Four Theorems on the Psychometric Function

Keith A. May*, Joshua A. Solomon

Division of Optometry and Visual Science, City University London, London, United Kingdom

Abstract

In a 2-alternative forced-choice (2AFC) discrimination task, observers choose which of two stimuli has the higher value. The psychometric function for this task gives the probability of a correct response for a given stimulus difference, Δx. This paper proves four theorems about the psychometric function. Assuming the observer applies a transducer and adds noise, Theorem 1 derives a convenient general expression for the psychometric function. Discrimination data are often fitted with a Weibull function. Theorem 2 proves that the Weibull "slope" parameter, β, can be approximated by $\beta^{\text{Noise}} \times \beta^{\text{Transducer}}$, where β^{Noise} is the β of the Weibull function that fits best to the cumulative noise distribution, and $\beta^{\text{Transducer}}$ depends on the transducer. We derive general expressions for β^{Noise} and $\beta^{\text{Transducer}}$, from which we derive expressions for specific cases. One case that follows naturally from our general analysis is Pelli's finding that, when $d' \propto (\Delta x)^b$, $\beta \approx \beta^{\text{Noise}} \times b$. We also consider two limiting cases. Theorem 3 proves that, as sensitivity improves, 2AFC performance will usually approach that for a linear transducer, whatever the actual transducer; we show that this does not apply at signal levels where the transducer gradient is zero, which explains why it does not apply to contrast detection. Theorem 4 proves that, when the exponent of a power-function transducer approaches zero, 2AFC performance approaches that of a logarithmic transducer. We show that the power-function exponents of 0.4–0.5 fitted to suprathreshold contrast discrimination data are close enough to zero for the fitted psychometric function to be practically indistinguishable from that of a log transducer. Finally, Weibull β reflects the shape of the noise distribution, and we used our results to assess the recent claim that internal noise has higher kurtosis than a Gaussian. Our analysis of β for contrast discrimination suggests that, if internal noise is stimulus-independent, it has *lower* kurtosis than a Gaussian.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: keith@keithmay.org

Introduction

On each trial of a 2-alternative forced-choice (2AFC) discrimination task, observers are presented with two stimuli, one (often called the *pedestal*) with stimulus value $x = x_p$, and one (the *target*) with value $x = x_p + \Delta x$, where x represents a value along some stimulus dimension, such as contrast, luminance, frequency, sound intensity, etc., and Δx represents a (usually) positive increment in x. The observer has to say which stimulus contained the higher value, $x_p + \Delta x$. For this task, the function relating stimulus difference, Δx, to the probability of a correct response, P, is called the *psychometric function*. The form of the psychometric function can reveal characteristics of the underlying mechanisms, helping to constrain the set of possible models. In this paper we present four theorems that help us to understand the properties of the psychometric function and clarify the relationship between the psychometric function and the underlying model.

In order to fit the psychometric function to data, we need a mathematical function whose parameters can be adjusted to fit the kind of data set usually obtained. A widely used function is the Weibull function, and two of our theorems relate specifically to this function. Letting Ψ_{Weibull} represent the Weibull function, and letting P_{Weibull} represent its output (i.e., the predicted proportion correct), the Weibull function is given by

$$\begin{aligned} P_{\text{Weibull}} &= \Psi_{\text{Weibull}}(\Delta x) \\ &= 1 - 0.5 \exp\left[-(\Delta x/\alpha)^{\beta}\right]. \end{aligned} \quad (1)$$

α is the "threshold" parameter, the stimulus increment that gives rise to a proportion correct of $\Psi_{\text{Weibull}}(\alpha) = 1 - 0.5/e = 0.816....$ In what follows, we will frequently refer to this threshold performance level as P_{θ}, so this term should be read as the constant, $1 - 0.5/e = 0.816....$ β is often referred to as the "slope" parameter, because it is proportional to the gradient of the Weibull function at $\Delta x = \alpha$ when $\Psi_{\text{Weibull}}(\Delta x)$ is plotted on a log abscissa.

In 2AFC visual contrast discrimination experiments where the contrasts of both stimuli are at least as high as the detection threshold, β usually falls between 1 and 2, with a median of around 1.4 (see Table 1 and Figure 1). As the pedestal contrast approaches zero (making it a 2AFC contrast detection task), β increases to a value of around 3 [1–5].

When $\Psi_{\text{Weibull}}(\Delta x)$ is plotted on a log abscissa, changing the value of α shifts the function horizontally, but otherwise leaves it unchanged (Figure 2A), and changing the value of β linearly stretches or compresses the function horizontally, leading to a change of slope (Figure 2B). On this log abscissa, the Weibull function always has the same basic shape, up to a linear horizontal scaling. When $\Psi_{\text{Weibull}}(\Delta x)$ is plotted on a linear abscissa,
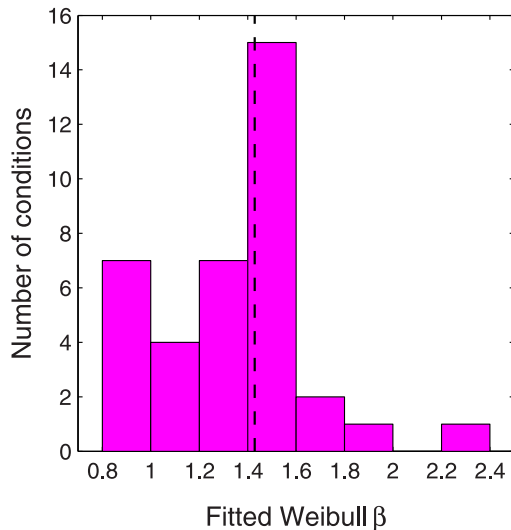
Figure 1. Distribution of fitted Weibull β values in Table 1. The fitted β values from the suprathreshold (non-starred) conditions of Table 1 were dropped into bins with edges that stepped from 0.8 to 2.4 in jumps of 0.2 (the histogram thus excludes one outlier, the value 6.70 for Henning et al. 's [47] observer GBH at 8.37 cpd). For this histogram, we used the β values that had been fitted using a nonzero lapse rate parameter where available, as this is more likely to reflect the true β. The median of this hybrid population (some including a lapse rate parameter, some not) was 1.43 (indicated by the vertical dashed line).
doi:10.1371/journal.pone.0074815.g001

changing the value of α linearly stretches or compresses the function horizontally as well as changing the threshold (Figure 2C), while changing the value of β changes the shape of the function in a way that cannot be described as a linear horizontal scaling (Figure 2D).

Since β is proportional to the slope of the Weibull function on a log abscissa, the low value of β for contrast discrimination (compared with detection) often leads to the psychometric function for discrimination being described as "shallow", and that for detection as "steep". However, psychometric functions for contrast discrimination can be *steeper* than for detection when plotted on a *linear* contrast abscissa (e.g., Nachmias & Sansbury [6], Figure 2; Foley & Legge [1], Figure 1). We must therefore be vigilant not to be misled by the common practice of referring to β as the "slope" parameter. β does control the slope of the Weibull function on a log abscissa, and this fact plays a key role in the proof of Theorem 2 of this paper, but the psychometric function is often plotted on a linear abscissa, and, in this case, α and β both affect the slope (Figures 2C and 2D); on a linear abscissa, α additionally controls the threshold and β additionally controls the overall shape of the psychometric function. Thus, when considering a linear abscissa, it would be more appropriate to describe β as the "shape" parameter, rather than the "slope" parameter.

The Weibull function defined in Equation (1) asymptotes to perfect performance ($P_{\text{Weibull}} = 1$). This is rarely achieved by human observers due to lapses of concentration, etc., and this can lead to a dramatic underestimation of β if the observer makes just one lapse on an easy trial [7]. Because of this problem, many researchers use a version of the Weibull function that includes a "lapse rate" parameter, λ:

$$P_{\text{Weibull}} = (1-\lambda) - (0.5-\lambda)\exp\left[-(\Delta x/\alpha)^\beta\right]. \quad (2)$$

This function asymptotes to $P_{\text{Weibull}} = (1-\lambda)$, and reduces to Equation (1) in the case of $\lambda = 0$. The psychometric function described by Equation (2) would result if the observer performed according to Equation (1) on a proportion $(1-2\lambda)$ of trials, and guessed randomly on the remaining trials.

The Weibull function was originally proposed by Weibull [8] as a useful, general-purpose statistical distribution. Its widespread use as a psychometric function can be traced back to Quick [9], who was apparently unaware of Weibull's prior work. Quick proposed this function because, given certain assumptions, the Weibull function makes it easy to calculate how detection performance will be affected by adding extra stimulus components or increasing the size or duration of the stimulus, an approach that has become known as *probability summation* [10–13]. Quick focused on yes/no detection tasks, where the observer has to make a binary decision about a single stimulus (as opposed to the 2AFC tasks that we consider in this paper, in which the observer makes a binary decision about a pair of stimuli), but a similar analysis can be applied to 2AFC tasks [2].

Most treatments of probability summation with the Weibull function invoke the "high threshold assumption" that a zero-contrast stimulus never elicits a response in the detection mechanism, so detection errors are always unlucky guesses. This assumption makes a number of predictions that have turned out to be false [2,14–16]. Furthermore, the convenient mathematics of probability summation with the Weibull function only applies to *detection*. For suprathreshold discrimination, where both stimuli are easily detectable, these computational benefits do not apply. Despite this, many researchers have continued to use the Weibull function to fit data from both detection and discrimination experiments for three perfectly valid reasons: it is well-known, fits well to the data, and is built into QUEST [17], probably the most widely used adaptive psychophysical method.

Different models of visual processing will deliver different mathematical forms for the psychometric function. Therefore, because of the widespread practice of fitting a Weibull function to data, it is of interest to know what happens when we fit a Weibull function to a psychometric function that is not a Weibull. In Theorem 2 of this paper, we derive a general analytical expression that gives a very accurate approximation of β when the Weibull function is fitted to non-Weibull psychometric functions.

Although the usage of the Weibull function has its origin in outdated theoretical views, the Weibull function has very recently become more relevant again, due to the work of Neri [18]. He argues that the internal noise on the decision variable has a Laplace distribution, which, as we explain later in this Introduction, can lead to a psychometric function that has the form of a Weibull function with $\beta = 1$.

First, we consider how the psychometric function might arise from the properties of the observer. In 2AFC discrimination experiments, the observer can be modelled using a transducer, followed by constant additive noise. The transducer converts the stimulus value, x, into some internal scalar signal value, $R(x)$. R is called the *transducer function*. A noise sample from a stationary, stimulus-invariant distribution is then added to the internal signal, $R(x)$, to give a noisy internal signal value. If the noise has zero mean, then $R(x)$ will be the mean internal signal for stimulus value x. The observer compares the noisy internal signal values from the two stimuli, and chooses the stimulus that gave the higher value.

From the experimenter's perspective, the observer behaves as if a sample of noise, ϵ, is added to the difference of mean signals, z, given by

Table 1. Fitted Weibull function parameters for 2AFC contrast discrimination.

Study	Condition/observer	Pedestal	$\lambda=0$			λ fitted			
			α	β	W	α	β	λ	W
Bird et al. [34]	CMB	0.03	0.00735	1.12	0.245	0.00735	1.12	5×10^{-13}	0.245
	CMB	0.3	0.0779	1.11	0.260	0.0692	1.21	0.0309	0.231
	GBH	0.03	0.00737	0.734	0.246	0.00643	0.832	0.0251	0.214
	GBH	0.3	0.0574	0.952	0.191	0.0541	0.993	0.0128	0.180
Foley & Legge [1]	JMF, 0.5 cpd	0.00400	0.00165	1.35	0.412	0.00165	1.35	3×10^{-9}	0.412
	JMF, 2 cpd	0.00230	0.00111	1.56	0.484	0.00111	1.56	5×10^{-12}	0.484
	JMF, 8 cpd	0.00300	0.00125	1.44	0.418	0.00123	1.46	0.0071	0.410
	GW, 0.5 cpd	0.00400	0.00134	1.50	0.335	0.00134	1.50	2×10^{-12}	0.335
	GW, 2 cpd	0.00229	0.000923	1.58	0.404	0.000923	1.58	1×10^{-12}	0.404
	GW, 8 cpd	0.00330	0.00117	1.40	0.353	0.000996	1.94	0.0544	0.301
Henning et al. [47]	CMB 2.09 cpd	0.15	0.0421	1.49	0.281	0.0421	1.49	1×10^{-12}	0.281
	CMB 8.37 cpd	0.15	0.0461	1.81	0.307	0.0379	2.21	0.0796	0.253
	GBH 2.09 cpd	0.15	0.0363	1.49	0.242	0.0363	1.49	2×10^{-12}	0.242
	GBH 8.37 cpd	0.15	0.0401	1.21	0.267	0.0244	6.70	0.0645	0.163
Henning & Wichman [40]	GBH*	0*	0.0219*	4.26*	—*				
	GBH*	0.01*	0.0102*	13.1*	1.02*				
	GBH*	0.02*	0.00562*	1.67*	0.281*				
	GBH	0.04	0.00705	0.987	0.176				
	GBH	0.08	0.0156	1.16	0.195				
	GBH	0.16	0.0322	1.75	0.201				
	GBH	0.32	0.0773	1.45	0.241				
	NAL*	0*	0.00619*	4.84*	—*				
	NAL*	0.00141*	0.00492*	5.90*	3.48*				
	NAL*	0.00283*	0.00407*	2.28*	1.44*				
	NAL*	0.00566*	0.00224*	1.43*	0.395*				
	NAL	0.0113	0.00272	0.902	0.241				
	NAL	0.0226	0.00707	0.990	0.312				
	NAL	0.0453	0.0150	0.943	0.331				
	NAL	0.0905	0.0233	1.28	0.257				
	NAL	0.181	0.0424	1.59	0.234				
	NAL	0.362	0.0658	1.33	0.182				
	TCC*	0*	0.00838*	6.38*	—*				
	TCC*	0.005*	0.00443*	2.14*	0.886*				

Table 1. Cont.

