# Brightness as a Function of Current Amplitude in Human Retinal Electrical Stimulation

*Scott H. Greenwald*,<sup>1,2</sup> *Alan Horsager*,<sup>3</sup> *Mark S. Humayun*,<sup>4</sup> *Robert J. Greenberg*,<sup>1</sup> *Matthew J. McMahon*,<sup>1</sup> *and Ione Fine*<sup>2</sup>

**PURPOSE.** With the goal of eventually restoring functional vision in patients with retinal degenerative diseases, USC/Second Sight Medical Products, Inc. chronically implanted blind human subjects with a prototype epiretinal prosthesis consisting of a  $4 \times 4$  array of 16 stimulating electrodes. To accurately represent a visual scene, a visual prosthesis must convey luminance information across a range of brightness levels. To achieve this, the brightness of phosphenes produced by an individual electrode should scale appropriately with luminance, and the same luminance should produce equivalently bright phosphenes across the entire electrode array. The goal was to examine how apparent brightness changes as a function of stimulation intensity across electrodes.

**METHODS.** As described in previous studies, electrical stimulation of intact cells of the neural retina using this prosthetic device reliably elicits visual percepts in human subjects blinded by retinitis pigmentosa. Here, apparent brightness for a range of electrical amplitudes was measured using both subjective magnitude rating and brightness-matching procedures in chronically implanted human subjects.

**R**ESULTS. It was found that apparent brightness can be described as a power function of stimulation intensity. The same model can also predict brightness matching across electrodes.

CONCLUSIONS. These results suggest that a relatively simple model for scaling current across electrodes may be capable of producing equivalently bright phosphenes across an entire array. (ClinicalTrials.gov number, NCT00279500.) (*Invest Ophthalmol Vis Sci.* 2009;50:5017-5025) DOI:10.1167/iovs.08-2897

**R** etinitis pigmentosa and age-related macular degeneration are two of the more frequent causes of blindness in the developed world.<sup>1-3</sup> Both diseases are progressive and begin

Supported by Second Sight Medical Products, Inc., Fletcher Jones Foundation, National Institutes of Health Grants NEI EY012893 and EY014645, and Research to Prevent Blindness.

Submitted for publication September 16, 2008; revised February 22 and May 18, 2009; accepted August 31, 2009.

Disclosure: S.H. Greenwald, Second Sight Medical Products, Inc. (I, E, P); A. Horsager, Second Sight Medical Products, Inc. (F, P, R); M.S. Humayun, Second Sight Medical Products, Inc. (F, I, C, P, R); R.J. Greenberg, Second Sight Medical Products, Inc. (E, P); M.J. McMahon, Second Sight Medical Products, Inc. (E, P); I. Fine, Second Sight Medical Products, Inc. (F, P)

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be marked "*advertise-ment*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Corresponding author: Ione Fine, Department of Psychology, University of Washington, Guthrie Hall, Room 233, Box 351525, Seattle, WA 98195-1525; ionefine@u.washington.edu.

with the degeneration of photoreceptors. In later stages of these diseases, bipolar, amacrine, and ganglion cells are still present, though their numbers are significantly decreased<sup>4-6</sup> and their spatial organization and circuitry are significantly disorganized.<sup>7,8</sup> There are more than 180 different gene mutations that result in photoreceptor diseases for which there is currently no cure or treatment.<sup>1</sup> Ideally, it would be possible to develop a treatment for these conditions that would not require targeting each genetic defect independently.

Several groups are developing implantable microelectronic visual prostheses that produce percepts by electrically stimulating remaining retinal neurons. To date, several groups have succeeded in generating visual percepts via electrical stimulation with implanted acute, semi-acute, and long-term retinal prostheses in human patients.<sup>9-14</sup> The ultimate goal of these projects is to generate useful vision in blind patients by transforming a video stream into a spatial and temporal sequence of electrical pulses that represents meaningful visual information. However, creating a perceptually meaningful pattern of stimulation is dependent on a detailed understanding of the perceived intensity of any given stimulation pattern; to date, the literature examining the perceptual consequences of electrical stimulation remains relatively sparse.<sup>10-12,15-20</sup>

A visual prosthesis should produce regions of constant brightness across a range of brightness levels, and ideally these brightness levels should be consistent with the apparent brightness of objects as they appear to those with normal vision. Our goal was to examine how apparent brightness changes as a function of stimulation intensity in two blind human subjects chronically implanted with a prototype epiretinal prosthesis consisting of a  $4 \times 4$  array of 16 stimulating electrodes.

In experiment 1, subjects rated the apparent brightness of pulse stimuli on individual electrodes using a reference pulse of fixed amplitude. We found that apparent brightness as a function of current amplitude can be described using a simple power function.

In experiment 2, a brightness-matching technique was used to compare apparent brightness across pairs of electrodes. We found that the apparent brightness of a given electrode can be related to other electrodes on the array using the same simple power function model.

The results from these two experiments suggest that a relatively simple model for scaling current across electrodes may, to a first approximation, be capable of producing equivalently bright phosphenes across an entire array.

# **GENERAL METHODS**

#### Subjects

Here we describe data from two subjects, a subset of six subjects who have undergone implantation since February 2002. The other four subjects were excluded for a variety of reasons: one subject because of

From <sup>1</sup>Second Sight Medical Products, Inc., Sylmar, California; <sup>2</sup>Department of Psychology, University of Washington, Seattle, Washington; and <sup>3</sup>Zilkha Neurogenetic Institute and <sup>4</sup>Doheny Eye Institute, University of Southern California, Los Angeles, California.

