Predicting Visual Sensitivity in Retinal Prosthesis Patients

Alan Horsager,1,2 Scott H. Greenwald,3 James D. Weiland,4 Mark S. Humayun,4 Robert J. Greenberg,3 Matthew J. McMahon,3 Geoffrey M. Boynton,4 and Ione Fine5

Purpose. With the long-term goal of restoring functional vision in patients with retinal degenerative diseases, the eyes of blind human subjects were implanted chronically with epiretinal prostheses consisting of two-dimensional electrode arrays that directly stimulated cells of the neural retina.

Methods. Psychophysical techniques were used to measure the brightness of electrically generated percepts on single electrodes using a variety of electrical stimulation patterns.

Results. It was possible to predict the sensitivity of the human visual system to a wide variety of retinal electrical stimulation patterns using a simple and biologically plausible model.

Conclusions. This is the first study to demonstrate that, on the single-electrode level, retinal electrical stimulation in humans can produce visual qualia that are predictable using a quantitative model, a prerequisite for a successful retinal prosthesis. (ClinicalTrials.gov number, NCT00279500.) (Invest Ophthalmol Vis Sci. 2009;50:1483-1491) DOI:10.1167/iovs.08-2595

Retinitis pigmentosa and age-related macular degeneration are frequent causes of blindness in the developed world.1-3 Both diseases begin with the degeneration of photoreceptors, though in later stages the number of bipolar, amacrine, and ganglion cells are significantly decreased4-6 and their spatial organization and circuitry are significantly disorganized.7 More than 150 different gene mutations result in photoreceptor diseases for which there are currently no Food and Drug Administration (FDA)–approved treatments,8 motivating the development of technologies to restore visual function that do not require targeting each genetic defect independently. One approach is to target bipolar and ganglion cells with engineered photo gates9-11 and light-sensitive proteins.12-17 However, to create perceptually meaningful images, it is necessary to predictively generate a range of brightness levels over both space and time. Although the literature examining the perceptual consequences of electrical stimulation has a long history,19,21,22-34 there is still relatively little data in humans systematically quantifying the effects of retinal electrical stimulation as a function of stimulation current levels and the temporal stimulation pattern.

In the current study, at the single electrode level, retinal electrical stimulation resulted in predictable visual qualia that can be described with a relatively simple linear-nonlinear model that predicts the relationship between electrical stimulation and sensitivity for a wide variety of temporally varying stimulation patterns. This model not only can be used to determine the “optimal” pattern of stimulation, given a variety of engineering constraints (such as stimulating at safe levels of charge density and minimizing overall charge), but may also provide some insight into the neural pathways that underlie the perceptual effects of electrical stimulation.

Materials and Methods

Subjects

We examined two patients who underwent ocular implantation of 16-electrode retinal prostheses (Second Sight Medical Products, Inc., Sylmar, CA) that remained in place for a prolonged period. These two patients, S05 and S06, were 59 and 55 years old at implantation (in 2004), respectively. Before surgery, subject S05 had bare light perception (BLP) in the implanted eye, was blind for 8 years before implantation, and was 59 years of age when the prosthesis was implanted. Subject S06 had no light perception (NLP), was blind for 10.5 years before implantation, and was 55 years of age at implantation. These tests were performed during a period of approximately 90 to 1170 days after implantation in the case of S05 and 30 to 1110 days after implantation in the case of S06.

These two patients were a subset of six patients who have undergone implantation since February 2002. The other four patients were excluded for a variety of reasons: one because of geographic location, two because of unrelated medical conditions, and one because the array cable became exposed. Because the cardiac status of this patient precluded general anesthesia, the multiwire cable connecting the array...
a subdermal coil implanted in the patient’s temporal skull. This signal was then sent through a parallel system of wires to the epiretinally implanted electrode array. Note that power and signal information could be independently controlled for each electrode.

