches. *Nature Rev Neurosci*. 2009;10:173–185.
11. Wen H, Shi J, Zhang Y, Lu KH, Cao J, Liu Z. Neural encoding and decoding with deep learning for dynamic natural vision. *Cereb Cortex*. 2018;28:4136–4160.
12. Baylor DA, Nunn BJ, Schnapf JL. The photocurrent, noise and spectral sensitivity of rods of the monkey Macaca fascicularis. *J Physiol*. 1984;357:575–607.
13. Ankerst M, Breunig MM, Kriegel H-P, Sander J. OPTICS: ordering points to identify the clustering structure. *SIGMOD Rec*. 1999;28:49–60.
14. Ram A, Jalal S, Jalal AS, Kumar M. A density based algorithm for discovering density varied clusters in large spatial databases. *Int J Comput Appl*. 2010;3:1–4.
15. Xu Q-S, Liang Y-Z. Monte Carlo cross validation. *Chemometr Intell Lab Syst*. 2001;56:1–11.
16. Cortes C, Vapnik V. Support-vector networks. *Mach Learn*. 1995;20:273–297.
17. Knerr S, Personnaz L, Dreyfus G. Single-layer learning revisited: a stepwise procedure for build-

ing and training a neural network. In: *Neurocomputing*. New York: Springer; 1990:41–50.
18. Zhang P, Zam A, Jian Y, et al. In vivo wide-field multispectral scanning laser ophthalmoscopy-optical coherence tomography mouse retinal imager: longitudinal imaging of ganglion cells, microglia, and Muller glia, and mapping of the mouse retinal and choroidal vasculature. *J Biomed Opt*. 2015;20:126005.
19. Liu Z, Elashoff D, Piantadosi S. Sparse support vector machines with L0 approximation for ultra-high dimensional omics data. *Artif Intell Med*. 2019;96:134–141.
20. Zhang Y, Kim IJ, Sanes JR, Meister M. The most numerous ganglion cell type of the mouse retina is a selective feature detector. *Proc Natl Acad Sci USA*. 2012;109:E2391–2398.
21. Stasheff SF. Emergence of sustained spontaneous hyperactivity and temporary preservation of OFF responses in ganglion cells of the retinal degeneration (rd1) mouse. *J Neurophysiol*. 2008;99:1408–1421.
22. Trenholm S, Awatramani GB. Origins of spontaneous activity in the degenerating retina. *Front Cell Neurosci*. 2015;9:277.
23. Umino Y, Solessio E, Barlow RB. Speed, spatial, and temporal tuning of rod and cone vision in mouse. *J Neurosci*. 2008;28:189–198.
24. Ghandeharizadeh S, Faloutsos C, Delis A, Association for Computing M. *Proceedings of the 1999 ACM SIGMOD International Conference on Management of Data: SIGMOD '99, Philadelphia, PA, USA, June 1-3, 1999*. New York: Association for Computing Machinery; 1999:xii, 602 p.
25. Mowat FM, Petersen-Jones SM, Williamson H, et al. Topographical characterization of cone photoreceptors and the area centralis of the canine retina. *Mol Vis*. 2008;14:2518–2527.
26. Ferris FL, Kassoff A, Bresnick GH, Bailey I. New visual acuity charts for clinical research. *Am J Ophthalmol*. 1982;94:91–96.
27. van der Maaten LHG. Visualizing Data using t-SNE. *J Mach Learn Res*. 2008;9:2579–2605.
28. Stone JL, Barlow WE, Humayun MS, de Juan E, Milam AH. Morphometric analysis of macular photoreceptors and ganglion cells in retinas with retinitis pigmentosa. *Arch Ophthalmol*. 1992;110:1634–1639.
29. Chuang AT, Margo CE, Greenberg PB. Retinal implants: a systematic review. *Br J Ophthalmol*. 2014;98:852–856.
30. Bi A, Cui J, Ma YP, et al. Ectopic expression of a microbial-type rhodopsin restores visual responses in mice with photoreceptor degeneration. *Neuron*. 2006;50:23–33.