Investigative Ophthalmology & Visual Science, November 2009, Vol. 50, No. 11 Copyright © Association for Research in Vision and Ophthalmology



**FIGURE 1.** (A) Electrode array. The electrode array consisted of 260- or 520- $\mu$ m electrodes arranged in a checkerboard pattern with center-tocenter separation of 800  $\mu$ m. The entire array covered approximately 2.84 × 2.84 mm retinal space, subtending approximately 9.5° of visual angle. (B) Prosthesis system schematic. The stimulus sets were programmed using technical computing software on a personal computer, which communicated the stimulus parameters to an external visual processing unit (not shown). Signal and power information was then passed through an external inductive coupling device (not shown) that attached magnetically to 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 the power and signal information could be independently controlled for each electrode. (C) Schematic of the single pulse that was used in experiment 1. Stimulation consisted of a biphasic cathodic-first charge-balanced square wave pulse. The durations of the individual cathodic and anodic phases were 0.975 ms, and each phase was separated by a 0.975-ms interpulse interval. Anodic and cathodic phases were always matched in amplitude.

geographic location, two because of unrelated medical conditions, and one because the device was explanted before this experiment (the array cable became exposed, and, because the cardiac status of this patient precluded general anesthesia, the multiwire cable connecting the array to the external stimulator was cut and the intraocular portion of the array was left in place). This study protocol was granted an Investigational Device Exemption by the US Food and Drug Administration and was approved by the Institutional Review Board at the University of Southern California. This research adhered to the tenets of the Declaration of Helsinki.

#### The Retinal Prosthesis

The implant consisted of an intraocular unit containing 16 platinum stimulating electrodes arranged in a  $4 \times 4$  pattern within a silicone substrate (Fig. 1A) and an extraocular unit containing electronics for wireless data and power transmission (Fig. 1B). In all subjects, this epiretinally implanted array was positioned over the macular region and held in place by a small tack. For these two subjects, the array electrodes were 260 and 520  $\mu$ m in diameter and were arranged in an alternating checkerboard pattern. The center-tocenter separation between electrodes was approximately 800 µm. The electrodes therefore covered either a  $2.84 \times 2.84$ -mm area of retinal surface, corresponding to a visual angle of approximately 9.5°. The distant return electrode was placed on the electronics case. Several loops of excess wire were left in the orbit so that the eye could move freely without stressing the cable. A 16-wire SC cable connected the intraocular electrode array to the extraocular unit by traversing the sclera.

As described elsewhere, the extraocular component of the implant, which converts a radio frequency signal into electrical stimulation patterns, was surgically implanted in the temporal bone, similarly to a cochlear implant.<sup>21</sup> Implant data and power were transmitted via an inductive wireless link using an external antenna magnetically aligned over the electronic implant. The desired pulse pattern was sent to a custom-built video processing unit (Second Sight Medical Products, Inc., Sylmar, CA) that coded the data as a serial data stream and transmitted it to the implant via the wireless link. The stimulator was connected to the array by a multiwire cable and was controlled by a computer-based external system that allowed independent control over each electrode. A reverse telemetry function in the implant allowed direct measurement of the impedance of each electrode.<sup>17</sup> Each subject's nonsurgical eye was patched during all experiments to ensure that the results were not affected by any residual vision in that eye.

# EXPERIMENT 1: BRIGHTNESS RATING FOR SINGLE ELECTRODES

# Methods

**Stimuli.** Stimulation for test and reference pulses always consisted of a single biphasic, cathodic-first, charge-balanced square wave pulse with a pulse duration of 0.975 ms and a 0.975-ms interpulse interval (see Fig. 1C). For safety reasons, all pulse trains were charge-balanced using anodic pulses of equal width and amplitude. The reference pulse was fixed at a current amplitude chosen to be roughly 2.5 times the threshold amplitude for a single pulse on that electrode. Charge densities were always below 1 mC/cm<sup>2</sup>.

**Procedure.** We used a classic brightness-matching procedure based on that of Stevens.<sup>22</sup> Before beginning each testing session, subjects were repeatedly stimulated with the reference pulse and were told, "This reference pulse has brightness of 10, and we will present it to you before we begin each trial. Your task is to compare the brightness of the test pulse in each trial to the brightness of this reference pulse. If the test pulse seems to be twice as bright as the reference pulse, then give it a rating of 20. If the test pulse seems to be half as bright as the reference pulse, then give it a rating of 5."

Once the subject reported feeling confident about having a clear idea of the brightness of the reference pulse, we began the experiment. All subject ratings were provided verbally.

On each trial, subjects were first presented with the reference pulse and were reminded that this pulse should be considered as having a brightness of 10. This reference pulse was quickly ( $\sim$ 1 second) followed by the test pulse. Short auditory cues marked the onset of the presentation of both the reference pulse and the test pulse. Subjects were then asked to verbally rate the apparent brightness of the test pulse compared with the reference pulse.

The test pulse was always presented on the same electrode as the reference pulse and had a current amplitude that varied pseudorandomly from trial to trial using the method of constant stimuli. Subjects were not told which test pulse current value had been presented on each trial, and no feedback was provided. Each test current amplitude was presented four times, and we calculated the mean and the SE of brightness ratings for each stimulation amplitude across these four repetitions.

#### Results

Subjects typically reported that phosphenes appeared white or yellow and were round or oval. At suprathreshold, percepts

FIGURE 2. Brightness-rating judgments for both subjects. Four electrodes are shown for each subject. The x-axis represents the current amplitude of the test pulse. Solid symbols: mean brightness rating for that test amplitude. Asterisk: current amplitude of the reference pulse, which was defined as having an apparent brightness of 10. Black dotted lines: power fit with three free parameters. Black dashed lines: power fit with two free parameters. Gray solid line: linear fit with one free parameter. Black solid line: power fit with one free parameter. Single standard errors are shown.



were reported as brighter, and the shape occasionally became more complex than a simple round or oval shape. Shapes were reported as approximately 0.5 to 2 inches in diameter at arm's length, corresponding to roughly 2° to 3° of visual angle. For both subjects, we also carried out a version of this experiment (data not shown) in which subjects were asked to rate both brightness and size on each trial. Subjects reported no difficulty or confusion when performing the two separate tasks, and brightness ratings were consistent with those obtained when performing the brightness task in isolation. For S1, size ratings were strongly correlated with brightness judgments, but for S2, the slope relating current amplitude to apparent size was much flatter than the slope obtained for brightness judgments. This variability in apparent size as a function of stimulation current between our two subjects was curious, but with only two subjects it was impossible to say whether this was attributed to differences in implantation between the two subjects (they were implanted with identical arrays), individual differences in the evoked percepts that were induced, or differences in how they interpreted the size task (for example, reporting the size of the brightest portion of the evoked phosphene or the composite phosphene).