Study	Condition/observer	Pedestal	λ=0			λ fitted			
			α	β	W	α	β	λ	W
	TCC	0.01	0.00339	0.912	0.339				
	TCC	0.016	0.00787	1.17	0.492				
	TCC	0.032	0.0126	1.52	0.393				
	TCC	0.08	0.0301	1.64	0.377				
	TCC	0.16	0.0381	1.27	0.238				
	TCC	0.32	0.0686	1.10	0.214				
Meese et al. [4]	Pedestal −∞ dB*	0*	0.00855*	3.32*	−*				
	Pedestal −10 dB*	0.00316*	0.00557*	2.44*	1.76*				
	Pedestal −5 dB*	0.00562*	0.00346*	1.47*	0.615*				
	Pedestal 0 dB	0.01	0.00340	1.47	0.340				
	Pedestal 5 dB	0.0178	0.00654	1.48	0.368				
	Pedestal 10 dB	0.0316	0.0110	1.40	0.348				
	Pedestal 15 dB	0.0562	0.0176	1.58	0.313				
	Pedestal 20 dB	0.1	0.0233	1.47	0.233				
	Pedestal 25 dB	0.178	0.0339	1.47	0.191				
	Pedestal 30 dB	0.316	0.0536	1.36	0.170				
Nachmias & Sansbury [6]	CS	0.0079	0.00387	1.27	0.489				
Mean of suprathreshold (non-starred) conditions				1.32	0.298		1.82		0.297
Median of suprathreshold conditions				1.38	0.274		1.49		0.267

This table shows Weibull parameters fitted to 2AFC contrast discrimination data from six studies. The data from Meese et al. [4] are for their Binocular condition (plotted as squares in their Figure 5); these data were kindly provided by Tim Meese. For the other five papers, we read off the data points from digital scans of the figures (Bird et al. [34], Figure 1; Foley and Legge [11], Figure 1; Henning et al. [47], Figure 1; Henning & Wichmann [40], Figure 4; Nachmias & Sansbury [6], Figure 2). In most cases, these figures plotted the proportion correct, P, for several different contrast differences, Δc, and we fitted the Weibull function using a maximum-likelihood method; specifically, we fitted the Weibull function by maximizing the expression $\sum_{\Delta c} P \log[\Psi_{\text{Weibull}}(\Delta c)] + [1 - P]\log[1 - \Psi_{\text{Weibull}}(\Delta c)]$, where Ψ_{Weibull} is the Weibull function whose parameters were being fitted. In Henning & Wichmann's [40] paper, the figures plotted the Δc values corresponding to 60%, 75%, and 90% correct on the *fitted psychometric functions*, so we had to fit Weibull functions to points sampled from Henning & Wichmann's own fitted psychometric functions, rather than to the raw data. Where possible, we fitted both the lapse-free Weibull function of Equation (1), and the Weibull function of Equation (2), which includes a fitted lapse rate parameter, λ. Parameters for the former fit appear under the heading "$\lambda = 0$", and those for the latter appear under the heading "λ fitted". In many cases, the data did not sufficiently constrain λ because there were no data points on the saturating portion of the psychometric function; in addition, Meese et al.'s Weibull fits did not include a lapse rate parameter. The Weber fraction, W, is given by α / x_p, where x_p is the pedestal value. The cases where the pedestal is below detection threshold are starred in the 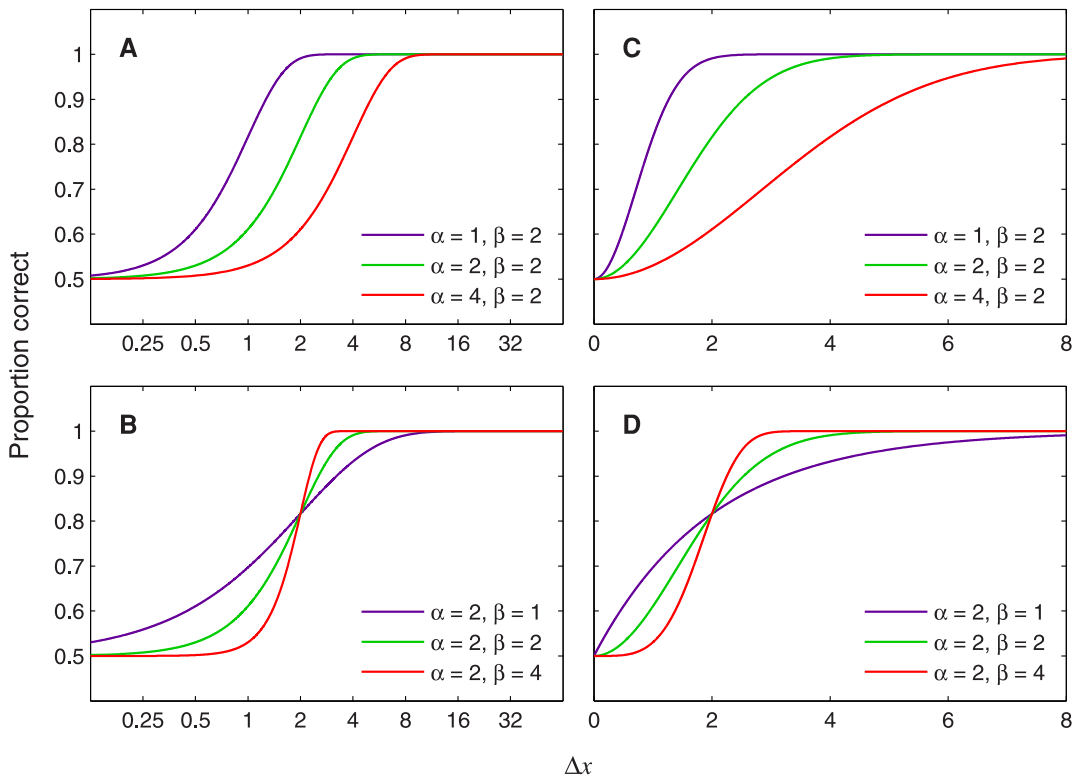table, and these were excluded from the means and medians. The means and medians at the bottom of the table are calculated from those studies for which the pedestal level exceeds the detection threshold, so that both stimuli were clearly visible.
doi:10.1371/journal.pone.0074815.t001

Figure 2. Effect of varying Weibull α and β on log and linear abscissas. (A) Varying α on a log abscissa: The curve shifts horizontally. (B) Varying β on a log abscissa: The curve is linearly stretched or compressed horizontally. (C) Varying α on a linear abscissa: The curve undergoes a linear horizontal stretch *and* a change of threshold. (D) Varying β on a linear abscissa: The shape changes in a way that cannot be described as a linear scaling.
doi:10.1371/journal.pone.0074815.g002

$$z = h(\Delta x) = R(x_p + \Delta x) - R(x_p). \qquad (3)$$

The observer is correct when $z + \epsilon > 0$, i.e. when $\epsilon > -z$. The probability, P, of this happening is given by

$$P = \int_{-z}^{\infty} f(\epsilon)d\epsilon, \qquad (4)$$

where f is the probability density function (PDF) of the noise, ϵ. This integral corresponds to the shaded area in Figure 3A. f has to be even-symmetric, even if the noise added to the output of the transducer is not. This is because the noise sample on z is equal to the noise sample on the target minus the noise sample on the nontarget. This is equivalent to swapping the sign of the nontarget noise sample and adding it to the target noise sample. The sign-reversed noise sample on the nontarget will have a PDF with mirror symmetry relative to the PDF of the noise sample on the target, so the sum of these two values will have an even-symmetric PDF. From the even symmetry of f we have

$$P = F(z) = \int_{-\infty}^{z} f(\epsilon)d\epsilon, \qquad (5)$$

where F is the cumulative distribution function (CDF) of the observer's noise on the internal difference signal, and corresponds

to the shaded area in Figure 3B. So the psychometric function for 2AFC discrimination, expressed as a function of z, will trace out the positive half of the internal noise CDF, increasing from 0.5 to 1 as z increases from 0 to ∞.

Figure 4 plots the CDFs and PDFs for several different forms of noise distribution (the mathematical definitions of these distributions will be given later). These CDFs (plotted as functions of z) do not have a sigmoidal shape: The point of inflection is at zero on the abscissa. This is because the point of inflection corresponds to the peak of the derivative, and the derivative of these functions is the noise PDF, which peaks at 0 in each case.

In summary, F is the CDF of the internal noise, and takes an input of z (Equation (5)); z is the output of h, a function that is determined by the transducer and pedestal, and takes an input of Δx (Equation (3)). The composition of these two functions, $F \circ h$, gives the observer's psychometric function when it is plotted as a function of Δx. We use Ψ to represent this composition of functions:

$$
\begin{aligned}
P &= \Psi(\Delta x) \\
&= (F \circ h)(\Delta x) \\
&= F[h(\Delta x)] \qquad (6) \\
&= F(z) \\
&= F\left[R(x_p + \Delta x) - R(x_p)\right].
\end{aligned}
$$

If we fit the Weibull function, Ψ_{Weibull}, of Equation (1) to the psychometric function, Ψ, of Equation (6), then the Weibull slope parameter, β, will be determined by both the noise CDF, F, and
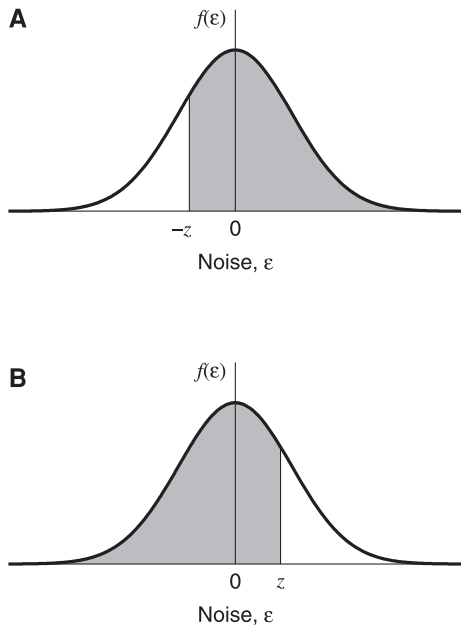
Figure 3. Graphical representation of the probability of a correct response. The shaded areas in A and B correspond to the integrals in Equations (4) and (5), respectively. The smooth curves trace out the PDF of the noise, $f(\epsilon)$, on the internal difference signal, z. As explained in the text, f has to be even-symmetric, and this means that the two integrals in Equations (4) and (5) are equal. The shaded areas correspond to the probability of a correct response. The psychometric function (expressed as a function of z) is the CDF of the noise, increasing from 0.5 to 1 as z increases from 0 to ∞.
doi:10.1371/journal.pone.0074815.g003

the transducer, R. In Theorem 2, we show that, to a good approximation, β can be partitioned into a product of two factors, β^{Noise} and $\beta^{\text{Transducer}}$. β^{Noise} estimates the β of the Weibull function that fits best to the noise CDF, while $\beta^{\text{Transducer}}$ depends on the transducer function. Weibull β is found by multiplying

these two factors together. We derive general analytical formulae for both factors, and then derive, from these formulae, specific expressions for β^{Noise} for a variety of noise distributions, and specific expressions for $\beta^{\text{Transducer}}$ for several commonly used transducer functions.

Our work greatly extends a result previously published by Pelli [19]. He showed that, for 2AFC detection or discrimination,

$$\beta \approx \beta^{\text{Noise}} \times b. \qquad (7)$$

where b is the slope of d' [20] against Δx on log-log axes. Pelli derived this relationship using the concrete example of *contrast detection*, but it is a purely mathematical relationship (outlined in his "Analysis" section, Ref. [19], p. 121), which makes no assumptions about the underlying model, and could equally well be applied to *discrimination* along any unspecified stimulus dimension by replacing the contrast term, c, with Δx in his Equations (14) to (21).

Pelli's analysis ran as follows. Given the *definition* of d' for 2AFC,

$$d' = \sqrt{2}\Phi^{-1}(P) \qquad (8)$$

(where Φ is the cumulative Gaussian), and the *observation* or *assumption* $d' = (\Delta x/\Delta x_{0.76})^b$ (where $\Delta x_{0.76}$ is the value of Δx corresponding to a proportion correct of $\Phi(1/\sqrt{2}) \approx 0.76$, giving $d' = 1$, and b is the log-log slope of d' against Δx), we have

$$P = \Phi\left[(\Delta x/\Delta x_{0.76})^b \big/ \sqrt{2}\right]. \qquad (9)$$

Note that Equation (9) has the same form as Equation (6) if the pedestal, x_p, is zero, the transducer is a power function, $R(x) \propto x^b$, and the internal noise CDF is the cumulative Gaussian (as is usually assumed). If we let β^{Noise} represent the β of the Weibull function, $\Psi_{\text{Weibull}}(\cdot)$, that fits best to the cumulative Gaussian, $\Phi(\cdot)$, then, substituting this Weibull function for $\Phi(\cdot)$ in Equation (9) yields a Weibull function with β given by $\beta^{\text{Noise}} \times b$, which is Relation (7).
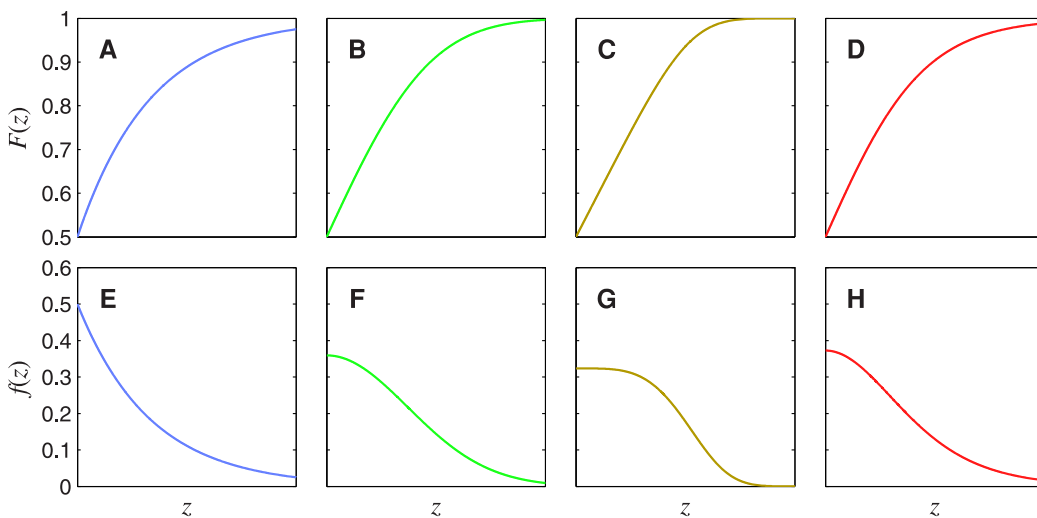

Figure 4. CDFs and PDFs of four different noise distributions. The top row shows noise CDFs, $F(z)$, for (A) a Laplace distribution (generalized Gaussian with $\rho = 1$), (B) a Gaussian distribution (generalized Gaussian with $\rho = 2$), (C) a generalized Gaussian with $\rho = 4$, (D) a logistic distribution. Each panel in the bottom row shows the PDF, $f(z)$, corresponding to the CDF above it. Only the positive halves of the distributions are shown (i.e. $z > 0$). Note that the use of these colours for the different noise distributions is maintained in Figures 7, 8, 10, 11, 12, 15, and 16.
doi:10.1371/journal.pone.0074815.g004

In our terms, the "b" part of Relation (7) is $\beta^{\text{Transducer}}$, the factor determined by the transducer; we will show that, in the case of a power-function transducer and zero pedestal, our general expression for $\beta^{\text{Transducer}}$ reduces to b. We obtain Weibull β by multiplying β^{Noise} and $\beta^{\text{Transducer}}$ together, resulting in an estimated β given by $\beta^{\text{Est}} = \beta^{\text{Noise}} \times \beta^{\text{Transducer}}$, which is equal to $\beta^{\text{Noise}} \times b$ in the scenario just described. In this paper, we derive general analytical expressions for β^{Noise} and $\beta^{\text{Transducer}}$ so that we can easily estimate Weibull β for any combination of noise distribution and transducer function, not just the specific case considered by Pelli.

In many situations, the observer can be modelled using a linear filter. This is equivalent to using a linear transducer, $R(x) = rx$, where r is a constant. For this transducer, Equation (6) gives

$$\Psi(\Delta x) = F(r\Delta x) \text{ for a linear transducer, } R(x) = rx. \quad (10)$$

Thus, the linear observer's psychometric function (plotted on a linear abscissa, Δx) will have the same basic shape as the internal noise CDF, F, just differing by a horizontal scaling factor, r. So, if the observer behaves in a linear fashion, the psychometric function plotted on linear axes gives us a direct plot of the shape of the internal noise CDF. In this situation, since β controls the Weibull function's shape on linear axes, the β that fits best to the psychometric function will be the β that fits best to the noise CDF (the sensitivity parameter, r, will determine the best-fitting α, since α controls the Weibull function's horizontal scaling on linear axes).

The internal noise is usually assumed to be Gaussian, but Neri [18] has recently disputed this assumption. Using reverse correlation methods, he attempted to measure both the "deterministic transformation" (in our terms, the transducer function for contrast), and the shape of the internal noise distribution. He concluded that, for temporal 2AFC detection of a bright bar in noise, the contrast transducer was linear, and the internal noise had a Laplace distribution (whose CDF and PDF are given in Figures 4A and 4E, respectively). This is a radical departure from the Gaussian assumption that has usually been made since the invention of signal detection theory in the 1950s [14]. The Laplace distribution has higher kurtosis (i.e., has a sharper peak and heavier tails) than the Gaussian (compare Figure 4E with 4F). As we shall see later on, for positive z, the Laplace distribution has a CDF that takes the form of a Weibull function with $\beta = 1$. Since the psychometric function has the same shape as the internal noise CDF for a linear observer, Neri's proposal that the transducer is linear and the internal noise has a stimulus-independent Laplace distribution predicts that the observer's psychometric function should, like the Laplace CDF, be a Weibull function with $\beta = 1$. As noted earlier (and shown in Table 1 and Figure 1), this does not generally seem to be the case – with noise-free stimuli, β is around 3 for contrast detection and, even for suprathreshold contrast discrimination, where β is substantially lower, it is still usually found to be greater than 1; later, we shall show that, assuming additive noise, these β values are more consistent with a distribution that has *lower* kurtosis than a Gaussian.