during the first few weeks after implantation, thresholds and matching judgments tended to be fairly variable. The gradual increase in the stability of thresholds (and probably a slight reduction in threshold) is likely to be partially due to the subjects’ becoming increasingly expert observers; similar effects have been noted in traditional psychophysical experiments. These experiments were begun several sessions after implantation, at a point where additional learning effects were likely to be fairly minimal. It should be noted that we used a large number of catch trials and monitored and compensated for false-positive responses to prevent changes in subjects’ criteria over time from affecting the thresholds. Half of the trials were stimulus-absent catch trials, and these trials were interleaved randomly with the stimulus-present trials. The catch trials were used to compensate for any change in criterion with time or practice. Current amplitude was varied using a three-up, one-down staircase procedure to find the threshold current amplitude needed for the subjects to see the stimulus-present trials, corrected for the false-alarm rate. If the subject responded correctly three times in a row, the task was made more difficult by decreasing the current amplitude. If the subject answered incorrectly on any trial, the task was made easier by increasing the current amplitude. During each staircase, only amplitude varied. All other parameters (frequency, pulse width, pulse train duration, and the number of pulses) were held constant. Each threshold was based on fitting a Weibull function to a minimum of 125 trials, and error bars were estimated by Monte-Carlo simulation. See Supplementary Data 1 for examples of psychometric functions collected for both threshold and suprathreshold data. Weibull functions generally provided good fits to threshold data, and failures to find a good fit appeared to be nonsystematic and due to noise. We did not observe any systematic variation of slope across the experimental conditions tested.

**Suprathreshold Brightness Matching Judgments.** Suprathreshold brightness-matching was performed on single electrodes by using a two-interval, forced-choice procedure. Each trial contained two intervals, with each interval containing a pulse train of a different frequency. Each interval was separated by 1050 ms. For example, interval 1 might contain a 15 Hz pulse train and interval 2 might contain a 45 Hz pulse train. Subjects were asked to report which interval contained the brighter stimulus. The order of the intervals was randomized on each trial to minimize potential presentation order biases (see also Supplementary Data 2). The time between each trial varied depending on the subject’s response time, but generally ranged between 3 and 5 seconds. A one-up, one-down staircase method was used to adjust the amplitude of the higher frequency pulse train based on the observer’s response. Occasionally (~1/10 cases), a dark phosphene rather than a white or yellow phosphene was elicited for a particular electrode. Because all comparisons were performed within a single electrode, the subjects were never asked to match the brightness of a bright to a dark
phosphene. When brightness-matching a dark phosphene, subjects reported which interval contained the darker phosphene.

The first brightness match was made by fixing the amplitude of a “standard” 5-Hz pulse train (a single pulse within a 200-ms window) to be two or three times threshold amplitude, and finding the amplitude needed for a 15-Hz “test” pulse train to match the brightness of the standard pulse train. Using the measured brightness-matched value of the 15-Hz pulse train, the 15-Hz pulse train then became the standard pulse train and was compared in brightness to a 45-Hz test pulse train and so on. It should be noted that this technique led to the accumulation of small errors as the standard changed, but unfortunately software limitations made it impossible either to use a fixed standard or perform all pair-wise comparisons. On a few electrodes we collected data “backwards,” beginning with a 225-Hz stimulus as the standard, and the results were qualitatively similar.

Each brightness match was based on a minimum of 80 trials. A cumulative normal was used to find the point of equibrightness, and error bars were again estimated by using an adaptive-sampling Monte Carlo simulation.28 Each individual psychometric function was inspected to make sure that an adequate fit was obtained, and data were recollected if fits were inadequate (based either on the estimated error or visual inspection). A cumulative normal appeared to provide a good fit to the data, see Supplementary Data 1.

Using this method, we were able to obtain an isobrightness curve that represented the current amplitude needed to maintain the same subjective brightness across a wide range of frequencies.

Stimulus Set. In each of our two subjects, we measured detection thresholds for 10 different categories of stimulation pattern (see Figs. 5, 4, 5) and suprathreshold perceived brightness for six different categories of stimulation pattern (see Figs. 6, 7). Data were collected from 12 electrodes across the two subjects. Across these 12 electrodes, we collected 534 threshold and 116 suprathreshold measurements in total. See Supplementary Data 3 and 4 for additional threshold and suprathreshold data.