Figure 2 shows brightness rating judgments for four electrodes for each of the two subjects on linear axes. (Note that rating data such as these are often presented on logarithmic axes: one reason for this is that in log-log axes a power function becomes a straight line whose slope is proportional to the exponent.) On each curve, the solid symbols represent mean subject ratings for each test electrode current amplitude, and the star represents the reference pulse (on the same electrode). Each data point represent four rating judgments. Two data sets in a single subplot represent repeated measurements on that electrode. Generally, these repeated sessions were separated by 21 days. In one case, for subject S1 (C3), we repeated the same set of measurements on the same electrode twice in a single session. In the case of S1, rating judgments seemed to vary significantly across repeated measurements, even when these two sets of measurements were carried out on the same day. Possible explanations for these inconsistencies include changes in the rating scale for S1 (both within and across sessions), adaptation effects within a single session, and movement of the electrode array across sessions. It should be noted that despite these differences across sessions, S1 did reliably report the standard as having a brightness of 10 and reliably (with relatively small standard errors) reported increasing brightness with increasing current intensity, suggesting that she understood the task. S2 showed remarkable test-retest reliability even across separate sessions.

The curves through the data represent four model fits in descending order of complexity. The black dotted lines represent the best-fitting power function with a multiplicative scalar and additive intercept  $(B = aC^b + d)$ , where B is the brightness rating made by the subject and C is the current amplitude of the test electrode. The black dashed lines represent the best-fitting power function ( $B = aC^{b}$ ) for each electrode with the assumption of an intercept of zero. In other words, the subject would report a brightness of zero when there was no stimulation, and, equally, the subject did not have a hard threshold whereby very low current amplitude values led to a brightness rating of zero. The gray solid lines represent the best-fitting linear function (B = aC), where again we assumed a zero intercept. The black solid lines once again represent the best-fitting power function without intercept  $(B = aC^b)$ ; however, in this model, b was fixed to be the median of the best-fitting values of b across all four electrodes for that subject. Parameter values a, b, and d were determined using a standard least-squared errors minimization technique.

It can be seen that these fits were very similar; in most cases, the separate curves were almost completely overlapping. Best-fitting parameter values and percentages of variance accounted for are shown in Table 1.

 TABLE 1. Best Fitting Parameter Values and Percentage of Variance Accounted for Using Four Different Models to Describe Brightness

 Rating Data

	Electrode	Power Fit with Intercept				Power Fit without Intercept				Linear Fit			Power Fit with Fixed b		
Subject		а	b	d	$R^2$	а	b	d	$R^2$	а	b	$R^2$	а	b	$R^2$
S1	D4	0.43	0.69	0.01	0.92	0.43	0.69	0.00	0.92	0.09	1.00	0.83	0.14	0.91	0.87
S1	D4	0.13	1.00	0.01	0.97	0.13	1.00	0.00	0.97	0.13	1.00	0.97	0.22	0.91	0.97
<b>S1</b>	B2	0.19	0.91	0.01	0.95	0.20	0.91	0.00	0.95	0.12	1.00	0.94	0.19	0.91	0.95
S1	B2	0.31	0.85	0.00	0.96	0.31	0.85	0.00	0.96	0.14	1.00	0.94	0.23	0.91	0.95
S1	C4	0.09	1.07	0.01	0.97	0.09	1.07	0.00	0.97	0.14	1.00	0.96	0.22	0.91	0.95
S1	C3	0.12	0.98	0.00	0.98	0.12	0.98	0.00	0.98	0.11	1.00	0.98	0.18	0.91	0.98
S1	C3	4.30	0.37	-10.80	0.96	0.54	0.67	0.00	0.94	0.08	1.00	0.85	0.14	0.91	0.89
Average		0.80	0.92	0.01	0.96	0.21	0.92	0.00	0.96	0.12	1.00	0.94	0.20	0.91	0.95
S2	D4	2.42	0.35	-0.01	0.97	2.41	0.35	0.00	0.97	0.08	1.00	-0.80	1.37	0.46	0.89
S2	D4	2.83	0.31	0.00	0.96	2.83	0.31	0.00	0.96	0.06	1.00	-1.39	1.27	0.46	0.78
S2	B2	0.94	0.49	0.00	0.97	0.94	0.49	0.00	0.97	0.06	1.00	0.26	1.09	0.46	0.97
S2	B2	1.12	0.46	0.00	0.95	1.11	0.46	0.00	0.95	0.06	1.00	0.25	1.12	0.46	0.95
S2	C4	0.62	0.58	0.00	0.97	0.62	0.58	0.00	0.97	0.07	1.00	0.64	1.16	0.46	0.94
S2	C4	2.28	0.33	0.00	0.93	2.29	0.33	0.00	0.93	0.06	1.00	-1.05	1.17	0.46	0.83
S2	C2	30.65	0.11	-40.60	0.94	1.05	0.49	0.00	0.89	0.06	1.00	0.37	1.22	0.46	0.88
Average		5.84	0.38	-5.80	0.96	1.61	0.43	0.00	0.95	0.06	1.00	-0.25	1.20	0.46	0.92

For subject S1, there was little difference in the percentage of variance accounted for  $(R^2)$  between any of the models. For subject S2, the fits for the power functions accounted for significantly more variance than the linear fit, but there was little difference in the percentage of variance accounted for between any of the power functions.