Although the whole of this paper is couched in terms of the transducer model, it is not necessary to accept the transducer model to find the results useful; we just have to assume that the psychometric function has a form *consistent* with a particular combination of internal noise distribution and transducer function. For example, the intrinsic uncertainty model produces psychometric functions that are *consistent* with additive noise following an expansive power-function transducer with exponent that increases with channel uncertainty [21], but the model itself has no explicit

transducer. Alternatively, suppose the observer carries out the discrimination task by making noisy estimates of each stimulus value and comparing them. Due to the noise, repeated presentations of the same stimulus value, x, will give a distribution of estimated values, $\hat{x}$, around the mean estimate. If we can find a function, R, such that the shape and width of the distribution of $R(\hat{x})$ is independent of x, then the observer is equivalent to a transducer model with additive noise. In this class of model, the stimulus value, x, is transduced to give $R(x)$, and then stimulus-independent noise is added to the signal. But we do not have to assume that this is literally how the observer works – the noisy estimates of the stimulus values could have arisen from all sorts of mechanisms, not just a transducer followed by additive noise.

In keeping with our terminology of P_θ for the threshold performance level, we introduce the terms z_θ and Δx_θ to represent the values of z and Δx at threshold, i.e. the values of z and Δx when the proportion correct is P_θ, which we define as $1 - 0.5/e$.

Theorem 1: A General Expression for the Psychometric Function in Terms of the Stimulus Values and the Threshold

Introduction

Equation (6) gives a general equation for the psychometric function in terms of the transducer function, R, and the noise CDF, F. The sensitivity of the system (which determines the discrimination threshold, Δx_θ) can be adjusted either by changing the gain of the transducer function (i.e., stretching or compressing R along its vertical axis), or by adjusting the spread of the noise CDF (i.e., stretching or compressing F along its horizontal axis), or both. Since the units in which we express the internal signal are arbitrary, researchers will usually either (1) fix the spread of the noise CDF at some convenient standard value (say, unit variance), and vary the transducer gain to achieve the desired threshold, or (2) fix the gain of the transducer at some convenient standard value (say, unit gain), and vary the spread of the noise CDF to achieve the desired threshold. However, for our purposes, it is more convenient to reformulate Equation (6) so that *both* the spread of the noise CDF *and* the gain of the transducer are set to convenient values, and the threshold is specified directly. This allows us to consider general forms of the transducer and noise, without having to worry about specifying the gain of the transducer or spread of the noise correctly – the reformulated equation will take care of the spread of the psychometric function automatically. Theorem 1 derives an expression for the psychometric function that meets these requirements.

Statement of Theorem 1

Theorem 1 has three parts:

1. The expression for the psychometric function, $\Psi(\Delta x)$, in Equation (6) can be rewritten as

$$\Psi(\Delta x) = F\left(F^{-1}(P_\theta) \frac{R(x_p + \Delta x) - R(x_p)}{R(x_p + \Delta x_\theta) - R(x_p)} \right), \quad (11)$$

where Δx_θ is the stimulus difference corresponding to a performance level of P_θ.

2. If we change the gain of the transducer by replacing the function $R(x)$ with $rR(x)$, this will have no effect on the psychometric function, $\Psi(\Delta x)$, in Equation (11).
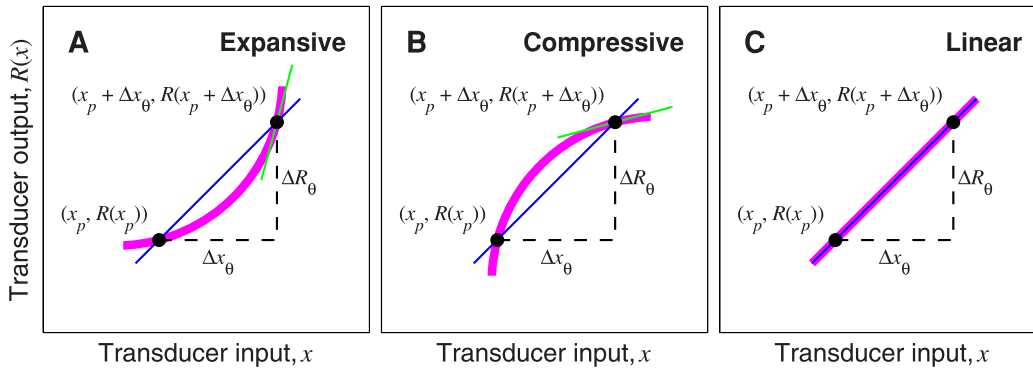
Figure 5. Geometrical interpretation of the expression for $\beta^{\text{Transducer}}$. In each panel, the thick, magenta curve represents the transducer function. The horizontal axes represent the transducer input, and the vertical axes represent the transducer output. x_p is the pedestal level, and Δx_θ is the discrimination threshold. The gradient of the blue line, $\Delta R_\theta / \Delta x_\theta$, is equal to V, defined in Equation (39). The green line is the tangent to the transducer at point $(x_p + \Delta x_\theta, R(x_p + \Delta x_\theta))$; its gradient is equal to U, defined in Equation (38). The ratio U/V is equal to $\beta^{\text{Transducer}}$. For an expansive transducer (panel A), $U > V$, so $\beta^{\text{Transducer}} > 1$. For a compressive transducer (panel B), $U < V$, so $\beta^{\text{Transducer}} < 1$. For a linear transducer (panel C), $U = V$, so $\beta^{\text{Transducer}} = 1$.
doi:10.1371/journal.pone.0074815.g005

3. Similarly, if we change the spread of the noise CDF by replacing the function $F(z)$ with $F(z/s)$, this will have no effect on $\Psi(\Delta x)$ in Equation (11).

Proof of Theorem 1

First, let us substitute the threshold values of P and Δx into Equation (6):

$$P_\theta = F\left[R(x_p + \Delta x_\theta) - R(x_p)\right]. \qquad (12)$$

Equation (12) can be rearranged to give

$$\frac{F^{-1}(P_\theta)}{R(x_p + \Delta x_\theta) - R(x_p)} = 1. \qquad (13)$$

Since the left hand side of Equation (13) is equal to 1, we can multiply anything by this expression, and leave it unchanged. Multiplying the argument of F in (6) by this expression, we obtain Equation (11), which proves Part 1 of the theorem. If we replace the transducer, $R(x)$, in Equation (11) with one that has a different gain, $rR(x)$, the r's will obviously cancel out, leaving the psychometric function, Ψ, unchanged, which proves Part 2 of the theorem. To prove Part 3 of the theorem, consider what happens if we replace the function, $F(z)$, in Equation (11) with one that has a different spread, $F(z/s)$. Then the inverse function is given by $sF^{-1}(P)$, and the s's cancel out:

$$\Psi(\Delta x) = F\left(\frac{sF^{-1}(P_\theta)\frac{R(x_p + \Delta x) - R(x_p)}{R(x_p + \Delta x_\theta) - R(x_p)}}{s}\right)$$

$$= F\left(F^{-1}(P_\theta)\frac{R(x_p + \Delta x) - R(x_p)}{R(x_p + \Delta x_\theta) - R(x_p)}\right),$$

which is identical to Equation (11). $\square$

Discussion of Theorem 1

Equation (11) gives us an expression for the psychometric function (parameterized by the threshold, Δx_θ) in which we can use any convenient standard form of the transducer function or noise distribution, without having to worry about setting the right gain or spread.

Although, for most of this paper, we define the threshold as the stimulus difference that gives rise to a performance level, P_θ, defined as $1 - 0.5/e$, Theorem 1 actually holds for any value that P_θ could have taken.

Note that, in the special case of a zero pedestal ($x_p = 0$) and a transducer that gives zero output for zero input ($R(0) = 0$), Equation (11) reduces to

$$\Psi(\Delta x) = F\left(F^{-1}(P_\theta)\frac{R(\Delta x)}{R(\Delta x_\theta)}\right) \text{ if } x_p = 0 \text{ and } R(0) = 0. \qquad (14)$$

Theorem 2. An Expression That Estimates the Best-Fitting Weibull β for Unspecified Noise and Transducer

Statement of Theorem 2

Theorem 2 has two parts:

1. If the parameters of the Weibull function, $\Psi_{\text{Weibull}}(\Delta x)$, of Equation (1) can be set to provide a good fit to Equation (6), then the best-fitting beta will be well approximated by

$$\beta^{\text{Est}} = \beta^{\text{Noise}} \times \beta^{\text{Transducer}}, \qquad (15)$$

where β^{Noise} and $\beta^{\text{Transducer}}$ are given by the following expressions:

$$\beta^{\text{Noise}} = 2eF^{-1}(P_\theta)f\left(F^{-1}(P_\theta)\right), \qquad (16)$$

$$\beta^{\text{Transducer}} = \frac{R'(x_p + \Delta x_\theta)\Delta x_\theta}{R(x_p + \Delta x_\theta) - R(x_p)}, \qquad (17)$$

and f and R' are the derivatives of, respectively, F and R with respect to their inputs.

2. β^{Noise} is an estimate of the β of the Weibull function that fits best to the noise CDF, F, in Equation (6).

Proof of Theorem 2

By assumption, the Weibull function provides a close fit to Ψ of Equation (6), so the gradient of Ψ at threshold will closely match the gradient of the best-fitting Weibull function at threshold. Therefore, since β is proportional to the gradient of the Weibull function at threshold with an abscissa of $\log(\Delta x)$, we can derive a close approximation to β from the gradient of Ψ at threshold on this abscissa. To create a log abscissa, let $y = \ln(\Delta x)$, so that

$$\Delta x = e^y. \tag{18}$$

If we substitute e^y for Δx in Equation (1), we find that the gradient of the Weibull function on the log abscissa, y, is given by

$$\frac{dP_{\text{Weibull}}}{dy} = \frac{\beta}{2} \left(\frac{\Delta x}{\alpha}\right)^{\beta} \exp\left[-\left(\frac{\Delta x}{\alpha}\right)^{\beta}\right]. \tag{19}$$

For the Weibull function at threshold performance ($P_{\text{Weibull}} = P_\theta$), it follows that $\Delta x = \alpha$. Substituting α for Δx in (19) gives

$$\left.\frac{dP_{\text{Weibull}}}{dy}\right|_{P_{\text{Weibull}} = P_\theta} = \frac{\beta}{2e}, \tag{20}$$

and so

$$\beta = 2e \left.\frac{dP_{\text{Weibull}}}{dy}\right|_{P_{\text{Weibull}} = P_\theta}. \tag{21}$$

To evaluate Equation (21), we use the chain rule to expand the derivative:

$$\beta = 2e \left.\frac{dP_{\text{Weibull}}}{dz} \times \frac{dz}{dy}\right|_{P_{\text{Weibull}} = P_\theta}. \tag{22}$$

As noted above, the assumed good fit of the Weibull function, Ψ_{Weibull}, of Equation (1) to Ψ of Equation (6) means that the output, P_{Weibull}, of the Weibull function is close to the output of Ψ, which is the proportion correct, P. Substituting P for P_{Weibull} in Equation (22) therefore gives us a good estimate of β, which we call β^{Est}:

$$\beta^{\text{Est}} = 2e \left.\frac{dP}{dz} \times \frac{dz}{dy}\right|_{P = P_\theta}. \tag{23}$$

From Equation (6), we see that $P = F(z)$, so dP/dz is given by $f(z)$, the noise PDF (which is the derivative of $F(z)$ with respect to z). At threshold, $z = z_\theta$, and so,

$$\beta^{\text{Est}} = 2e f(z_\theta) \times \left.\frac{dz}{dy}\right|_{P = P_\theta}. \tag{24}$$

We will see that the first part of Equation (24), $2ef(z_\theta)$, is proportional to β^{Noise}, the β-estimate of the Weibull function that fits best to the noise CDF, and the second part, dz/dy at threshold,

is proportional to $\beta^{\text{Transducer}}$ defined above. Most of the work involves deriving an expression for dz/dy at threshold.

Using Equation (18) to substitute for Δx in Equation (3), we get

$$z = R(x_p + e^y) - R(x_p). \tag{25}$$

Let us define x_t as the target stimulus value:

$$x_t = x_p + \Delta x = x_p + e^y. \tag{26}$$

Using Equation (26) to substitute for $x_p + e^y$ in Equation (25), we have

$$z = R(x_t) - R(x_p). \tag{27}$$

Then,

$$\begin{aligned} \frac{dz}{dy} &= \frac{dz}{dx_t} \times \frac{dx_t}{dy} \\ &= R'(x_t)e^y \\ &= R'(x_p + \Delta x)\Delta x, \end{aligned} \tag{28}$$

where R' is the derivative of R with respect to its input. At threshold, we can substitute Δx_θ for Δx in Equation (28), giving

$$\left.\frac{dz}{dy}\right|_{P = P_\theta} = R'(x_p + \Delta x_\theta)\Delta x_\theta. \tag{29}$$

Using Equation (29) to substitute for dz/dy in Equation (24), we obtain

$$\beta^{\text{Est}} = 2e f(z_\theta) R'(x_p + \Delta x_\theta)\Delta x_\theta. \tag{30}$$

From Equation (6), we have $P = F(z)$, so, considering the values of P and z at threshold,

$$z_\theta = F^{-1}(P_\theta). \tag{31}$$

Using Equation (31) to substitute for z_θ in Equation (30), we have

$$\beta^{\text{Est}} = 2e f\left(F^{-1}(P_\theta)\right) R'(x_p + \Delta x_\theta)\Delta x_\theta. \tag{32}$$

To evaluate this expression as written in Equation (32), we need to know the gain of the transducer and the spread of the noise CDF, or at least their ratio. However, if we know the shape of the transducer (apart from the gain), and we know the shape of the noise CDF (apart from the spread), we can work out the ratio of gain to spread from Δx_θ. But it is much more convenient to reformulate Equation (32) so that this is taken care of, and we can arbitrarily set the spread of the noise CDF and the gain of the transducer to any convenient values. We can use the same trick that we used in Theorem 1: We multiply the expression in Equation (32) by the left hand side of Equation (13), which equals 1. After doing this, and rearranging the terms, we obtain

$$\beta^{\text{Est}} = 2e F^{-1}(P_\theta) f\left(F^{-1}(P_\theta)\right) \times \frac{R'(x_p + \Delta x_\theta)\Delta x_\theta}{R(x_p + \Delta x_\theta) - R(x_p)}. \tag{33}$$

Equation (33) can be written in the form given by Equations (15) to (17), which proves the Part 1 of the theorem.
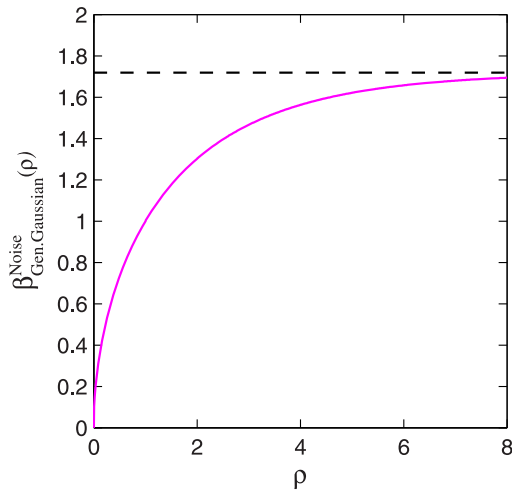
Figure 6. $\beta^{\text{Noise}}_{\text{Gen.Gaussian}}(\rho)$ **plotted as a function of ρ.** This curve plots the predicted β when the Weibull function is fitted to the CDF of generalized Gaussian distributions with a range of different ρ values. The graph asymptotes to a value of $e-1$ (see Appendix S1), indicated by the horizontal dashed line. The shape of the generalized Gaussian distribution is determined by ρ. ρ-values of 1 and 2 are special cases: $\rho=1$ gives a Laplace distribution, and $\rho=2$ gives a Gaussian distribution.
doi:10.1371/journal.pone.0074815.g006

We now prove Part 2, that β^{Noise} is the estimated β of the Weibull function that fits best to the noise CDF, F. First, note that all linear transducers have the form $R(x)=rx$. This gives $R'(x)=r$, and so, from Equation (17), $\beta^{\text{Transducer}}=1$, regardless of the value of r, x_p or Δx_θ. Therefore, from (15), $\beta^{\text{Est}}=\beta^{\text{Noise}}$ for a linear transducer. Now, consider the linear transducer $R(x)=x$. For this transducer, Equation (6) gives $\Psi(\Delta x)=F(\Delta x)$. The estimate of β when the Weibull function, $\Psi_{\text{Weibull}}(\Delta x)$, is fitted to $\Psi(\Delta x)$ is given by $\beta^{\text{Est}}=\beta^{\text{Noise}}$, as it will be for any linear transducer. Since, in this case, $\Psi(\Delta x)=F(\Delta x)$, the Weibull function has also been fitted to the noise CDF, and the estimated β of this fitted function is given by $\beta^{\text{Est}}=\beta^{\text{Noise}}$.□