Because of the limited availability of our subjects, we were unable to collect data on all electrodes. The only criterion applied to choose electrodes are shown in Supplementary Data 3.

Determining the Optimal Parameters of the Model
Optimization was performed with a subset of the full set of data: two electrodes for each of the two patients (S05, B3 and C2; S06, B1 and C2). Threshold and suprathreshold fits and parameter values for these electrodes are shown in Supplementary Data 3.

The parameter values $\tau_1$, $\tau_2$, and $\tau_3$ were optimized across the seven threshold and three suprathreshold experiments with a standard least-squares error minimization technique. The parameters $\epsilon$ (linear shift as a function of charge) and $\beta$ (power nonlinearity) were fit separately for threshold and suprathreshold levels of stimulation. When suprathreshold data were fitted, $\epsilon$ and $\beta$ were allowed to vary across different levels of apparent brightness.

The parameter that represented the model output at threshold, $\hat{\theta}$, was allowed to vary across each experiment on a given electrode. Variation in $\hat{\theta}$ accounts for differences in mean sensitivity between the two patients, differences in sensitivity across electrodes, and slight changes in electrode sensitivity over time. The set of data in this article were collected over slightly more than a 2-year period. During this period, we observed gradual changes in sensitivity that appeared to be mainly due to slight changes in the position of the electrode array over time.25 We also saw some variation in threshold within individuals across different testing sessions. Similar variability (of roughly the same magnitude) has been reported for sensitivity (perimetry) data in visually normal control subjects and appears to increase with age.25,26
RESULTS

Patients typically reported that phosphenes appeared white or yellow and were round or oval. At suprathreshold, percepts were reported as being approximately 0.5 to 2 in. in diameter near the point of equal brightness. The subjects were told to base responses on changes in brightness and ignore other qualitative changes, and they reported that the task was easy.

One concern was that we might see changes of sensitivity over the time course of a single session. We therefore performed control experiments, testing for changes in sensitivity over time on S6 (S5 was not available for testing at the time of these control experiments) to examine whether there were changes in sensitivity between (1) the first and second intervals of a single trial, (2) between the beginning and end of a run (~150 trials), and (3) between the beginning and end of an entire testing session. We also used previously collected data on both S5 and S6 to examine desensitization across the two intervals of a trial and between the beginning and end of a session. As reported in Supplementary Data 5, we saw evidence of a slight loss in sensitivity between the two intervals of a single trial. In our brightness-matching task the presentation order of the intervals was randomized, thereby minimizing interval order effects. Based on the results of a control experiment described in Supplementary Data 5, conservative calculations estimate that any resulting bias never exceeded 3% of the estimated brightness match and was generally much smaller (<1 μA for almost all data points). Given the small size of these biases, for simplicity sake, we did not include interinterval interactions in our model. As described in Supplementary Data 5, we found no evidence of desensitization within a run or within a session. As a result, we feel justified in treating each trial as an independent event.

Predicting Visual Sensitivity for Novel Electrodes and Temporal Patterns

After optimizing the model using a subset of the full set of data, we averaged the best-fitting parameters values for \( \tau_1, \tau_2, \tau_3, \beta, \) and \( \varepsilon \) across all the four electrodes used for optimization and used these mean values to predict threshold and suprathreshold data for novel electrodes. For these novel electrodes, the only parameter allowed to vary across each experiment was the threshold parameter, \( \theta. \) Values of \( \theta \) for these novel electrodes are shown in Supplementary Data 4.