These data could probably also be fit well using other nonlinear functions: the power of our data was not sufficient to differentiate power fits from other similar models. One advantage of modeling our data using a power function with zero intercept is that the exponent b, which describes the shape of the curve relating apparent brightness to current amplitude, is then independent of the scaling parameter a. As described, the amplitude of the reference pulse was chosen relatively arbitrarily to be approximately  $2.5 \times$  threshold. The independence of a and b allows the scaling factor a to compensate for the fact that the amplitude of our reference pulse was chosen relatively arbitrarily, leaving the parameter b to describe the nonlinearity in the amplitude-brightness function. If we had, for example, chosen the reference pulse to represent an apparent brightness of 5 rather than 10, then subjects' ratings of the brightness of the test pulses (which were rated relative to the reference pulse) would have been halved. This would result in a being halved, but b would remain unchanged. We found that for S1, b varied between 0.69 and 1.07 with a mean value of 0.92 and a median value of 0.91. For S2, b varied between 0.31 and 0.58 with a mean value of 0.43 and a median value of 0.46.

A two-factor (subject  $\times$  electrode size) ANOVA comparing the parameter *b* used to described brightness-rating slopes (for the model with zero intercept) only found a significant effect of subject (*P* < 0.0001), with no effect of electrode size (*P* > 0.05) and no significant interaction between subject and electrode size.

# **EXPERIMENT 2: BRIGHTNESS MATCHING**

Although brightness ratings have the advantage of providing insight into the apparent brightness of the percept, these ratings are inherently subjective. As a result, brightness ratings are likely to vary substantially across sessions and across subjects. As shown in experiment 1, we see significant variance in rating judgments even within a single session in a single subject. It is also likely that rating judgments may be difficult for elderly patients showing some memory or cognitive dysfunction.<sup>23,24</sup> Given that retinal implants target diseases such as retinitis pigmentosa and macular degeneration, which primarily affect older populations, methods of calibrating the brightness of electrodes that are robust against cognitive and memory decline have a distinct advantage.

In experiment 2, we used a brightness-matching technique in which subjects were asked to report which of two phosphenes appeared brighter. As well as comparing brightnessmatching performance within a single electrode, we also compared brightness across pairs of electrodes.

#### Methods

**Stimuli.** Both reference and test stimuli consisted of 200-ms, 15-Hz pulse trains that contained biphasic, cathodic-first, charge-balanced square wave pulses (Fig. 3A). The durations of the individual cathodic and anodic phases were 0.975 ms, and each phase was separated by a 0.975-ms interpulse interval, as in experiment 1. Our motivation for using pulse trains was that prosthetic stimulation is likely to use pulse trains rather than discrete pulses. Fifteen hertz was below the critical flicker fusion limit of our two subjects, who reported perceivable flickering in stimuli of 20 Hz and greater. The appearance of our stimuli was of a single stimulus (which might contain multiple phosphenes) with a moderate amount of flicker that did not interfere with task performance; 15 Hz was chosen as a relatively low frequency (desirable because of system-specific engineering constraints) that produced only a moderate sensation of flicker.

Because we were using pulse trains rather than single pulses (as in experiment 1) we used a smaller range of current amplitudes in the brightness-matching experiment to remain below the conservative long-term charge density limit of  $0.35 \text{ mC/cm}^2$ .

**Procedure.** For each subject, the same reference electrode was used throughout the experiment. The reference electrode for S1 was C3, and the reference electrode for S2 was C2. The brightness of the reference electrode was compared with that of six test electrodes for each subject.

The choice of the reference electrode was based on two criteria: the electrode threshold was chosen to be close to the median threshold value across all 16 electrodes, and the electrode was chosen to be within the innermost  $2 \times 2$  square of the array.

Subjects made brightness judgments between a pulse train presented on the reference electrode and a pulse train presented on the



**FIGURE 3.** (A) Schematic of the pulse train that was used in experiment 2. Stimulation consisted of a biphasic cathodic-first, charge-balanced square-wave pulse train. The durations of the individual cathodic and anodic phases were 0.975 ms, and each phase was separated by a 0.975-ms interpulse interval. Anodic and cathodic phases were always matched in amplitude. Pulses were presented at a rate of 15 Hz for a duration of 200 ms. (B) Example data showing how PSE (*dashed lines*) was calculated. The PSE was defined as the current amplitude on the test electrode where subjects reported that the test electrode was brighter on 50% of trials. The size of each marker was proportional to the number of trials collected for that test current amplitude. Data shown here are for S1, in the condition where the reference electrode (C3) was compared with electrode B4. The reference pulse was fixed at 41.1  $\mu$ A.

test electrode using a two-interval, forced-choice procedure in which the subject reported which of two intervals appeared brighter on each trial. Each interval lasted 200 ms; there was a 900-ms delay between intervals (we used a slightly longer delay between intervals than is generally used in vision experiments chosen to minimize adaptation effects<sup>19</sup>), and presentation order for the reference and test electrode pulses was randomized for each trial.

Within each run (100 trials), the current amplitude of the reference pulse train was held constant, and the current amplitude of the test pulse train was adjusted using a 1 up-1 down staircase procedure based on which stimulus the subject had reported as brighter in the previous trial. To make comparisons across a range of brightness levels, runs were carried out with the reference pulse train taking five different brightness levels spanning 23.3  $\mu$ A to 119.7  $\mu$ A. For each current value on the reference electrode, as the current amplitude of the test pulse train increased, so did the probability of the subject reporting that the test pulse train was brighter.

We used a Monte Carlo simulation procedure to find the best-fitting cumulative normal function to find the error value of the estimated point of subjective equality (PSE; the amplitude of the test pulse train for which the test and reference appear equally bright).

#### Results

Figure 3B shows an example of a single data run in which the test electrode B4 was brightness matched to the reference electrode C3 (41.1- $\mu$ A, 15-Hz pulse train). The *x*-axis represents the current amplitude of the test pulse train, and the *y*-axis represents the probability of the subject reporting that the test pulse train was brighter than the reference pulse train. We used a Monte Carlo simulation procedure<sup>25,26</sup> to find the best-fitting cumulative normal function, which was used to calculate the estimated point of subjective equality. The data run shown here contained 100 trials; the size of each data point is proportional to the number of trials at that test pulse train amplitude.

Figure 4 shows brightness matching between the reference and test electrodes for both subjects. Each point represents the amplitude on the test electrode required to reach the PSE for a fixed current amplitude on the reference electrode. The *x*-axis represents the current amplitude of the reference electrode, and the *y*-axis represents the current amplitude of the test electrode.