Discussion of Theorem 2

To get an intuition into how Weibull β is partitioned into the two terms, β^{Noise} and $\beta^{\text{Transducer}}$, let us refer back to Equation (21). This equation shows that β is proportional to dP_{Weibull}/dy at threshold. We used the chain rule to express dP_{Weibull}/dy as $dP_{\text{Weibull}}/dz \times dz/dy$, which is approximately equal to $f(z_\theta) \times dz/dy$. $f(z_\theta)$ depends only on the noise distribution, and is proportional to β^{Noise}; dz/dy at threshold generally depends on the transducer, the pedestal and the threshold, and is proportional to $\beta^{\text{Transducer}}$; their product is proportional to Weibull β. This is essentially where Equations (15) to (17) come from. The equations were tidied up by specifying the constants of proportionality, and defining β^{Noise} and $\beta^{\text{Transducer}}$ in such a way that they are independent of any horizontal scaling of the noise distribution, or any vertical scaling of the transducer function. Thus, the β^{Noise} term will be the same for, for example, all Gaussian distributions,
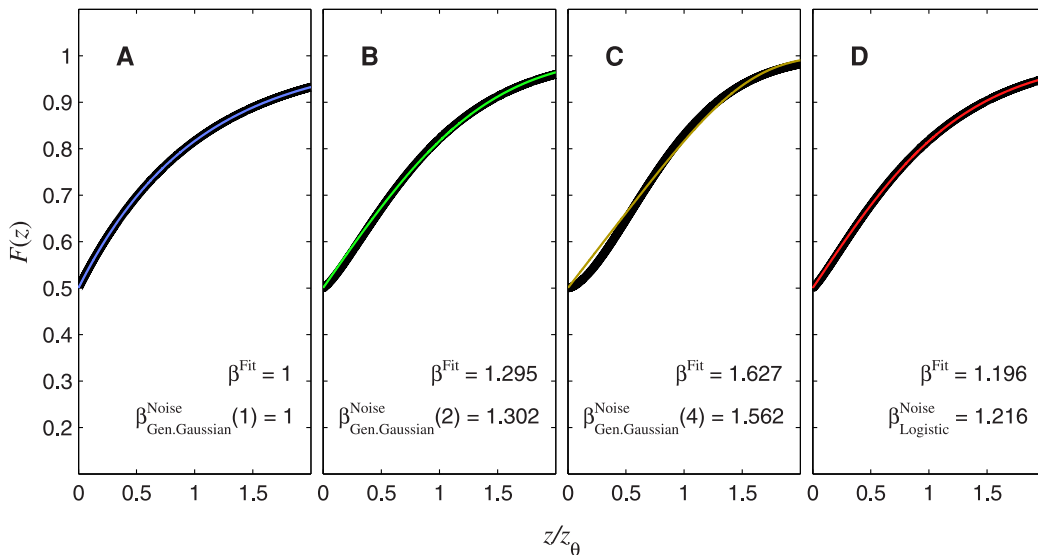

Figure 7. Noise CDFs from Figure 4 plotted against the best-fitting Weibull functions. The thin, coloured curves shown in (A) to (D) are the CDFs from Figures 4A to 4D, respectively. The thick, black curves are the Weibull functions that give the best (maximum-likelihood) fit across the range of inputs shown on the horizontal axis. This fit was carried out by maximizing the expression $\sum_z F(z)\log[\Psi_{\text{Weibull}}(z)] + [1-F(z)]\log[1-\Psi_{\text{Weibull}}(z)]$, where F is the noise CDF, and Ψ_{Weibull} is the Weibull function whose parameters were being fitted. The Weibull function provides a perfect fit to the Laplace CDF (A), an excellent fit to the Gaussian (B), and logistic (D) CDFs, and an acceptable fit to the generalized Gaussian with $\rho=4$ (C); this partly justifies our use of P as an estimate of P_{Weibull} in Equation (23). The β^{Fit} values are the β parameters of the fitted Weibull functions. The β^{Noise} values are our analytical estimates of β^{Fit}, given by Equations (52) to (54) for panels (A) to (C), respectively, and Equation (60) for panel (D). In each case, β^{Noise} provides a close match to β^{Fit}. The parameter in brackets in each $\beta^{\text{Noise}}_{\text{Gen.Gaussian}}$ term is the shape parameter, ρ (see Equation (50)). As noted in the text, the CDFs all have a point of inflection at zero. With the exception of panel A, the best-fitting Weibull functions have a point of inflection slightly above zero (β would have to be 1 or less for the steepest point to occur at zero). Nevertheless, the Weibull functions still provide good fits.
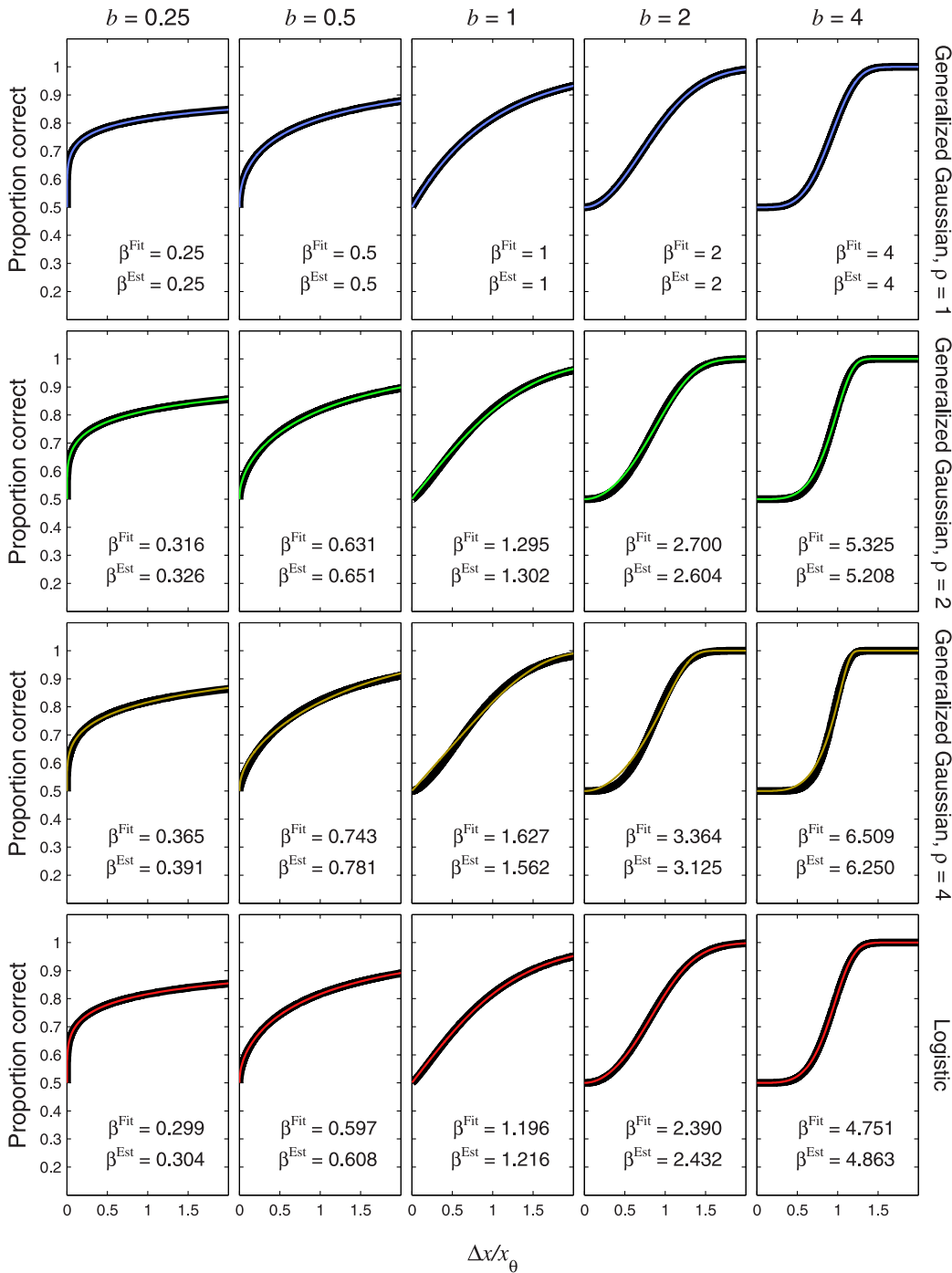doi:10.1371/journal.pone.0074815.g007

Figure 8. Psychometric functions resulting from power-function transducers and zero pedestal. The thin, coloured curves show the psychometric function of Equation (65), plotted as a function of $\Delta x/\Delta x_\theta$. Different rows of panels show psychometric functions with different noise CDFs, F, given by the Laplace distribution (top row of panels), the Gaussian (second row), the generalized Gaussian with $\rho=4$ (third row) or logistic (bottom row). Different columns of panels show psychometric functions for different transducer exponents, b, as indicated at the top of the figure. The thick, black curves show the best-fitting (maximum-likelihood) Weibull functions. The curves in the middle column ($b=1$, top to bottom) are identical to Figures 7A to 7D, respectively. This is because $b=1$ gives a linear transducer, and so the psychometric functions for $b=1$ will have the same shape and same fitted β as the CDF (see Equation (10)). Each panel displays the β value of the best-fitting Weibull function (β^{Fit}) and the estimate, $\beta^{\text{Est}} = \beta^{\text{Noise}} \times b$, where β^{Noise} is given by Equation (52), (53), (54) or (60), as appropriate.
doi:10.1371/journal.pone.0074815.g008

whatever the spread, and the $\beta^{\text{Transducer}}$ term will be the same for, for example, all power functions with a particular exponent, whatever the gain.

Equation (17) expresses $\beta^{\text{Transducer}}$ as a function of the threshold stimulus difference, Δx_θ. Alternatively, for nonzero pedestals, we can reformulate Equation (17) as a function of the Weber fraction, W, defined as the ratio $\Delta x/x_p$ at threshold:
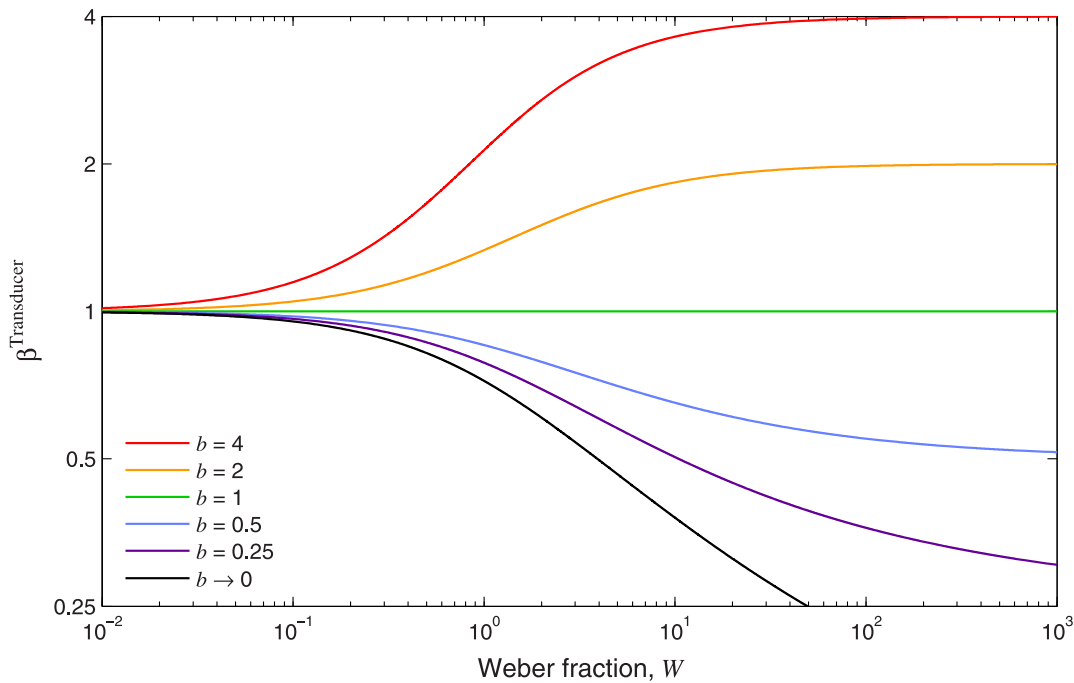
Figure 9. $\beta^{\text{Transducer}}$ for the power-function transducer with nonzero pedestal, plotted as a function of Weber fraction. Each curve gives $\beta^{\text{Transducer}}$ for a different transducer exponent, b. $\beta^{\text{Transducer}}$ asymptotes towards 1 as W decreases, and towards b as W increases. For typical Weber fractions of less than 0.3 (see Table 1), $\beta^{\text{Transducer}}$ does not deviate much from 1. The bottom curve, in black, shows the limiting case, as $b \to 0$. All the plotted functions except the one for $b \to 0$ are given by Equation (67). In Theorem 4B, we prove that the limiting case as $b \to 0$ is identical to the curve corresponding to a logarithmic transducer; this curve is given by Equation (70).
doi:10.1371/journal.pone.0074815.g009

$$W = \Delta x_\theta / x_p. \qquad (34)$$

From Equation (34), we obtain $\Delta x_\theta = W x_p$, and, using this expression to substitute for Δx_θ in Equation (17), we can rewrite the expression for $\beta^{\text{Transducer}}$ in terms of W:

$$\beta^{\text{Transducer}} = \frac{R'\big(x_p(1+W)\big) W x_p}{R\big(x_p(1+W)\big) - R(x_p)} \text{ if } x_p \neq 0. \qquad (35)$$

The Weber fraction can only be defined if $x_p \neq 0$. If $x_p = 0$ and $R(0) = 0$, Equation (17) reduces to

$$\beta^{\text{Transducer}} = \frac{R'(\Delta x_\theta)\Delta x_\theta}{R(\Delta x_\theta)} \text{ if } x_p = 0 \text{ and } R(0) = 0. \qquad (36)$$

When the stimulus dimension of interest is contrast, a discrimination experiment with a zero pedestal is called a contrast detection experiment.

One important property of $\beta^{\text{Transducer}}$ is that it is always greater than 1 for a fully expansive transducer function (i.e., one for which the slope always increases away from zero with increasing input), and is always less than 1 for a fully compressive transducer function (i.e., one for which the slope always decreases towards zero with increasing input). Here we provide a geometrical argument (illustrated in Figure 5) to explain why this is the case.