Figure 3 shows subject thresholds (gray squares) and model predictions (solid line) for a single biphasic pulse presented on a novel electrode for both subjects. Note the logarithmic horizontal axis. Figure 4 shows threshold data and predictions on a novel electrode for 200-ms pulse trains whose frequency varied between 5 and 225 Hz. Pulse trains...
consisted of either 0.075 (Fig. 4B)- or 0.975 (Fig. 4C)-ms charge-balanced biphasic pulses. Figure 5 shows threshold data and predictions for pulse trains containing either 2 (Fig. 5B) or 15 (Fig. 5C) pulses, whose frequency was varied between 3 and 3333 Hz. Figure 6 shows suprathreshold data and model predictions for 200-ms pulse trains consisting of either 0.075 (Fig. 6B)- or 0.975 (Fig. 6C)-ms pulses, whose frequency varied between 5 and 135 Hz. Additional threshold and suprathreshold data and predictions are shown in Supplementary Data 4. The model and parameter values generalized successfully to predict data on novel electrodes.

We then examined the ability of the model to predict responses to novel pulse train waveforms not used to optimize model parameters (Fig. 7). We again used the same fixed values for $\tau_1$, $\tau_2$, $\tau_3$, $\beta$, and $\epsilon$, based on the electrodes and stimulus patterns used for optimization, and the only parameter allowed to vary across each experiment was the threshold parameter $\theta$. The novel waveforms consisted of repeated bursts of three pulses with a variable interburst delay. Data for this novel waveform were collected at both threshold and suprathreshold levels of stimulation on novel electrodes not used for the original model fits (see Supplementary Data 4 for threshold predictions). The model and parameter values generalized, to successfully predict these data from a novel stimulation pattern on a novel electrode.

**Model Power**

The power of this model was significantly higher than that of a less constrained model where $\tau_1$, $\tau_2$, $\tau_3$, $\beta$, and $\theta$ were all allowed to vary across each experiment and electrode ($F$ test, $I_{\text{ratio}} = 0.6483, \alpha < 0.01$). The power of the model was also significantly better than that of a variety of simplified versions of the model, as well as slight variants of this model, see Supplementary Data 4. There is of course an infinite supply of alternative models of the same or greater complexity, however none of the alternative models that we examined performed as well as the model described here.

Although there were some small deviations between the model and the data, these deviations were certainly no larger than the deviations that are generally found for comparable models of psychophysical performance for temporal light stimuli e.g.,34,37 There were some systematic deviations between the model and performance for long pulses at suprathreshold levels of stimulation (Figs. 4B, 4C). It is perhaps not surprising that our model did not generalize completely to suprathreshold levels of stimulation with long pulses, given that neurophysiological data on the effects of electrical stimulation suggest that presynaptic cells have a much larger influence on neuronal responses to such stimuli.38 Models of greater complexity were capable of better capturing the long pulse data, but led to overfitting across the full data set (see Supplementary Data 4).

This model is also highly constrained compared with analogous models that have been used to model human responses to temporally varying light patterns.34,39 In these psychophysical models, a similar number of parameters are required, a smaller range of temporal patterns are generally modeled, and parameters are typically allowed to vary across subjects. This model is also constrained relative to a similar model of tempo-
We found that perceptual responses to retinal electrical stimulation could be predicted with a surprisingly simple model that resembles models of ganglion cell firing behavior during contrast adaptation. In human temporal integration of light stimuli, and auditory processing in cochlear implant users. Our model, like those describing the perception of light stimuli, presumably approximates the responses of neuronal populations. Observed thresholds were comparable to in vitro thresholds in primates for equivalent pulse widths and electrode size. (Also see deBalthasar et al., for a direct comparison of thresholds in our patients to those obtained within in vitro animal models.) Subjects with normal vision can reliably...
FIGURE 8. Efficiency predictions for a 500-ms pulse train. In each panel the x-axis represents pulse width on a logarithmic axis, and the y-axis represents current density. Red dashed lines: represent a current amplitude limit of 200 µA; yellow dashed lines: represent the constraint that stimulation must occur above the critical flicker frequency of 50 Hz; blue dashed lines: represent the constraint of a charge density limit of 0.35 mC/cm². Light shading: pulse widths and frequencies that fall outside these constraints. The z-axis represents current (A), charge density (B), and overall charge across the entire pulse train (C). Given these example constraints, the model predicts that the most charge efficient stimulation pattern is a 50-Hz pulse train consisting of 0.089-ms pulses (C, *).

detect flashes of 100 photons⁴⁰; a very small increase in the firing rate of ganglion cells is sufficient to mediate detection. It is possible that thresholds in our subjects are mediated by a relatively small number of spikes. These spikes might, of course, occur either in a single cell or across several cells. At suprathreshold our model presumably approximates the population response of a larger number of cells each producing one or multiple spikes.