As would be expected, when the reference electrode was brightness matched to itself, the curve of subjective equal brightness fell along a line of unity slope. In experiment 1, we found that brightness as a function of current could be described using Stevens' power function ( $B = aC^b$ ). The natural extension of Stevens' model to this brightness comparison data was made by assuming that the brightness of both the test and the reference electrode could be described using Stevens' power function, i.e., with the equations  $B_T = a_T C_T^{b_T}$  and  $B_R = a_R C_R^{b_R}$  for the test and reference electrode, respectively. By rearrangement, the current on the test electrode ( $C_T$ ) needed to match the brightness of the reference electrode at a given current amplitude ( $C_R$ ) can be described as  $C_T = a_T/a_R C_R^{(b_R/b_T)}$ . The dotted lines show a fit to our brightness matching data using the model  $C_T = aC_R^{b}$ , where both *a* and *b* were allowed to vary (*a* represents  $a_T/a_R$ , and *b* represents

BS2

**FIGURE 4.** Brightness-matching data for both subjects. For each subject, we measured PSEs with pulse amplitude on the reference electrode fixed at five different amplitude levels. The *x*-axis represents the amplitude of the pulse on the reference electrode. The *y*-axis represents the PSE on each of six test electrodes, and the reference electrode brightness matched to itself. *Large symbols*: 520-µm electrodes; *small symbols*: 520-µm electrodes; *Dasbed line*: equal amplitude on test and reference electrode.





AS1

			Linear Fit				
Subject	Electrode	a	b	$R^2$	a	<b>R</b> <sup>2</sup>	
<u>\$1</u>	B2	0.07	1.59	1.00	0.99	0.91	
S1	BÁ	0.31	1.25	0.96	0.98	0.94	
S1	A4	3.04	0.65	0.95	0.63	0.77	
S1	C4	1.37	0.81	1.00	0.60	0.96	
S1	C3	0.91	1.02	1.00	0.98	1.00	
S1	C1	0.29	1.25	0.99	0.90	0.96	
Average		1.00	1.10	0.98	0.85	0.92	
82	B2	2.88	0.83	0.99	1.35	0.95	
S2	B3	0.99	1.06	0.91	1.29	0.91	
S2	A4	1.27	0.94	0.86	0.96	0.86	
S2	D4	0.45	1.10	0.95	0.72	0.94	
S2	D2	2.15	0.74	0.94	0.66	0.84	
S2	C2	1.12	0.98	1.00	1.02	1.00	
Average		1.48	0.94	0.94	1.00	0.92	

 TABLE 2. Best Fitting Parameter Values and Percentage of Variance Accounted for Using Two Different Models to Describe Brightness

 Matching Data

 $b_R/b_T$ ). The solid lines represent a fit based on the model  $C_T = aC_R$ . The assumption that b = 1 is, of course, equivalent to  $b_R = b_T$  (the same power function exponent is applied to both electrodes being compared).

Best-fitting parameter values and percentage of variance accounted for with these two different fits are shown in Table 2. The linear model where b=1 again accounted for a large amount of the variance, suggesting that reasonable brightness matching can be carried out across electrodes by measuring a single parameter, a.

The value of a has a simple intuitive interpretation: a slope of 0.9 means that for any current amplitude on the reference electrode, the test electrode only required 90% as much current as the reference electrode to appear equally bright. We found that slopes varied between 0.65 to 1.59 for S1 and between 0.74 to 1.10 for S2 (i.e., the least sensitive electrode required 145% more current than the most sensitive electrode to create an appearance of equal brightness for S1 and required 49% more current for S2).

We found that the slopes of these brightness-matching functions was significantly shallower for large (520  $\mu$ m) than for small (260  $\mu$ m) electrodes, implying that less current was required on a large electrode to match the brightness of a small electrode. A two-factor (subject  $\times$  electrode size) ANOVA on brightness-matching slopes found a significant effect of both subject (P < 0.05) and electrode size (P < 0.05), with no significant interaction between factors. As described, no effect of electrode size was noted in the brightness-rating task. However, fewer electrodes were tested in that task, and there was likely to have been greater variability in those data, as described below. In an earlier paper,<sup>20</sup> we reported that thresholds were the same for 260- and 520-µm electrodes. However, as can be seen in Figure 4, differences in sensitivity across different sizes of electrode at near threshold levels of stimulation were much smaller than at high current levels. This may also explain why we did not find a significant correlation between slope and threshold (P > 0.05).

# COMPARISON OF MEASUREMENT VARIABILITY BETWEEN BRIGHTNESS RATING AND MATCHING MEASUREMENTS

The black solid line in Figure 5A shows a single brightness rating function (replotted from Fig. 2) for electrode S2 D4, with gray triangles and lines representing  $\pm 1$  SD in the bright-

ness ratings. To compare variability in brightness-rating judgments to brightness-matching judgments, it was necessary to convert variability in rating scores to variability in current amplitude. To do this, we interpolated to find the points along the curves representing  $\pm 1$  SD along the *y*-axis and found the corresponding amplitude values along the *x*-axis. If subjects were exquisitely accurate in assigning brightness ratings to changes in amplitude, we would expect this uncertainty range to be very small. If subjects showed large amounts of variability in assigning brightness ratings, then this uncertainty range would have been large.

The black solid line in Figure 5B shows a single brightness matching function (re-plotted from Fig. 4) for the same electrode. Gray triangles and lines represent  $\pm 1$  SD in the brightness match. We once again interpolated to find the uncertainty range for the test electrode. If subjects were exquisitely sensitive to changes in amplitude, we expected the current range over which subjects were uncertain whether the test or the reference pulse was brighter to be very small. If subjects were insensitive to changes in brightness, this uncertainty range would have been large.