First, note that we can rewrite Equation (17) as

$$\beta^{\text{Transducer}} = \frac{U}{V}, \qquad (37)$$

where

$$U = R'\big(x_p + \Delta x_\theta\big), \qquad (38)$$

and

$$V = \frac{R\big(x_p + \Delta x_\theta\big) - R\big(x_p\big)}{\Delta x_\theta} = \frac{\Delta R_\theta}{\Delta x_\theta}, \qquad (39)$$

with

$$\Delta R_\theta = R\big(x_p + \Delta x_\theta\big) - R\big(x_p\big). \qquad (40)$$

These quantities are illustrated for an expansive transducer in Figure 5A, where the thick, magenta curve represents the transducer. The filled circles mark points $\big(x_p, R(x_p)\big)$ and $\big(x_p + \Delta x_\theta, R(x_p + \Delta x_\theta)\big)$. The gradient of the blue line connecting these two points is V, defined in Equation (39). The short, green, line segment is the tangent to the transducer at $\big(x_p + \Delta x_\theta, R(x_p + \Delta x_\theta)\big)$; its gradient is U, defined in Equation (38). It is clear from the diagram that, for an expansive transducer, like the one illustrated, the gradient of the transducer at $\big(x_p + \Delta x_\theta, R(x_p + \Delta x_\theta)\big)$ must always be steeper than the blue line, because, as we travel along the transducer function from
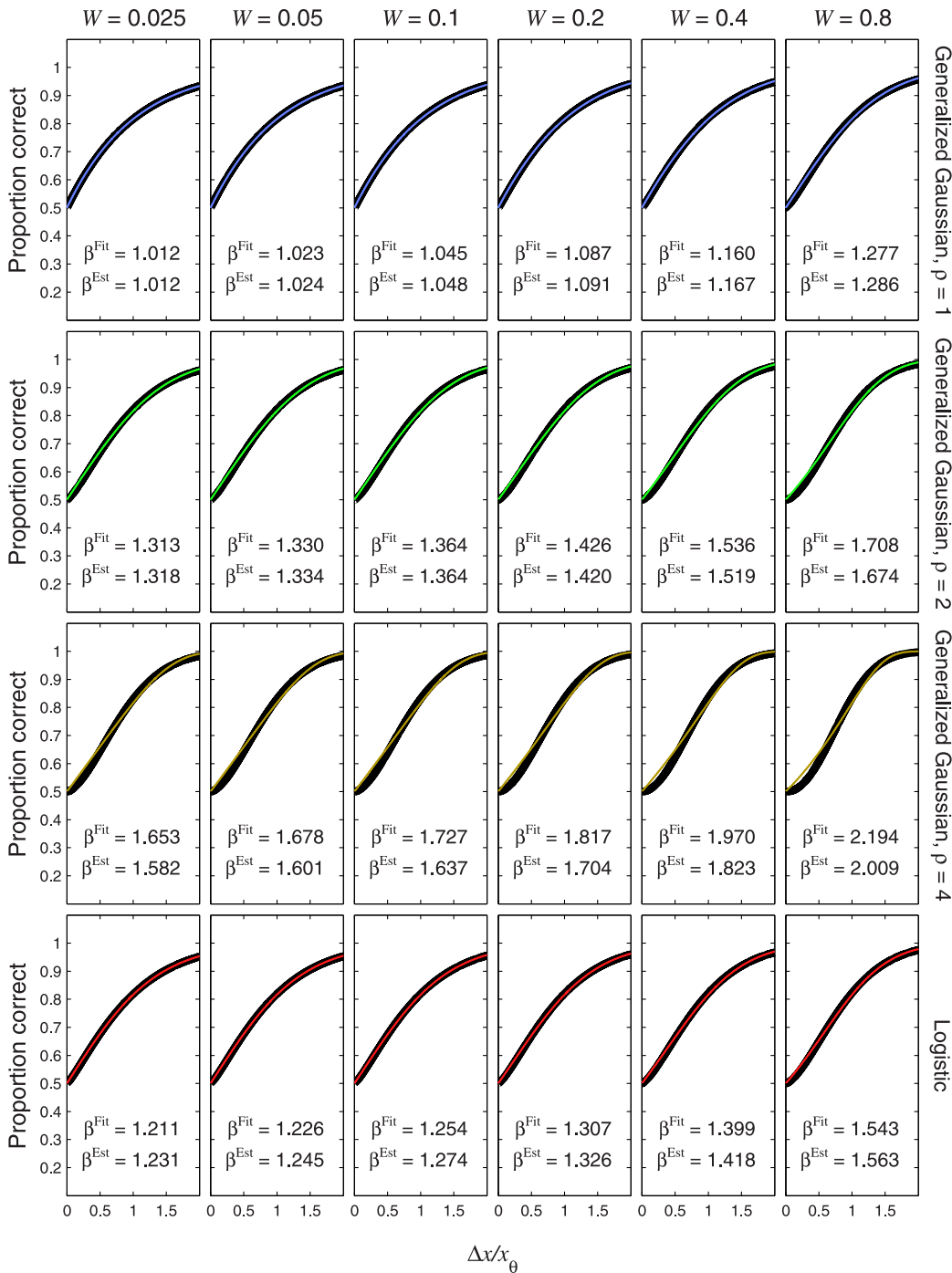
Figure 10. Psychometric functions resulting from transducer $R \propto x^2$ and nonzero pedestal. The thin, coloured curves show the psychometric function of Equation (69) with $b = 2$. Different rows of panels show psychometric functions with different noise CDFs, as indicated on the right of the figure. Different columns of panels show psychometric functions for different Weber fractions, W. The thick, black curves show the best-fitting (maximum-likelihood) Weibull functions. Each panel displays the β value of the best-fitting Weibull function (β^{Fit}) and the estimate, $\beta^{\text{Est}} = \beta^{\text{Noise}} \times \beta^{\text{Transducer}}_{\text{Power func, } x_p \neq 0}$, where $\beta^{\text{Transducer}}_{\text{Power func, } x_p \neq 0}$ is given by Equation (67), and β^{Noise} is given by Equation (52), (53), (54) or (60), as appropriate.
doi:10.1371/journal.pone.0074815.g010

$(x_p, R(x_p))$ to $(x_p + \Delta x_\theta, R(x_p + \Delta x_\theta))$, the transducer approaches the second point from below the blue line. Therefore, U must always be greater than V, so, from Equation (37), $\beta^{\text{Transducer}}$ must always be greater than 1.

Figure 5B illustrates the situation for a compressive transducer. Here, as we travel along the transducer function from $(x_p, R(x_p))$ to $(x_p + \Delta x_\theta, R(x_p + \Delta x_\theta))$, the transducer approaches the second point from above the blue line, and so the gradient of the transducer at the second point must be lower than the gradient of
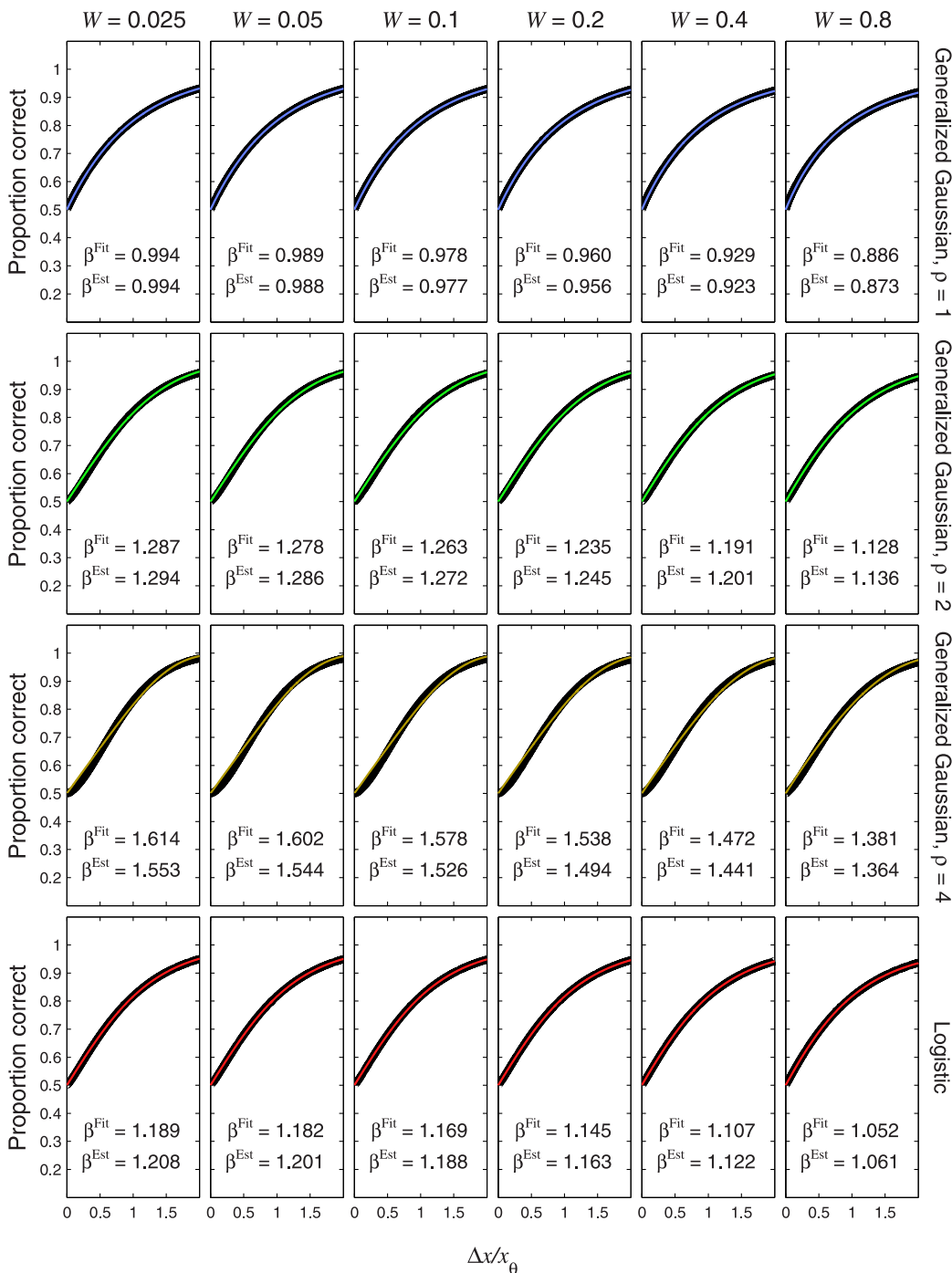
Figure 11. Psychometric functions resulting from transducer $R \propto x^{0.5}$ and nonzero pedestal. All details are the same as in Figure 10, except that the transducer exponent is 0.5.
doi:10.1371/journal.pone.0074815.g011

the blue line. Thus, U must always be less than V, so, from Equation (37), $\beta^{\text{Transducer}}$ must always be less than 1.

Finally, Figure 5C illustrates the situation for a linear transducer, i.e. one that is neither expansive nor compressive. Here, the gradient of the transducer is equal to the gradient of the blue line, so $U = V$, and therefore $\beta^{\text{Transducer}} = 1$. This provides a geometrical insight into the previously proved fact that $\beta^{\text{Transducer}} = 1$ for a linear transducer.

In conclusion, Weibull β can be partitioned into two factors: β^{Noise} (Equation (16)), which estimates the β of the Weibull function that fits best to the noise CDF, F; and $\beta^{\text{Transducer}}$ (Equation (17), (35) or (36)), which is determined partly (or, as we shall see, sometimes completely) by the shape of the transducer function, R. $\beta^{\text{Transducer}}$ is greater than 1 for an expansive transducer, less than 1 for a compressive transducer, and equal to 1 for a linear transducer. β^{Noise} is independent of the spread (i.e.

horizontal scaling) of the CDF (analogously, Weibull β is independent of the spread of the Weibull function on linear axes); $\beta^{\text{Transducer}}$ is independent of the gain (i.e. vertical scaling) of the transducer. Multiplying β^{Noise} and $\beta^{\text{Transducer}}$ together gives us β^{Est}, the estimate of Weibull β. The expressions for β^{Noise} and $\beta^{\text{Transducer}}$ derived above are completely general. In later sections, we derive values for β^{Noise} given specific noise distributions, and expressions for $\beta^{\text{Transducer}}$ given specific transducers.

There are two possible sources of error in the Weibull β estimate, β^{Est}. Firstly, the derivation of the expression for β^{Est} relies on the use of dP/dz as an approximation of dP_{Weibull}/dz at threshold in the step from Equation (22) to (23), where P is the output of the psychometric function, Ψ, and P_{Weibull} is the output of the best-fitting Weibull function. The accuracy of β^{Est} relies on these two slopes being close at the threshold performance level. A second potential source of inaccuracy is that, even if these two slopes are very close at the threshold level, the overall psychometric function, Ψ, might still not be well fit by a Weibull function, in which case the best-fitting Weibull β could deviate substantially from β^{Est}. However, as we will show, in the range of conditions usually encountered, the Weibull function does provide a good fit to the psychometric function, so β^{Est} is accurate. In cases where Ψ is a Weibull function, the best-fitting Weibull function will fit exactly, and β^{Est} gives the exact value of the best-fitting Weibull β.

Deriving β^{Noise} for Specific Noise Distributions

As proved in Theorem 2, β^{Noise} is an estimate of the β of the Weibull function that fits best to the noise CDF. In this section, we evaluate the analytical expression for β^{Noise} (Equation (16)) for several different noise distributions. We also compare each value with the β value obtained by fitting the Weibull function to the noise CDF numerically. There is of course no single correct answer to the question of what is the best-fitting β – it depends on both the fitting criterion and the points on the psychometric function that are sampled. When Pelli [19] fitted the Weibull function to the Gaussian CDF, he minimized the maximum error over all positive inputs. We instead performed a maximum-likelihood fit over all inputs from 0 to twice the threshold (actually, we approximated this by sampling the psychometric function in discrete steps of one thousandth of the threshold). Our rationale for this approach was that fitting the psychometric function is usually done by maximum likelihood, and the threshold usually falls around the middle of the set of stimulus values.

Evaluating β^{Noise} for a generalized Gaussian noise CDF

Most psychophysical models use Gaussian noise. This is partly because the Gaussian is often easy to handle analytically, but also because, according to the Central Limit Theorem, the sum of independent sources of noise tends towards a Gaussian-distributed random variable, whatever the distribution of the individual noise sources. However, as noted earlier, Neri [18] has recently argued that internal sensory noise is closer to a Laplace distribution. Both the Gaussian and the Laplace are parameterizations of the generalized Gaussian, which we consider in this section.

The generalized Gaussian CDF is given by the following expression, with horizontal scaling (i.e. spread) determined by τ, and shape determined by ρ:

$$P = F_{\text{Gen.Gaussian}}(z;\rho,\tau) = \frac{1 + \text{sgn}(z)\gamma[(|z|/\tau)^\rho; 1/\rho]}{2}, \quad (41)$$

where $\text{sgn}(z) = 1$ for $z \geq 0$ and -1 for $z < 0$, and γ is the lower incomplete gamma function, defined as

$$\gamma(\zeta; v) = \frac{1}{\Gamma(v)} \int_0^\zeta t^{v-1} e^{-t} dt. \quad (42)$$

$\Gamma(v)$ in Equation (42) is the gamma function, which is a continuous generalization of the factorial, given by

$$\Gamma(v) = \int_0^\infty t^{v-1} e^{-t} dt. \quad (43)$$

Note, the lower incomplete gamma function is often defined without the normalization term, $\Gamma(v)$, but it is more convenient for us to define it as in Equation (42), because otherwise we would just have to divide by $\Gamma(v)$ anyway, complicating the expression for the generalized Gaussian in Equation (41); in addition, the MATLAB function gammainc evaluates the function as defined in Equation (42).

The variance of the generalized Gaussian distribution is given by

$$\sigma_z{}^2 = \frac{\tau^2 \Gamma(3/\rho)}{\Gamma(1/\rho)}. \quad (44)$$

We use the subscript, z, in Equation (44) to indicate that this is the variance of the noise on the difference of mean signals, z, as opposed to the variance of the noise on the transducer outputs, which we could call $\sigma_R{}^2$. As long as the noise on the two transducer outputs within a trial is uncorrelated and has zero mean, then we have $\sigma_z{}^2 = 2\sigma_R{}^2$, and so $\sigma_z = \sigma_R\sqrt{2}$, whatever form the noise CDF takes.