31. Busskamp V, Duebel J, Balya D, et al. Genetic reactivation of cone photoreceptors restores visual responses in retinitis pigmentosa. *Science*. 2010;329:413–417.
32. Sahel JA, Roska B. Gene therapy for blindness. *Annu Rev Neurosci*. 2013;36:467–488.
33. Laprell L, Tochitsky I, Kaur K, et al. Photopharmacological control of bipolar cells restores visual function in blind mice. *J Clin Invest*. 2017;127:2598–2611.
34. Drager UC, Olsen JF. Ganglion cell distribution in the retina of the mouse. *Invest Ophthalmol Vis Sci*. 1981;20:285–293.
35. Rompani SB, Mullner FE, Wanner A, et al. Different Modes of Visual Integration in the Lateral Geniculate Nucleus Revealed by Single-Cell-Initiated Transsynaptic Tracing. *Neuron*. 2017;93:767–776.e766.
36. Kwon M, Liu R. Linkage between retinal ganglion cell density and the nonuniform spatial integration across the visual field. *Proc Natl Acad Sci USA*. 2019;116:3827–3836.
37. Ortin-Martinez A, Nadal-Nicolas FM, Jimenez-Lopez M, et al. Number and distribution of mouse retinal cone photoreceptors: differences between an albino (Swiss) and a pigmented (C57/BL6) strain. *PLoS One*. 2014;9:e102392.
38. Anderson JR, Jones BW, Yang JH, et al. A computational framework for ultrastructural mapping of neural circuitry. *PLoS Biol*. 2009;7:e1000074.
39. Hwang CK, Chaurasia SS, Jackson CR, Chan GC, Storm DR, Iuvone PM. Circadian rhythm of contrast sensitivity is regulated by a dopamine-neuronal PAS-domain protein 2-adenylyl cyclase 1 signaling pathway in retinal ganglion cells. *J Neurosci*. 2013;33:14989–14997.
40. Pan F. Defocused image changes signaling of ganglion cells in the mouse retina. *Cells*. 2019;8:640.
41. Jacoby J, Schwartz GW. Three small-receptive-field ganglion cells in the mouse retina are distinctly tuned to size, speed, and object motion. *J Neurosci*. 2017;37:610–625.
42. Jouty J, Hilgen G, Sernagor E, Hennig MH. Non-parametric physiological classification of retinal ganglion cells in the mouse retina. *Front Cell Neurosci*. 2018;12:481.
43. Zuo L, Yu S, Briggs CA, Kantor S, Pan JY. Design and fabrication of a three-dimensional multi-electrode array for neuron electrophysiology. *J Biomech Eng*. 2017;139:121011.

44. Van Gelder RN, Chiang MF, Dyer MA, et al. Regenerative and restorative medicine for eye disease. *Nat Med*. 2022;28:1149–1156.
45. Sahel JA, Boulanger-Scemama E, Pagot C, et al. Partial recovery of visual function in a blind patient after optogenetic therapy. *Nat Med*. 2021;27:1223–1229.
46. Palanker D, Le Mer Y, Mohand-Said S, Sahel JA. Simultaneous perception of prosthetic and natural vision in AMD patients. *Nat Commun*. 2022;13:513.
47. Tochitsky I, Trautman J, Gallerani N, Malis JG, Kramer RH. Restoring visual function to the blind retina with a potent, safe and long-lasting photoswitch. *Sci Rep*. 2017;7:45487.
48. Hull K, Benster T, Manookin MB, Trauner D, Van Gelder RN, Laprell L. Photopharmacologic vision restoration reduces pathological rhythmic field potentials in blind mouse retina. *Sci Rep*. 2019;9:13561.
49. Barrett JM, Degenaar P, Sernagor E. Blockade of pathological retinal ganglion cell hyperactivity improves optogenetically evoked light responses in rd1 mice. *Front Cell Neurosci*. 2015;9:330.
50. Telias M, Denlinger B, Helft Z, Thornton C, Beckwith-Cohen B, Kramer RH. Retinoic acid induces hyperactivity, and blocking its receptor unmasks light responses and augments vision in retinal degeneration. *Neuron*. 2019;102:574–586.e575.