Figure 5C plots uncertainty range as a function of current amplitude (rating, thin lines) and current amplitude of the reference electrode (matching, thick lines) for both subjects. Only those electrodes for which both types of measurement were taken were included. Note that the brightness rating task used single pulses, whereas the brightness matching task used pulse trains. However, data (not shown) comparing brightness matches for individual pulses to pulse trains on the same electrode did not show any difference in measurement variability between the two types of stimulation. It can be seen that for any given electrode, subjects were able to make finer brightness discriminations when using the matching protocol than they were when using the rating protocol. Using the rating protocol, the size of subjects' standard deviations suggested that they would be able to reliably (with 64% accuracy) differentiate between less than four brightness levels, whereas the standard deviations from our matching protocol suggested that subjects might be able to differentiate up to 16 different brightness levels.

# DISCUSSION

We show here that apparent brightness increases systematically as a function of current amplitude in subjects who undergo implantation with an epiretinal electrode array.



FIGURE 5. Comparison of measurement variability between brightnessrating and brightness-matching measurements. (A) Brightness rating for a single electrode. (S2 D4) showing the current amplitudes span  $\pm 1$  SD (the uncertainty range). (B) The uncertainty range for brightness matching on the same electrode. (C) A comparison of the uncertainty range across the two measurement techniques for six electrodes. Thin lines: brightness rating. Thick lines: brightness matching. Each electrode is represented by a different color and line pattern.

# Subject Variability in the Rating Task

In the brightness rating experiment, S2 showed remarkable test-retest reliability even across separate sessions. For S1, rating judgments seemed to vary significantly across repeated measurements, even when both sessions were carried out on the same day. It should be noted that despite these differences across sessions, S1 did reliably report the standard as having a brightness of 10 and reliably (with relatively small standard errors) reported increasing brightness with increasing current intensity, suggesting that she understood the task.

It seems unlikely that this variability resulted from movement of the electrode array across sessions. Both subjects had arrays that were assessed as close to the retina throughout the data collection period (data were collected over 29 weeks in the case of S1 and 54 weeks in the case of S2), and there was no noticeable movement of the array over the course of data collection, as assessed by threshold measurements, fundus examination, and optical coherence tomography.<sup>20</sup> This explanation is also inconsistent with the finding that her rating measurements varied within a single session.

Another possibility was that these changes in apparent brightness were caused by adaptation effects within a session. However, in a previous study,<sup>19</sup> we systematically examined adaptation between two intervals of a single trial, between the beginning and end of a run (approximately 100 trials), and between the beginning and end of a session. No adaptation effects were found between the beginning and the end of a run or between the beginning and end of a 3-hour testing appointment, inconsistent with adaptation being responsible for the change in the brightness ratings of S1 across sessions. Small adaptation effects were found between two intervals of a single trial for a 1050-ms inter-interval delay (similar to the delay between the standard and test in our brightness-matching paradigm). However, this adaptation effect would be expected to result in a very slight and systematic increase in the slope relating apparent brightness as a function of current rather than variance across sessions.

We think it most likely that S1 changed her subjective rating scale across sessions. For example, elicited percepts were not uniformly bright, so her rating scale could have varied depending on whether she was rating the percept based on the average brightness or the brightest part of the percept. Subjects were not informed "how" to judge brightness for any individual electrode but were told to keep their criterion consistent within a given session.

## Shape of the Power Function for Rating Judgments

Brightness ratings can consistently be well fit by a power function with zero intercept, as is traditionally used for magnitude data. Nonlinearities in the amplitude-brightness function did vary across our two subjects, but within each subject good fits could be obtained using a fixed exponent (*b*) across all electrodes. Exponents for rating data ranged between 0.69 to 1.07 for S1 and 0.31 to 0.58 for S2.

Rating data for electrical stimulation of the skin and visual cortex are both better described with exponents greater than 1, implying an accelerating response at high-stimulation amplitudes.<sup>27</sup> However, rating data for most sensory stimuli (including our data) are best fit with exponents less than 1, implying response saturation at high intensities. Brightness-rating data for a light point source presented in darkness has an exponent of 0.5. For light stimuli, this exponent decreases as a function of the size of the light source: a 5° light stimulus has an exponent of 0.333.<sup>22,28</sup> One interesting possibility is that the difference in exponent between our two subjects might have corresponded to differences in the apparent size of their elicited percepts. Unfortunately, our inability to provide a fixed-size reference visual target to these blind subjects makes it technically challenging to compare apparent size across sub-

jects (McMahon MJ, et al. *IOVS* 2007;48:ARVO E-Abstract 4443).

Although underlying neuronal responses cannot be directly inferred from rating judgments, the wide range of power law exponents observed across different types of stimuli suggests that differences in transducer functions between sensory systems are critical in determining the shape of the rating function. This is supported by evidence showing that the power function evoked by stimulating the acoustic nerve electrically is very different from the function relating loudness to sound intensity<sup>29</sup> and by a variety of studies in the tactile domain suggesting that rating judgments may be linearly related to the underlying neuronal response.<sup>30</sup> This linear relationship does not imply that the intervening neural transforms are all linear but simply that the sum of their serial superpositions is linear.<sup>31</sup> One potential explanation for our finding that brightness-rating functions are considerably more linear for electrical stimulation than for light stimuli might be that early stages of adaptation within cones (and possibly further stages of adaptation within bipolar and amacrine cells) are bypassed during electrical stimulation, thereby reducing saturation effects for higher intensity stimuli. It would be interesting to see whether retinal spiking rates in animal models show analogous differences in the functions relating firing rates to light versus electrical stimuli.

Although all electrodes showed a monotonic increase in brightness as a function of current, different electrodes (even within a single observer) varied significantly in how brightness increased with current amplitude. As measured in the brightness-matching task, the least sensitive electrode required 145% more current than the most sensitive electrode to create an appearance of equal brightness for S1 and required 49% more current for S2.