The PDF of the generalized Gaussian distribution is given by the derivative of the CDF:

$$f_{\text{Gen.Gaussian}}(z;\rho,\tau) = \frac{\rho}{2\tau\Gamma(1/\rho)} \exp\left(-\left(\frac{|z|}{\tau}\right)^\rho\right). \quad (45)$$

As noted above, the shape of the distribution is determined by the parameter, ρ. When $\rho = 2$, Equation (45) describes the Gaussian PDF:

$$f_{\text{Gen.Gaussian}}(z;2,\tau) = \frac{1}{\sigma_z\sqrt{2\pi}} \exp\left(-\frac{z^2}{2\sigma_z{}^2}\right). \quad (46)$$

When $\rho = 1$, Equation (45) describes the Laplace PDF:

$$f_{\text{Gen.Gaussian}}(z;1,\tau) = \frac{1}{\sigma_z\sqrt{2}} \exp\left(-\frac{|z|}{\sigma_z/\sqrt{2}}\right). \quad (47)$$

For positive z, the inverse of the generalized Gaussian CDF is given by

$$z = F_{\text{Gen.Gaussian}}^{-1}(P;\rho,\tau)$$
$$= \tau\left[\gamma^{-1}(2P-1; 1/\rho)\right]^{1/\rho} \quad (48)$$

(we don't need to worry about negative z, because, for any monotonically increasing transducer, and positive Δx, z as defined
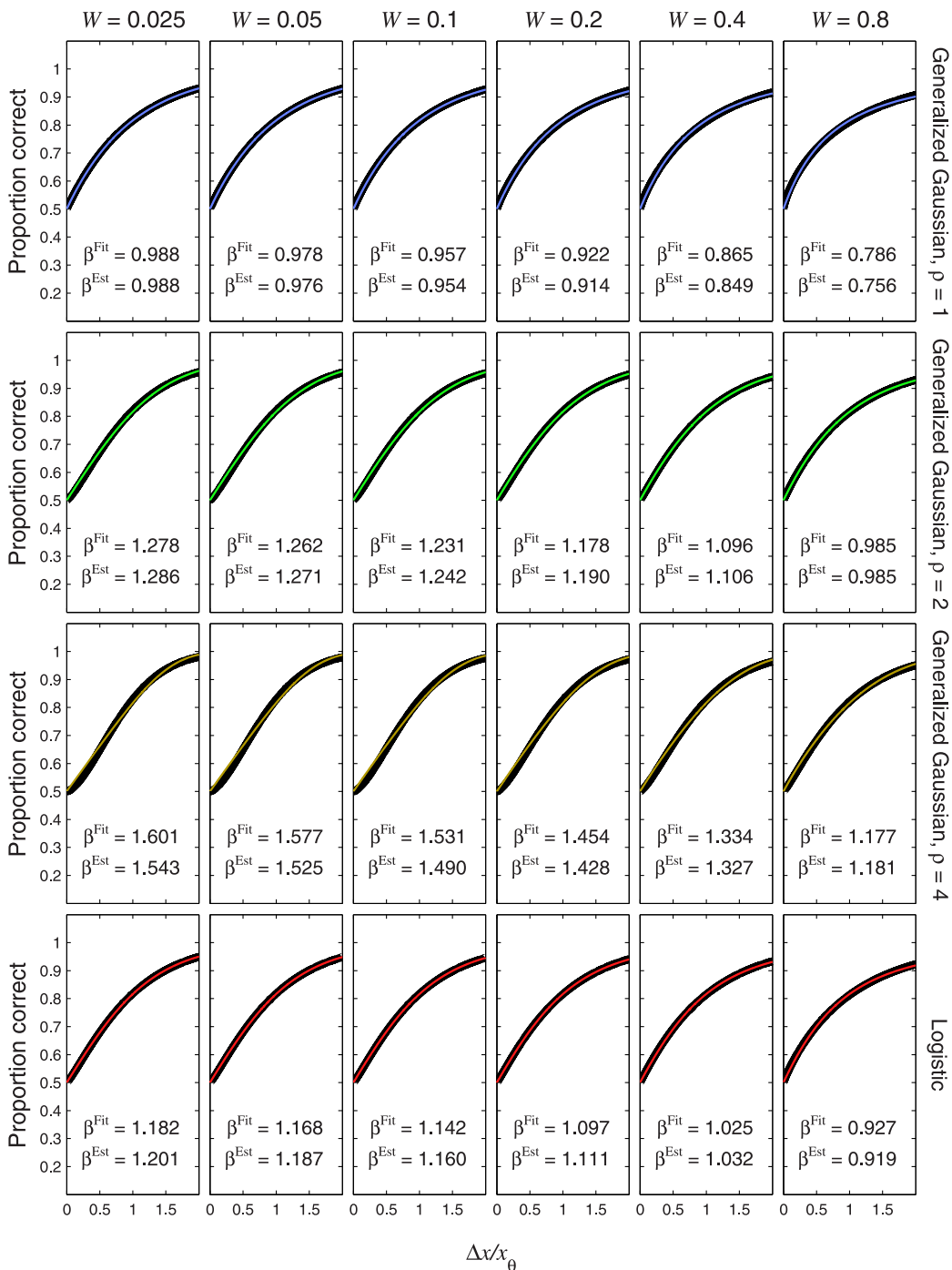
Figure 12. Psychometric functions resulting from a logarithmic transducer. The thin, coloured curves show the psychometric function of Equation (71). Different rows of panels show psychometric functions with different noise CDFs, as indicated on the right of the figure. Different columns of panels show psychometric functions for different Weber fractions, W. The thick, black curves show the best-fitting (maximum-likelihood) Weibull functions. Each panel displays the β value of the best-fitting Weibull function (β^{Fit}) and the estimate, $\beta^{\text{Est}} = \beta^{\text{Noise}} \times \beta^{\text{Transducer}}_{\log}$, where $\beta^{\text{Transducer}}_{\log}$ is given by Equation (70), and β^{Noise} is given by Equation (52), (53), (54) or (60), as appropriate.
doi:10.1371/journal.pone.0074815.g012

in Equation (3) is always positive). The inverse of the lower incomplete gamma function, γ^{-1}, in Equation (48) can be evaluated using the MATLAB function `gammaincinv`. At threshold, $P = P_\theta = 1 - 0.5/e$. Substituting these values into Equation (48), we get

$$F^{-1}_{\text{Gen.Gaussian}}(P_\theta; \rho, \tau) = \tau \left[\gamma^{-1}(1 - 1/e; 1/\rho) \right]^{1/\rho}. \qquad (49)$$

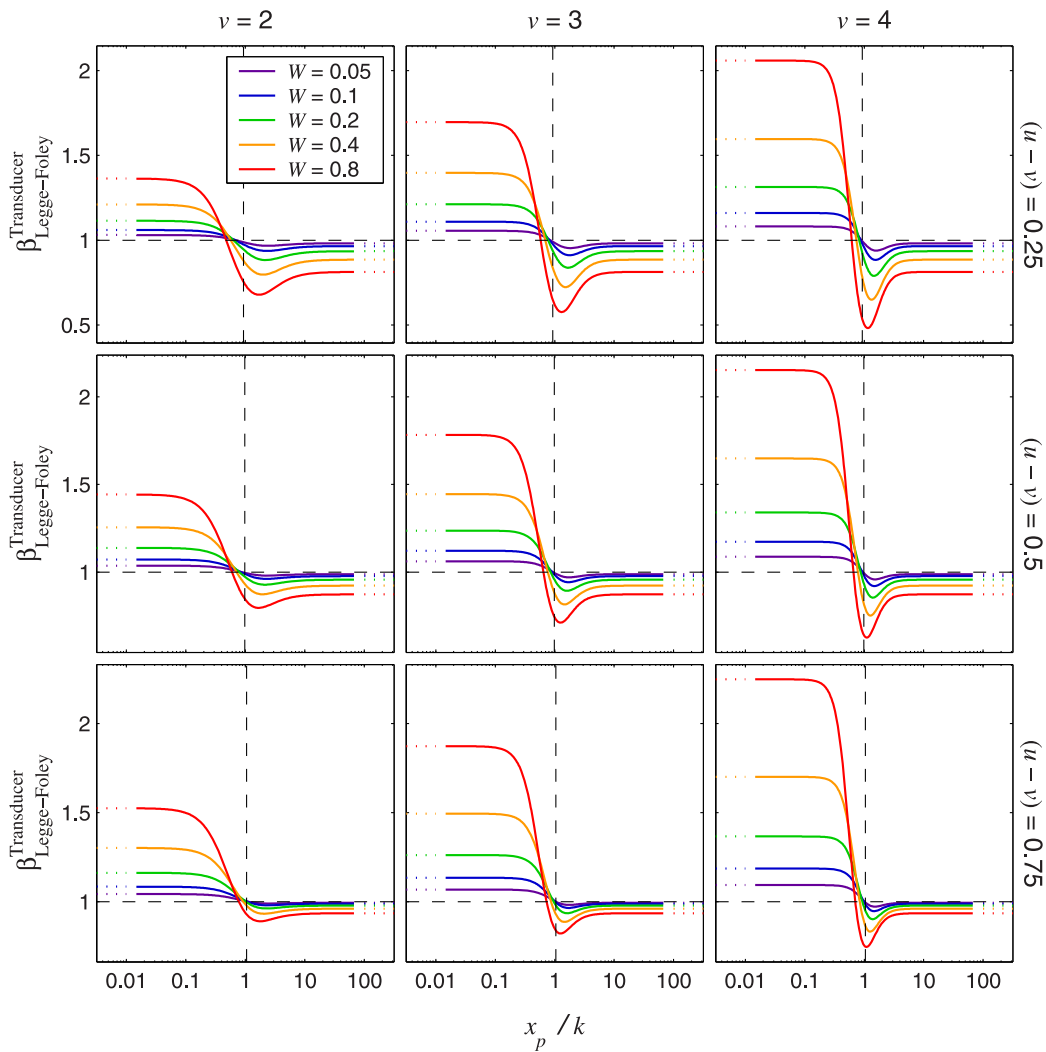We can use the expression for $F^{-1}_{\text{Gen.Gaussian}}(P_\theta; \rho, \tau)$ in Equation

Figure 13. $\beta^{\text{Transducer}}$ for the Legge-Foley transducer with nonzero pedestal. The curves were generated using Equation (74). Each column of panels has a particular value for v, and each row of panels has a particular value for the difference $(u-v)$. Within the panels, the Weber fraction, W, is indicated by the colour of the curve (the legend in the top-left panel applies to all panels). The curves approach horizontal asymptotes on the right (indicated by dotted lines), with vertical position given by Equation (67) with $b=u-v$. This is because, as mentioned earlier, as the input signal increases, the Legge-Foley transducer approaches a power function with exponent $(u-v)$. This asymptote can also be derived from Equation (74) by setting k/x_p to 0, which gives the limit as $x_p \to \infty$. On the left, the curves come close to approaching an asymptote with vertical position given by Equation (67) with $b=u$ because, at low contrasts, the Legge-Foley transducer approximates a power function with exponent u. These near-asymptotes are indicated by dotted lines on the left of each panel. They are not true asymptotes because, even for $x_p = 0$, the Legge-Foley transducer is not exactly equal to a power function over a finite range of inputs. The horizontal, dashed lines indicate $\beta^{\text{Transducer}}_{\text{Legge-Foley}} = 1$. The vertical dashed lines indicate the value of x_p/k corresponding to the point of inflection of the Legge-Foley transducer. An expression for this quantity is derived in Appendix S2. For typical values of u and v, including those in this figure, the point of inflection occurs very close to an input of k, giving $x_p/k \approx 1$. For pedestals above this value, both the target and pedestal will lie in the compressive region of the Legge-Foley transducer, so $\beta^{\text{Transducer}}_{\text{Legge-Foley}}$ must be less than 1. For this reason, none of the curves enter the top-right quadrant in any of the panels.
doi:10.1371/journal.pone.0074815.g013

(49) to substitute for $F^{-1}(P_\theta)$ in Equation (16), and we can use the expression for $f_{\text{Gen.Gaussian}}(\cdot; \rho, \tau)$ in Equation (45) to substitute for $f(\cdot)$ in Equation (16). The different instances of τ cancel out, giving us an expression for β^{Noise} for the generalized Gaussian noise distribution that is a function of ρ:

$$\beta^{\text{Noise}}_{\text{Gen.Gaussian}}(\rho) = \frac{\rho I^{1/\rho}}{\Gamma(1/\rho)} \exp(1-I), \qquad (50)$$

where

$$I = \gamma^{-1}(1 - 1/e; 1/\rho) \qquad (51)$$

The subscript, "Gen.Gaussian", on $\beta^{\text{Noise}}_{\text{Gen.Gaussian}}$ in Equation (50) indicates the general form of the noise CDF.

Figure 6 plots $\beta^{\text{Noise}}_{\text{Gen.Gaussian}}(\rho)$ as a function of ρ. As proved in Appendix S1, $\beta^{\text{Noise}}_{\text{Gen.Gaussian}}(\rho) \to e-1$ as $\rho \to \infty$. Values of $\beta^{\text{Noise}}_{\text{Gen.Gaussian}}(\rho)$ for $\rho = 1$, 2, and 4 are given by
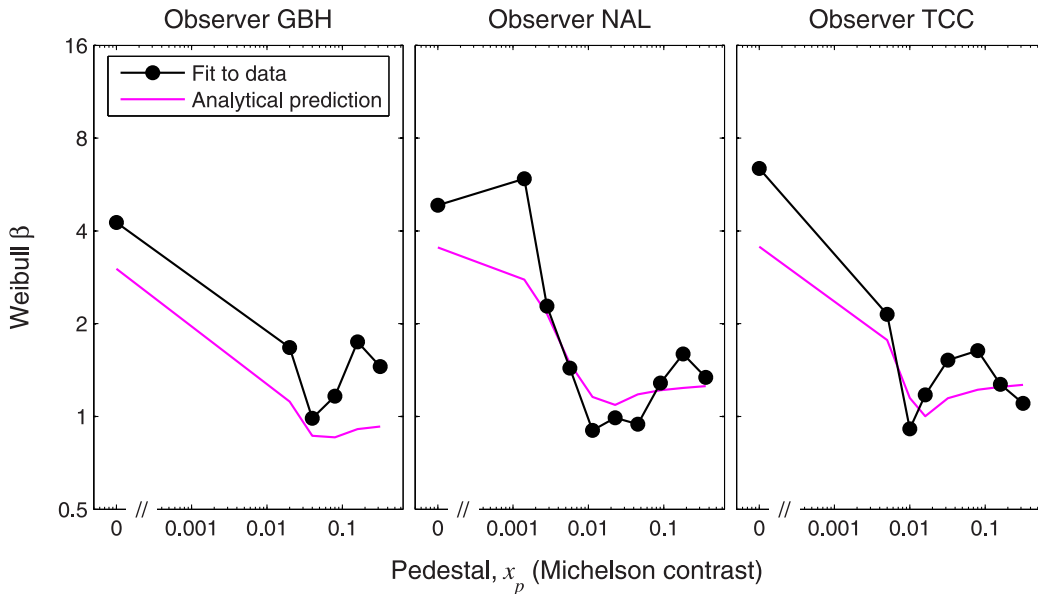
Figure 14. Dipper functions for Weibull β from Henning and Wichmann's data. Weibull β was fitted to Henning and Wichmann's [40] published data as described in the legend of Table 1. These β values are plotted in black lines and symbols, excluding observer GBH's β value of 13.1 for a pedestal contrast of 0.01, which is obviously an outlier. In each case, the function mapping pedestal contrast to β has a dipper shape. To see whether the dip occurred in the predicted location, we fitted a Legge-Foley transducer model to Henning and Wichmann's data separately for each observer. The model's predicted proportion correct, P_{Model}, was given by Equation (6) with the transducer function, R, given by the 4-parameter Legge-Foley transducer (Equation (72)), and the noise CDF, F, given by the generalized Gaussian (Equation (41)), which had ρ as a free parameter, and τ set so that $\sigma_z = 1$, using Equation (44) (thus we adjusted sensitivity by adjusting the transducer gain, rather than the noise CDF spread). For each pedestal value, Henning and Wichmann reported the contrast differences, Δx, corresponding to three different performance levels (proportion correct, $P_{\text{Data}} = 0.6$, 0.75, or 0.9), sampled from their fitted psychometric functions. We performed a maximum-likelihood fit of the Legge-Foley transducer model to the data, by adjusting the parameters to maximize the likelihood, $\sum_{(x_p, \Delta x)} P_{\text{Data}} \log(P_{\text{Model}}) + (1 - P_{\text{Data}})\log(1 - P_{\text{Model}})$. Fitted model parameter sets (u, v, k, r, ρ) were (3.78, 3.38, 0.0322, 15.3, 0.947) for GBH, (3.36, 3.02, 0.00968, 22.7, 2.09) for NAL, and (3.93, 3.51, 0.0102, 18.9, 2.15) for TCC. For each observer and pedestal value, we used a numerical search method to find the threshold, Δx_θ, corresponding to a proportion correct of P_θ, and then calculated the Weber fraction, W, using Equation (34). We then found $\beta_{\text{Legge-Foley}}^{\text{Transducer}}$ using Equation (74), and $\beta_{\text{Gen.Gaussian}}^{\text{Noise}}(\rho)$ using Equation (50). The analytical prediction of Weibull β is then given by $\beta_{\text{Legge-Foley}}^{\text{Transducer}} \times \beta_{\text{Gen.Gaussian}}^{\text{Noise}}(\rho)$, and this is plotted in magenta in the figure. Each observer's global minimum in Weibull β was close to that in the analytical prediction.
doi:10.1371/journal.pone.0074815.g014

$$\beta_{\text{Gen.Gaussian}}^{\text{Noise}}(1) = 1 \qquad (52)$$

$$\beta_{\text{Gen.Gaussian}}^{\text{Noise}}(2) = 1.302 \text{(to 4 significant figures)} \qquad (53)$$

$$\beta_{\text{Gen.Gaussian}}^{\text{Noise}}(4) = 1.562 \text{(to 4 significant figures)}. \qquad (54)$$

The value of β^{Noise} for the Laplace distribution (Equation (52)) is exactly 1. This is because the positive half of its CDF is a Weibull function with $\beta = 1$. This can be seen from the fact that $\gamma(\zeta; 1) = 1 - \exp(-\zeta)$, and so Equation (41) gives, for positive z,

$$F_{\text{Gen.Gaussian}}(z; 1, \tau) = 1 - 0.5 \exp(-z/\tau). \qquad (55)$$

The Weibull function with $\beta = 1$ therefore gives an exact fit to the Laplacian noise CDF, and the estimated Weibull β, given by $\beta^{\text{Noise}} = 1$, is exactly correct.