Quantitative models, as well as providing a description of behavioral performance, can also sometimes provide insight into neural organization.⁵¹–⁵⁴ In the case of retinal stimulation, neurophysiological interpretations are necessarily highly speculative, given the current body of knowledge describing membrane and synapse properties of degenerate retina and the scarcity of neurophysiological data regarding the effects of retinal stimulation.

Histologic evidence shows that bipolar cells of degenerated retinas have a general decrease and redistribution of glutamate receptors, leading to a decrease of mGluR6- and iGluR-mediated currents.⁵⁵,⁵⁶ However, retinal ganglion cells continue to show iGluR currents.⁵⁶ The ganglion cells of retinal degenerated retinas have a much higher spontaneous firing frequency than wild-type, which sometimes produces rhythmic bursts of activity as a result of self-signaling.⁵⁶,⁵⁷ However, expressing and activating channel rhodopsin (ChR2), a light-activated cation channel, in ON bipolar cells of rd1 mice results in excitatory responses in retinal ganglion cells of rd1 retinas similar to those of wild-type.⁵³ These findings suggest that the circuitry between bipolar and ganglion cells is not entirely disrupted and that this circuitry, when presented with stimuli, has some functional similarities to normal retina. Given this, we find it interesting that many of the stages in our model are quite similar to those used to describe in vitro models of the normal retina.

The parameter \( \tau_1 \) (equation 1) represents the time course of the first stage of current integration and was mainly constrained by the shape of the functions for single-pulse thresholds (Fig. 3B). Estimates varied between 0.24 and 0.65 ms, with a mean of 0.42 ms—a value very similar to electrophysiology estimates of the integration of current by ganglion cells in rodent and primate models.²⁶,³⁸,⁴⁹ It should be noted that intracellular current injection results in slower time constants, ranging from 3.9 to 94.6 ms.⁵⁸ Lipton and Tauck⁵⁹ have found the time constant of sodium channel activation in ganglion cells is approximately 0.1 ms; our \( \tau_1 \) parameter is therefore closer to the time constant of sodium channel activation than the time constant of the entire membrane.⁵⁹ In contrast, the time constant associated with spikes originating from mammalian bipolar cells seems to be much longer. Long-latency spiking in ganglion cells, occurring >8 to 60 ms after the beginning of electrical stimulation,⁶⁶,⁶⁸,⁶⁹ is thought to originate from bipolar cells, since it is suppressed by a cocktail of synaptic blockers. The measured time constant associated with these longer latency spikes varies between 20 and 26 ms, depending on electrode size (time constants for amphibian bipolar cells may be shorter⁶⁸). The time constant associated with the inhibitory input from amacrine cells is also much longer, on the order of 100 to 200 ms.⁶⁸ The fast integration time course of the first stage of our model therefore suggests that direct stimulation of ganglion cells may be the primary source of percepts in our subjects for the pulse durations and amplitudes that we used.

The parameters \( \epsilon \) and \( \tau_3 \) (equation 3) represent desensitization as a consequence of accumulated charge; \( \epsilon \) represents
the strength of desensitization and \( \tau_2 \) represents the time constant over which charge was integrated. These values were mainly determined by the difference in the data curve slopes between the fixed-duration 0.075- and 0.975-ms pulse trains for both threshold and suprathreshold data (Figs. 4, 6). Slopes were steeper for 0.075- than for 0.975-ms pulses, consistent with desensitization as a function of accumulated charge. \( \tau \) ranged from 2 to 3 with a mean of 2.25 for threshold stimulation, and between 8 and 10 with a mean of 8.75 for suprathreshold stimulation. Estimates of \( \tau_2 \) ranged between 58 and 57 ms with a mean of 45.25 ms. Our need to include desensitization as a function of charge to adequately fit our data is consistent with the finding that shifts in resting potentials can be produced in ganglion cells by injection of hyperpolarizing current.\(^{33} \) It is possible that these shifts in resting potential as a result of injection of hyperpolarizing current may be analogous to at least some of the processes that underlie slow-contrast gain control for light stimuli.\(^{51–55} \) Inhibition from presynaptic cells is also likely to play a role in the desensitization that we observed: inhibitory presynaptic influences on spiking in response to electrical stimulation have been described by Fried et al.,\(^{38} \) particularly for longer pulses. It seems likely that the desensitization stage of our model approximates a series of complex adaptive processes, with time courses varying between milliseconds to tens of seconds.\(^{31–55} \)