As described in earlier work, thresholds (the current required to reliably detect whether stimulation has occurred) also vary widely across subjects and across electrodes.<sup>18,20</sup> Potential factors that may affect sensitivity to electrical current include the degree of retinal degeneration, time from onset of total or partial blindness, differences in degeneration,<sup>7,8</sup> differences in sensitivity to electrical current across each subject's retina, and differences in the distance of the array from the retina.<sup>20</sup> These factors are also likely to be associated with differences in apparent brightness as a function of stimulation amplitude. As more patients undergo implantation and better retinal imaging techniques are developed, it should be possible to gain a better understanding of the relative importance of these various factors. Another important issue that will have to be addressed once more subjects undergo implantation is how the apparent size of percepts varies as a function of stimulation amplitude. As described, our subjects differed in how apparent size varied with stimulation amplitude. Regardless of whether apparent size remains relatively constant (similar to S2) or increases (S1) as a function of stimulation amplitude, some way of modeling the effect of stimulation current on percept size will have to be developed so as to maximize the "naturalness" of elicited percepts. Consistently evoking phosphenes of regular shape by which the brightness can be varied but the apparent size remains constant might be ideal; however, this may not be achievable.

#### Effect of Electrode Size

One factor that did show a significant effect on the measured slope for the brightness-matching task was the size of the electrode, with larger electrodes showing *shallower* slopes in the brightness-matching task than small electrodes, implying that less current was required on a large electrode to match the brightness of a small electrode.

It should be noted that this result is the opposite of what one might expect from simpler models of current density across electrodes, which predict that (assuming threshold/ brightness is proportional to charge density) large electrodes would require *more* current to match the reference electrode than small electrodes. This result also contradicts a recent literature review by Sekirnjak et al.,<sup>32</sup> who found, across a wide range of in vitro and in vivo studies (using electrodes varying between 2 and 1800  $\mu$ m in diameter), that log thresholds increase linearly with log electrode area with a slope of 0.7.

Given the large electrode sizes used in this experiment, it is likely that current density concentrated in a "ring" around the electrode edges.<sup>33–36</sup> As a result, the peak charge density near the edge of the electrode was likely to be higher than the charge density that would be predicted based on calculations using the entire electrode's surface area. Smaller electrodes would be expected to be less susceptible to this "ringing" than larger electrodes, resulting in a more even current distribution across the electrode surface and a lower peak charge density. If brightness is related to peak charge density rather than the mean charge density, this might explain why larger electrodes required less current than smaller electrodes to elicit a given brightness level.

It is also possible that the stimulation of a larger retinal area results in greater perceived brightness. This is the case for light stimuli: the detection threshold of small light stimuli is proportional to the area of that light stimulus multiplied by its luminance (Ricco's law). Ricco's law is thought to hold for stimuli of up to 0.5° in the parafovea,<sup>37,38</sup> where our electrode array was positioned. For much larger stimuli, the detection threshold is proportional to spot luminance. For stimuli of intermediate sizes, there appears to be a range of partial summation (Piper's law). Our 260- and 520- $\mu$ m electrodes subtended visual angles of 0.9° and 1.8° of visual angle on the retina, respectively. The stimuli generated by our electrodes might therefore easily have fallen within the range where Piper's law might be expected to hold.

It is also worth noting that differences in current distribution between large and small electrodes may have resulted in differential recruitment of neurons (both within a given cell class and across cell classes) between the two electrode sizes. Percepts generated by a low level of stimulation across a wide area compared with a smaller, more intense stimulation pattern might result in the same reported brightness but nonetheless might be based on very different populations of cells and be very different in appearance.

#### **CONCLUSIONS**

Given that the function relating current to brightness varies across electrodes, it is necessary to develop some mapping that will easily permit brightness to be equated across an entire array. As the numbers of electrodes in an array increase, it will be unfeasible to measure full brightness-rating or brightnessmatching functions for each individual electrode.

Here, we find that that brightness as a function of current amplitude can be described using a power function with a single scaling factor as a free parameter. Equivalently, it is possible to describe relative brightness across electrodes using a linear scaling with a single free parameter based on the slope of the brightness-matching function using a single electrode as a reference.

Of course, these results may not hold for other stimulation waveforms (though see Ref. 19), other prosthetic devices (such as those using much smaller electrodes or subretinal implantation), or in arrays that are less proximate to the retina (among other factors). However, these results, if they hold more generally, suggest that it should be possible to normalize brightness across an entire array of electrodes by measuring a single parameter for each electrode. A single brightness rating or matching judgment, if made at relatively high amplitude, would be sufficient to describe relative sensitivity across electrodes, and thereby provide a simple method of creating the appearance of equally bright phosphenes across an entire array across a wide range of brightness levels.

## Acknowledgments

The authors thank Anne-Marie Ripley and Grant Palmer for support on regulatory issues and for comments on the manuscript.