The coloured curves in Figures 7A, 7B, and 7C show the generalized Gaussian noise CDFs for $\rho = 1$, 2, and 4, respectively, and the thick, black curves show the best-fitting Weibull functions (maximum-likelihood fit over inputs from 0 to twice the threshold).

Also shown in each panel is the appropriate value of β^{Noise} from equations (52) to (54), and the best-fitting Weibull β, which we call β^{Fit}. As explained above, the match between β^{Fit} and β^{Noise} is exact for the Laplace ($\rho = 1$, Figure 7A), but the match is also very good for the other distributions. For the Gaussian ($\rho = 2$, Figure 7B), $\beta^{\text{Fit}} = 1.295$, very close to our analytically derived value of $\beta_{\text{Gen.Gaussian}}^{\text{Noise}}(2) = 1.302$. As discussed earlier, Pelli [19] fitted the Weibull to a Gaussian CDF using a different fitting method: He minimized the maximum error over all positive inputs. The β value he obtained from this fit was 1.247. As noted earlier, there is no single "correct" answer, but our maximum-likelihood fitting paradigm is probably more representative of the process of fitting a function to psychophysical data, and our obtained β of 1.295 is very close to the analytically obtained value. The match between β^{Noise} and β^{Fit} for $\rho = 4$ (Figure 7C) is also close, the deviation being far smaller than the margin of error usually encountered when measuring Weibull β [22–27].

Evaluating β^{Noise} for the logistic noise CDF

Sometimes, the logistic function is used instead of the Gaussian, for computational convenience (e.g. Ref. [28]). The logistic function is very similar in shape to the Gaussian. Its CDF is given by
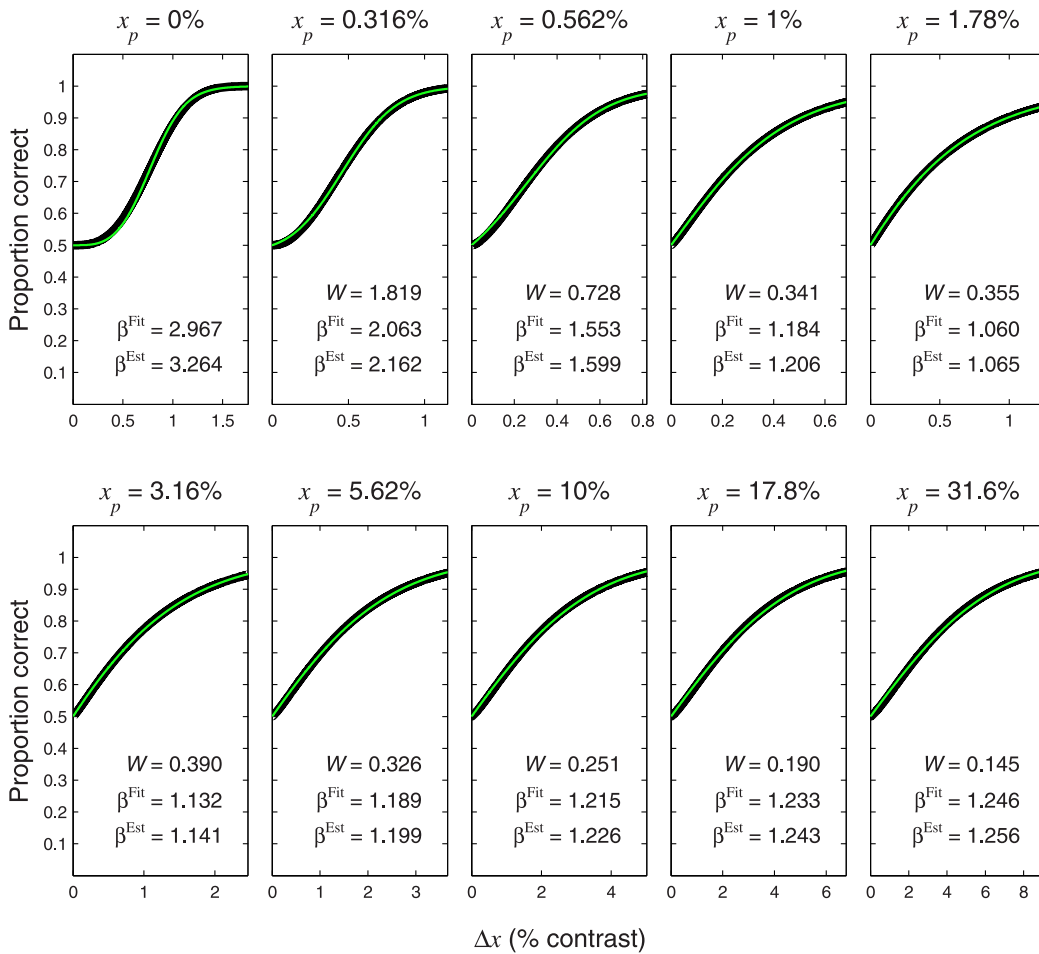
Figure 15. Psychometric functions resulting from a Legge-Foley transducer in Meese et al.'s binocular condition. The thin, green curves show the psychometric functions generated by Meese et al.'s [4] twin-summation model in their binocular condition; in this condition, their transducer is equivalent to the Legge-Foley transducer of Equation (72) with $u = 3.53$, $v = 3.07$, $k = 1.10$, $r = 1.05$. Note k is in units of % contrast, as used by Meese et al.; to convert to units of Michelson contrast, k should be divided by 100. The CDF of the noise on the internal difference signal, z, is a cumulative Gaussian with standard deviation given by $\sigma_z = 0.366$. Each panel gives the model's psychometric function for a different pedestal contrast, x_p, in Meese et al.'s binocular condition. The thick, black curves show the best-fitting (maximum-likelihood) Weibull functions. Each panel displays the β value of the best-fitting Weibull function (β^{Fit}) and the estimate, $\beta^{\text{Est}} = \beta^{\text{Noise}}_{\text{Gen.Gaussian}}(2) \times \beta^{\text{Transducer}}_{\text{Legge-Foley}}$, where $\beta^{\text{Transducer}}_{\text{Legge-Foley}}$ is given by Equation (73) for $x_p = 0$, and by Equation (74) for the other pedestal levels. The Weber fraction, W, in these equations was calculated from the model's threshold, found by inverting the model's psychometric function using a numerical search method, as explained in the text. Because this model fitted well to Meese et al.'s data, these Weber fractions are close to (but not exactly equal to) the actual Weber fractions obtained in the experiment, given in Table 1.
doi:10.1371/journal.pone.0074815.g015

$$P = F_{\text{Logistic}}(z; \tau) = \frac{1}{1 + \exp(-z/\tau)}. \qquad (56)$$

As noted by Strasburger [29], this function is identical to the hyperbolic tangent function, given by $F_{\text{tanh}}(z; \tau) = 0.5\{\tanh[z/(2\tau)] + 1\}$. Its variance is $\tau^2 \pi^2/3$. The PDF of the logistic distribution is given by the derivative of the CDF:

$$f_{\text{Logistic}}(z; \tau) = \frac{\exp(-z/\tau)}{\tau[1 + \exp(-z/\tau)]^2}. \qquad (57)$$

The inverse of the logistic CDF is given by

$$z = F^{-1}_{\text{Logistic}}(P; \tau) = -\tau \ln(1/P - 1), \qquad (58)$$

At threshold, $P = P_\theta = 1 - 0.5/e$. Substituting these values into Equation (58), we get

$$F^{-1}_{\text{Logistic}}(P_\theta; \tau) = \tau \ln(2e - 1). \qquad (59)$$

We can use the expression for $F^{-1}_{\text{Logistic}}(P_\theta; \tau)$ in Equation (59) to substitute for $F^{-1}(P_\theta)$ in Equation (16), and we can use the expression for $f_{\text{Logistic}}(\cdot)$ in Equation (57) to substitute for $f(\cdot)$ in Equation (16). This gives
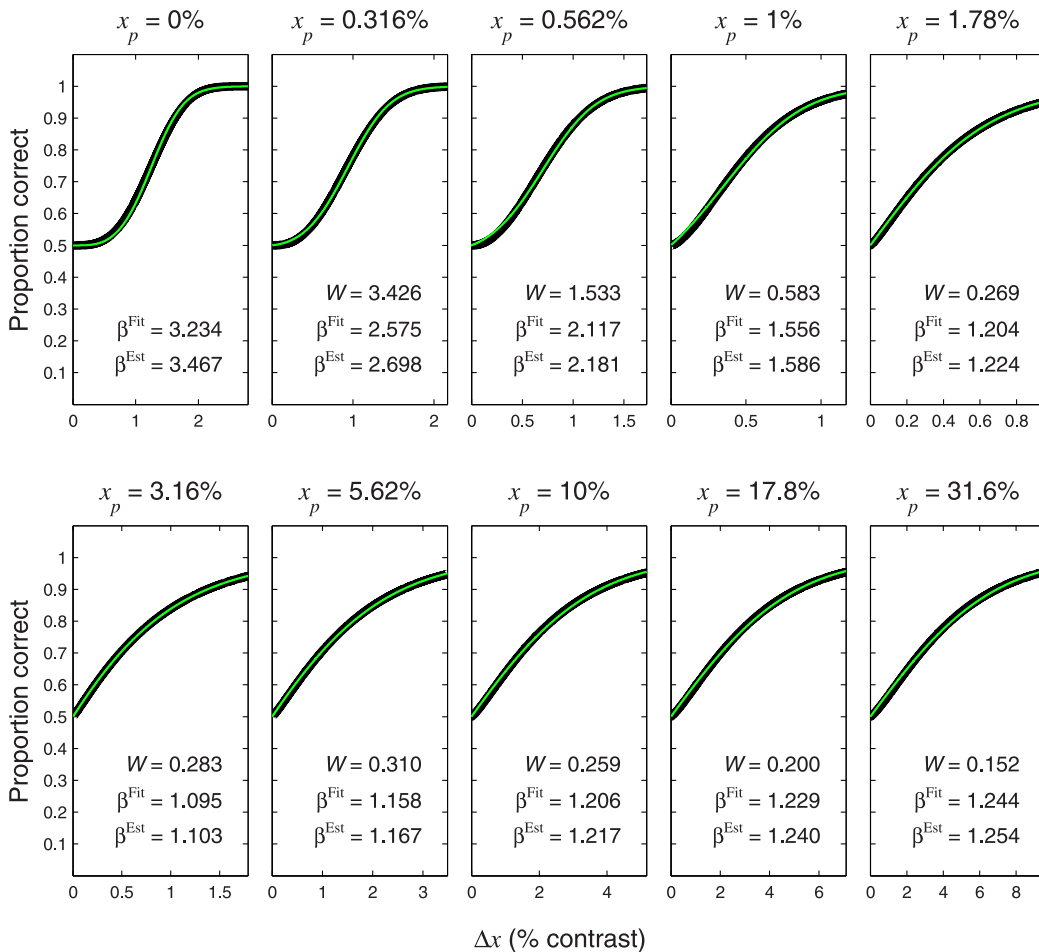
Figure 16. Psychometric functions resulting from a Legge-Foley transducer in Meese et al.'s monocular condition. The same as Figure 15, but for Meese et al.'s monocular condition. In this condition, their twin summation model is equivalent to the Legge-Foley transducer of Equation (72) with the same parameters as those given in the legend to Figure 15, except with $k = 1.89$ and $r = 1$.
doi:10.1371/journal.pone.0074815.g016

$$\beta^{\text{Noise}}_{\text{Logistic}} = \ln(2e-1)[1-1/(2e)]$$

$$= 1.216(\text{to 4 significant figures}). \tag{60}$$

As before, the subscript on β^{Noise} indicates the form of noise CDF. The accuracy of this approximation is confirmed in Figure 7D. The β parameter of the fitted Weibull function ($\beta^{\text{Fit}} = 1.196$) is very close to the estimated value from Equation (60).

Theorem 3. Tendency towards Linear Behaviour with Non-Zero Pedestals

Introduction

As shown earlier, for a linear transducer, $\beta^{\text{Transducer}} = 1$, and so $\beta^{\text{Est}} = \beta^{\text{Noise}}$, which takes a value of around 1.3 for Gaussian internal noise. So, if a transducer model has additive Gaussian noise and generates psychometric functions with a Weibull β of about 1.3, that might seem to suggest that it contains a linear transducer. However, a transducer model with additive Gaussian noise can in fact generate psychometric functions with $\beta \approx 1.3$ for suprathreshold contrast discrimination even when the transducer

departs wildly from a linear function [4]. Theorem 3 explains how this occurs.

Statement of Theorem 3

If the gradient of the transducer is not 0 or ∞ at the pedestal level, then, as $\Delta x_\theta \rightarrow 0$, $\beta^{\text{Transducer}} \rightarrow 1$.

Proof of Theorem 3

As noted earlier, $\beta^{\text{Transducer}} = U/V$, where U and V are given in Equations (38) and (39), respectively. The limit of V as $\Delta x_\theta \rightarrow 0$ is the derivative of R at x_p, i.e. $R'(x_p)$, by definition of the derivative, and the limit of U as $\Delta x_\theta \rightarrow 0$ is obviously $R'(x_p)$, so we have

$$\lim_{\Delta x_\theta \rightarrow 0} U = \lim_{\Delta x_\theta \rightarrow 0} V = R'(x_p). \tag{61}$$

Then, provided that $R'(x_p)$ is not 0 or ∞, we have

$$\lim_{\Delta x_\theta \rightarrow 0} \beta^{\text{Transducer}} = \lim_{\Delta x_\theta \rightarrow 0} \frac{U}{V} = \frac{\lim_{\Delta x_\theta \rightarrow 0} U}{\lim_{\Delta x_\theta \rightarrow 0} V} = \frac{R'(x_p)}{R'(x_p)} = 1. \; \Box \tag{62}$$

If $R'(x_p)$ is 0 or ∞, then $R'(x_p)/R'(x_p)$ is an indeterminate form, $0/0$ or ∞/∞, and cannot be evaluated. In this case, we cannot evaluate the limit of $\beta^{\text{Transducer}}$ by dividing the limit of U by the limit of V. The limit must instead be evaluated in some other way that will depend on the form of the transducer, and the limit in this case will not necessarily be 1.

Discussion of Theorem 3

Theorem 3 shows that, *whatever* the transducer function, as long as its gradient is not 0 or ∞ at the pedestal level, Weibull β will approach that for a linear transducer as sensitivity improves. Virtually all proposed transducers do have a finite, nonzero gradient for nonzero inputs; therefore, if the internal noise is approximately Gaussian, we would expect Weibull β to be close to 1.3 for suprathreshold contrast discrimination. Detection and discrimination data are often fitted with a power-function transducer or a Legge-Foley transducer (both considered below), and, with these transducers, the gradient is 0 or ∞ at an input level of zero. Thus, for these transducers, when the pedestal level is zero, Equation (62) does not apply, and β does not necessarily approach that for a linear transducer as sensitivity improves. This explains why, for contrast detection experiments (i.e. when the pedestal is zero), Weibull β has been found to deviate greatly from the value of 1.3 expected from a linear transducer with Gaussian noise.