The parameter \( \beta \) (equation 4) describes a power input-output nonlinearity, which presumably describes the input-output nonlinearity across a population of cells. It was mainly constrained by the slopes of the threshold and suprathreshold pulse train stimuli (Figs. 4, 6). An increase in the brightness of the percept to be matched led to a decrease in the slope of the response nonlinearity. It varied between 3.0 and 4.2, with a mean of 3.43 for threshold stimulation and ranged between 0.6 and 1.0, with a mean of 0.85, for suprathreshold data. One possibility is that as the intensity of stimulation increases, neurons with shallower input-output nonlinearities are recruited. Alternatively, this change in the power function may be driven by changes in the input-output nonlinearity within individual cells. Although it is necessary to be cautious in generalizing from responses to light stimuli to the effects of electrical stimulation, similar nonlinearities, including changes in slope as a function of increased contrast or injected hyperpolarizing current,\(^{32,55} \) are found in linear-nonlinear models describing spiking behavior in ganglion cells\(^{31,55} \) and human behavioral data for light stimuli.\(^{34} \)

The parameter \( \tau_3 \) (equation 5) represents the integration period of the final low-pass filter. It was primarily determined by the shapes of the curves of Figure 5. Thresholds decrease as a function of frequency for a fixed number of pulses, with an asymptote at approximately 100 to 200 Hz. The parameter \( \tau_3 \) ranged between 24 and 53 ms, with a mean of 26.25 ms. This time constant is consistent with the slow temporal integration that occurs in cortex,\(^{54,61} \) though of course there are many stages (or combination of stages) along the visual pathway that might mediate this slow integration.

A successful retinal prosthesis must produce percepts consisting of regions of constant brightness across a range of brightness levels, while satisfying a complex set of engineering constraints: Charge densities must remain relatively low, in that it is technically difficult to produce very high-current amplitudes, and absolute charge must be minimized to maximize battery life. Models of the perceptual effects of electrical stimulation over time, such as that described herein, will be critical in allowing stimulation protocols to be selected that best satisfy these many constraints.

Of course, a wide variety of challenges remain. For example, apparent brightness is not the only perceptual quality that needs to be considered. It is likely that different temporal patterns stimulate slightly different subpopulations of neurons, resulting in distinct percepts. Another constraint is that stimulation should ideally be at a rate that is high enough to avoid perceptual flicker. Moreover, our experiments only considered pulse trains of a few seconds. Longer periods of continuous stimulation (minutes or hours) may result in long-term adaptation, sensitization, and/or retinal rewiring.\(^{62} \)

Possible interactions with the extent of retinal degeneration are another important consideration, especially given that this model is based on only two subjects. It is not currently known whether implanting the electrode array shortly after the onset of blindness would improve performance or result in different temporal dynamics. Given the changes that occur within the diseased retina,\(^{7,52} \) timing of implantation may well be a significant factor. As more prostheses are implanted and tested, this will be an interesting question to address.

Finally, the model described here is limited in its scope. It simply predicts sensitivity over time at the single-electrode level. The extension of models such as ours to the spatial domain is an obvious next step. A successful prosthesis will require arrays that are stable on the retina, map to predictable locations in space, and are of high enough resolution to provide the quality of visual information needed to perform useful real-world tasks.

References