#### References

- Bunker CH, Berson EL, Bromley WC, Hayes RP, Roderick TH. Prevalence of retinitis pigmentosa in Maine. *Am J Ophthalmol.* 1984;97(3):357-365.
- Heckenlively JR, Boughman J, Friedman L. Diagnosis and classification of retinitis pigmentosa. In: Heckenlively JR, ed. *Retinitis Pigmentosa*. Philadelphia: JB Lippincott; 1988:21.
- Friedman DS, O'Colmain BJ, Munoz B, et al. Prevalence of agerelated macular degeneration in the United States. *Arch Ophthalmol.* 2004;122(4):564–572.
- Santos A, Humayun MS, de Juan E Jr, et al. Preservation of the inner retina in retinitis pigmentosa: a morphometric analysis. *Arch Ophthalmol.* 1997;115(4):511–515.
- 5. Humayun MS, Prince M, de Juan E Jr, et al. Morphometric analysis of the extramacular retina from postmortem eyes with retinitis pigmentosa. *Invest Ophthalmol Vis Sci.* 1999;40(1):143–148.
- Jones BW, Watt CB, Frederick JM, et al. Retinal remodeling triggered by photoreceptor degenerations. *J Comp Neurol.* 2003; 464(1):1-16.
- 7. Marc RE, Jones BW. Retinal remodeling in inherited photoreceptor degenerations. *Mol Neurobiol.* 2003;28(2):139–147.
- 8. Marc RE, Jones BW, Watt CB, Strettoi E. Neural remodeling in retinal degeneration. *Prog Retin Eye Res.* 2003;22(5):607-655.
- Humayun MS, de Juan E Jr, Weiland JD, et al. Pattern electrical stimulation of the human retina. *Vision Res.* 1999;39(15):2569– 2576.
- Rizzo JF 3rd, Wyatt J, Loewenstein J, Kelly S, Shire D. Perceptual efficacy of electrical stimulation of human retina with a microelectrode array during short-term surgical trials. *Invest Ophthalmol Vis Sci.* 2003;44(12):5362–5369.
- Rizzo JF 3rd, Wyatt J, Loewenstein J, Kelly S, Shire D. Methods and perceptual thresholds for short-term electrical stimulation of human retina with microelectrode arrays. *Invest Ophthalmol Vis Sci.* 2003;44(12):5355-5361.
- Yanai D, Weiland JD, Mahadevappa M, Greenberg RJ, Fine I, Humayun MS. Visual performance using a retinal prosthesis in three subjects with retinitis pigmentosa. *Am J Ophthalmol.* 2007; 143:820–827.
- Yanai D, Lakhanpal RR, Weiland JD, et al. The value of preoperative tests in the selection of blind patients for a permanent microelectronic implant. *Trans Am Ophthalmol Soc.* 2003;101:223– 228; discussion 228–230.
- Zrenner E, Stett A, Weiss S, et al. Can subretinal microphotodiodes successfully replace degenerated photoreceptors? *Vision Res.* 1999;39(15):2555-2567.
- Humayun MS, de Juan E Jr, Dagnelie G, Greenberg RJ, Propst RH, Phillips DH. Visual perception elicited by electrical stimulation of retina in blind humans. *Arch Ophthalmol.* 1996;114(1):40-46.
- Weiland JD, Humayun MS, Dagnelie G, de Juan E Jr, Greenberg RJ, Iliff NT. Understanding the origin of visual percepts elicited by electrical stimulation of the human retina. *Graefes Arch Clin Exp Ophtbalmol.* 1999;237(12):1007–1013.

- Humayun MS, Weiland JD, Fujii GY, et al. Visual perception in a blind subject with a chronic microelectronic retinal prosthesis. *Vision Res.* 2003;43(24):2573-2581.
- Mahadevappa M, Weiland JD, Yanai D, Fine I, Greenberg RJ, Humayun MS. Perceptual thresholds and electrode impedance in three retinal prosthesis subjects. *IEEE Trans Neural Syst Rebabil Eng.* 2005;13(2):201–206.
- Horsager A, Greenwald SH, Weiland JD, et al. Predicting visual sensitivity in retinal prosthesis patients. *Invest Ophthalmol Vis Sci.* 2009;50:1483–1491.
- 20. de Balthasar C, Patel S, Roy A, et al. Factors affecting perceptual thresholds in epiretinal prostheses. *Invest Ophthalmol Vis Sci.* 2008;49(6):2303-2314.
- Tucci DL, Niparko JK. Medical and surgical aspects of cochlear implantation. In: Niparko JK, ed. *Cochlear Implants: Principles and Practices*. Philadelphia: Lippincott Williams & Wilkins; 2000: 189-221.
- 22. Stevens SS. On the psychophysical law. *Psychol Rev.* 1957;64(3): 153-181.
- Gabrielsson A, Schenkman BN, Hagerman B. The effects of different frequency responses on sound quality judgments and speech intelligibility. J Speech Hear Res. 1988;31(2):166–177.
- Narendran MM, Humes LE. Reliability and validity of judgments of sound quality in elderly hearing aid wearers. *Ear Hear*. 2003;24(1): 4–11.
- Wichmann FA, Hill NJ. The psychometric function, II: bootstrapbased confidence intervals and sampling. *Percept Psychophys.* 2001;63(8):1314-1329.
- Wichmann FA, Hill NJ. The psychometric function, I: fitting, sampling, and goodness of fit. *Percept Psychophys.* 2001;63(8):1293–1313.
- Evans JR, Gordon J, Abramov I, Mladejovskyh MG, Dobelle WH. Brightness of phosphenes elicited by electrical stimulation of human visual cortex. *Sens Processes*. 1979;3(1):82–94.
- Stevens SS. Psychophysics: Introduction to Its Perceptual, Neural, and Social Prospects. New York: Wiley; 1975.
- Stevens SS. The psychophysics of sensory function. In: Rosenblith WA, ed. Sensory Communication. Cambridge, MA: MIT Press; 1961:1-34.
- Johnson KO, Hsaio SS, Yoshioka T. Neural coding and the basic law of psychophysics. *Neuroscientist.* 2002;8:111.
- Werner G, Mountcastle VB. Neural activity in mechanoreceptive cutaneous afferents: stimulus-response relations, Weber functions, and information transmission. J Neurophysiol. 1965;28:359–397.
- Sekirnjak C, Hottowy P, Sher A, Dabrowski W, Litke AM, Chichilnisky EJ. Electrical stimulation of mammalian retinal ganglion cells with multielectrode arrays. *J Neurophysiol.* 2006;95(6):3311– 3327.
- Wiley JD, Webster JG. Distributed equivalent-circuit models for circular dispersive electrodes. *IEEE Trans Biomed Eng.* 1982; 29(5):385-389.
- Wiley JD, Webster JG. Analysis and control of the current distribution under circular dispersive electrodes. *IEEE Trans Biomed Eng.* 1982;29(5):381–385.
- 35. Suesserman MF, Spelman FA, Rubinstein JT. In vitro measurement and characterization of current density profiles produced by nonrecessed, simple recessed, and radially varying recessed stimulating electrodes. *IEEE Trans Biomed Eng.* 1991;38(5):401–408.
- Rubinstein JT, Spelman FA, Soma M. Suesserman ME. Current density profiles of surface mounted and recessed electrodes for neural prostheses. *IEEE Trans Biomed Eng.* 1987;34(11):864– 875.
- 37. Shechter S, Hochstein S. Characteristics of the flux dimension cue in apparent motion correspondence. *Spat Vis.* 1992;6(1):11–24.
- Hallett PE, FH Marriott, Rodger FC. The relationship of visual threshold to retinal position and area. *J Physiol.* 1962;160:364– 373.