Consider what happens in general when the pedestal approaches zero. If we assume that $R(0)=0$, then, as x_p drops below Δx_θ, both U (Equation (38)) and V (Equation (39)) become dominated by the Δx_θ term, and $\beta^{\text{Transducer}}$ approaches the value given in Equation (36), which is not, in general, equal to 1. Thus, we would expect Weibull β to deviate substantially from the linear case for low pedestals. Meese, Georgeson and Baker [4] showed that this is indeed the case for visual contrast discrimination, and we examine their work in more detail later, in the section on the Legge-Foley transducer.

Deriving Psychometric Functions and Weibull β for Specific Nonlinear Transducers

As shown earlier, $\beta^{\text{Transducer}}=1$ for any linear transducer. For a nonlinear transducer, $\beta^{\text{Transducer}}$ will deviate from 1, and this is how the transducer has its effect on β^{Est}, the estimated Weibull β. Starting with one of the general expressions for $\beta^{\text{Transducer}}$ (Equation (17), (35) or (36), as appropriate), we can substitute a specific transducer function for the general function, R, to give a specific expression that describes $\beta^{\text{Transducer}}$ for that transducer. Similarly, starting with one of the general expressions for the psychometric function (Equation (11) or (14), as appropriate), we can substitute a specific transducer function for the general function, R, to give a specific expression for the psychometric function. In this section, we consider five commonly used scenarios: a power function with zero or nonzero pedestal, a log function, and a Legge-Foley function [30] with zero or nonzero pedestal.

Power-function transducers have been used to account for visual contrast discrimination data. As the pedestal increases from 0, the discrimination threshold first decreases, and then starts to increase with further increases in pedestal; this function, giving contrast discrimination threshold at each pedestal level, is known as a "dipper function". The initial dip can be explained by an expansive power function (i.e. one with exponent greater than 1) at low contrasts [1,6], while the increase in contrast discrimination threshold for larger pedestals can be explained by a compressive

power function (i.e., one with exponent less than 1) at high contrasts. The Legge-Foley transducer approximates an expansive power function at low contrasts and a compressive power function at high contrasts, and accounts for the whole dipper function [4,30]. We also include the log transducer in our analysis, firstly because discrimination at high pedestal levels has often been found to adhere closely to Weber's law in many different perceptual dimensions and sensory modalities [31–35] (a prediction of the log transducer with additive noise), and, secondly, because we have discovered an interesting link between the power function and the log transducer, which is presented in Theorem 4.

Power function and zero pedestal

The first case that we consider is the one examined by Pelli [19], i.e. $d' \propto (\Delta x)^b$. As noted earlier, this relationship between d' and Δx is consistent with a power function transducer ($R(x)=rx^b$) and zero pedestal ($x_p=0$). In this case, we can use Equation (36) to derive $\beta^{\text{Transducer}}$, and it follows easily that

$$\beta^{\text{Transducer}}_{\text{Power func}, x_p=0} = b \qquad (63)$$

(as with β^{Noise}, the subscript on $\beta^{\text{Transducer}}$ describes the specific case). Thus, using Equation (63) to substitute for $\beta^{\text{Transducer}}$ in Equation (15), we have

$$\beta^{\text{Est}} = \beta^{\text{Noise}} \times b, \qquad (64)$$

which is the relationship derived by Pelli [19] (Relation (7) of this paper).

From Equation (14), it follows that, for a power function transducer and zero pedestal, the model's true psychometric function is given by

$$\Psi_{\text{Power func}, x_p=0}(\Delta x) = F\left(F^{-1}(P_\theta) \left(\frac{\Delta x}{\Delta x_\theta} \right)^b \right) \qquad (65)$$

with the subscript on Ψ describing the specific case. Equation (65) gives us the option of expressing the stimulus difference in absolute units, Δx, or in "threshold units", $\Delta x/\Delta x_\theta$ – the two options differ only in a linear horizontal scaling. The latter is useful when dealing with general cases where the threshold is not specified; the psychometric function is often expressed in this way [19,29,36]. The coloured curves in Figure 8 show the psychometric function of Equation (65). Different rows of panels show psychometric functions for different noise CDFs, F, as indicated on the right of the figure. Different columns of panels show psychometric functions for different transducer exponents, b. The thick, black curves show the best-fitting Weibull functions. These provide a good fit to the true psychometric functions, justifying the premise of Theorem 2, which is that the Weibull function provides a good fit.

Each panel of Figure 8 also compares β of the best-fitting Weibull function (β^{Fit}) with the estimate, β^{Est}, given by $\beta^{\text{Est}} = \beta^{\text{Noise}} \times b$. In every case, β^{Est} is very close to the fitted value, the discrepancy being far smaller than the margin of error usually encountered in psychophysical measurements of psychometric function slope [22–27]. For each transducer (i.e. each column of Figure 8), the difference in β^{Est} between the different noise CDFs (i.e. between the different rows of Figure 8) is caused entirely by the different values of β^{Noise}. For example, the value of
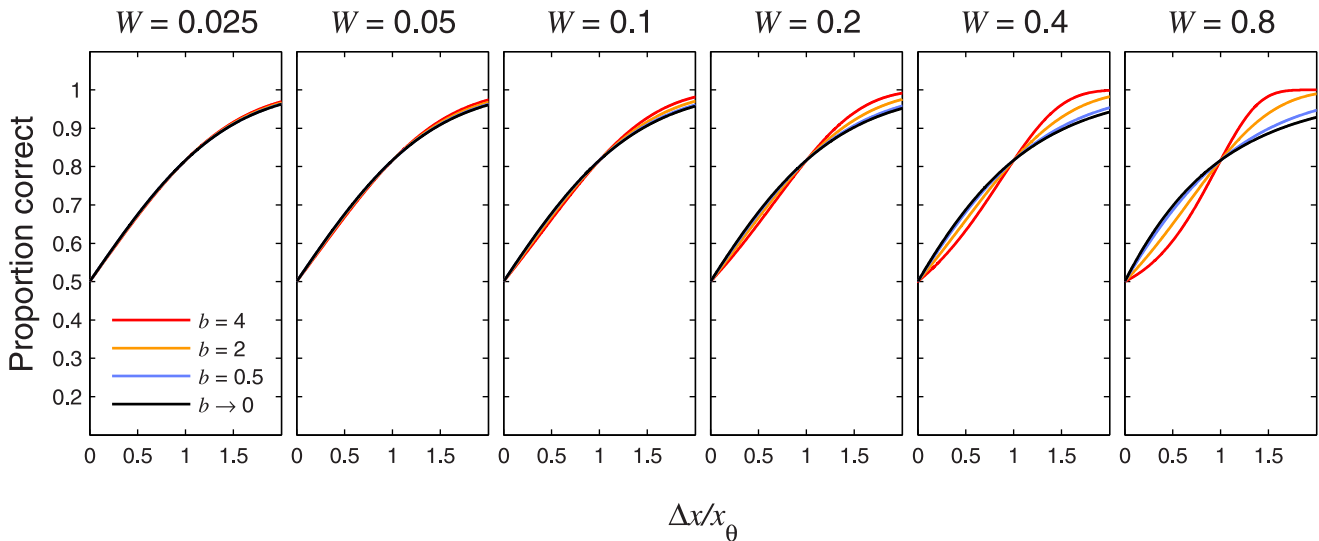
Figure 17. Psychometric functions for the power-function transducer with nonzero pedestal. The psychometric function for $b \to 0$ was generated using Equation (71); the others were generated using Equation (69). In both cases, we assumed Gaussian internal noise (i.e. F is the cumulative Gaussian). All the psychometric functions go through the point $(1, P_\theta)$, by definition of the threshold (the abscissa is in threshold units, i.e. $\Delta x / \Delta x_\theta$). Each panel shows psychometric functions for a particular Weber fraction. Each curve within a panel shows the psychometric function for a particular transducer exponent, b. The orange curve ($b = 2$) is the psychometric function plotted in green in the second-to-top row of Figure 10. The blue curve ($b = 0.5$) is the psychometric function plotted in green in the second-to-top row of Figure 11. The black line shows the limit as $b \to 0$. As proved in Theorem 4A, this limiting case is identical to the psychometric function for a log transducer. This is the psychometric function plotted in green in the second-to-top row of Figure 12. This figure illustrates two effects. Within each panel, we see how the psychometric function for the power-function transducer converges towards that for a log transducer as the exponent decreases (Theorem 4). Across panels (right-to-left), we see a demonstration of the effect proved in Theorem 3, whereby, with a nonzero pedestal, all psychometric functions converge towards that for a linear transducer as the discrimination threshold decreases (in this case, since we are plotting psychometric functions for Gaussian noise, the functions converge towards the pure noise CDF of Figure 4B as the Weber fraction decreases).
doi:10.1371/journal.pone.0074815.g017

β^{Est} for the Gaussian will always exceed that for the Laplace by a factor $\beta^{\text{Noise}}_{\text{Gen.Gaussian}}(2) \big/ \beta^{\text{Noise}}_{\text{Gen.Gaussian}}(1) = 1.302$.

Power function and nonzero pedestal

We now consider the case of a power function and any pedestal value, a generalization of the previous case. First, starting with Equation (17), we trivially obtain

$$\beta^{\text{Transducer}}_{\text{Power func}} = \frac{b(x_p + \Delta x_\theta)^{b-1} \Delta x_\theta}{(x_p + \Delta x_\theta)^b - x_p^b}, \quad (66)$$

which reduces to Equation (63) when $x_p = 0$. When $x_p \neq 0$, we can start with Equation (35), from which it follows straightforwardly that

$$\beta^{\text{Transducer}}_{\text{Power func, } x_p \neq 0} = \frac{b(1 + W)^{b-1} W}{(1 + W)^b - 1}. \quad (67)$$

Figure 9 plots $\beta^{\text{Transducer}}_{\text{Power func, } x_p \neq 0}$ as a function of the Weber fraction, W (defined in Equation (34)), for several values of the transducer exponent, b. These curves all converge to a value of 1 towards the left. This is because, for a power function transducer with nonzero pedestal, the gradient of the transducer at the pedestal level is not 0 or ∞, and so, as proved in Theorem 3, $\beta^{\text{Transducer}}_{\text{Power func, } x_p \neq 0} \to 1$ as $W \to 0$.

From Equation (11), we can see that, for a power function transducer with unspecified pedestal, the model's true psychomet-

ric function is given by

$$\Psi_{\text{Power func}}(\Delta x) = F\left(F^{-1}(P_\theta) \frac{(x_p + \Delta x)^b - x_p^b}{(x_p + \Delta x_\theta)^b - x_p^b} \right). \quad (68)$$

If $x_p \neq 0$, then we can divide through by x_p^b, and rewrite Equation (68) in terms of the Weber fraction, W:

$$\Psi_{\text{Power func, } x_p \neq 0}(\Delta x) =$$
$$F\left(F^{-1}(P_\theta) \frac{(1 + W(\Delta x / \Delta x_\theta))^b - 1}{(1 + W)^b - 1} \right). \quad (69)$$

Equation (68) is a general formula for the psychometric function, given a power-function transducer and any pedestal value. Equations (65) and (69) are simpler expressions for this psychometric function in the cases of zero and nonzero pedestals, respectively. As with Equation (65), Equation (69) gives us the option of expressing the stimulus difference in absolute units, Δx, or threshold units, $\Delta x / \Delta x_\theta$. The coloured curves in Figure 10 show the psychometric function of Equation (69) for a range of Weber fractions and a transducer exponent of 2. Different rows of panels show psychometric functions for different noise CDFs, and different columns of panels show psychometric functions for different Weber fractions, W. The thick, black curves show the best-fitting Weibull functions. Each panel also compares β of the best-fitting Weibull function (β^{Fit}) with the estimate, β^{Est}, given by $\beta^{\text{Est}} = \beta^{\text{Noise}} \times \beta^{\text{Transducer}}_{\text{Power func, } x_p \neq 0}$. Figure 11 is the same as Figure 10 except that the transducer exponent is 0.5. In every case, the
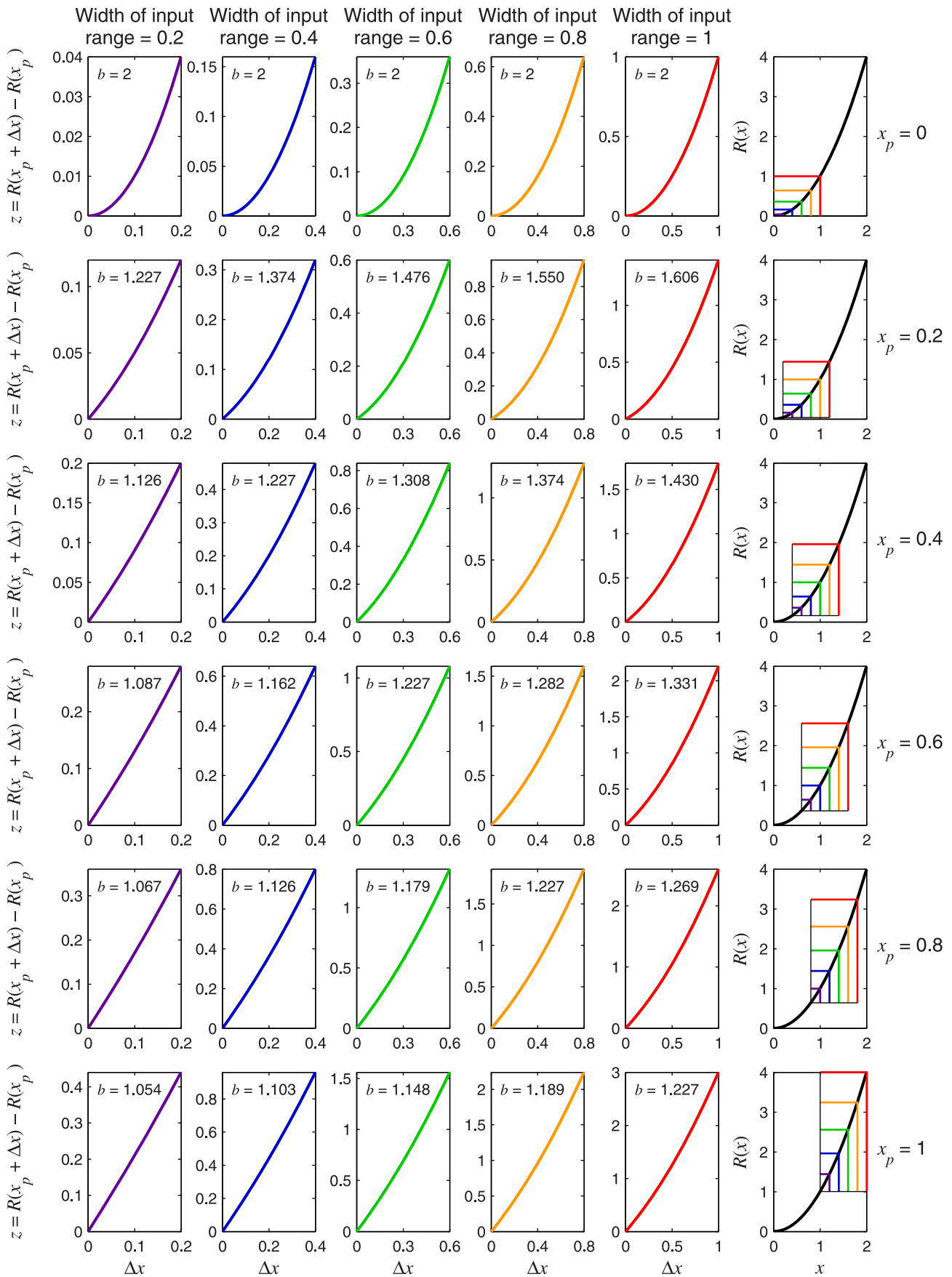
Figure 18. Effect of a pedestal on the linearity of an expansive power-function transducer. Each panel in the rightmost column shows the same expansive power-function transducer given by $R(x) = x^2$. The panels to the left show parts of this transducer sampled over different ranges of inputs: The width of the range is varied across columns of panels, and the lower limit of the range is varied across rows of panels. The lower limit would correspond to the pedestal value, x_p, in a discrimination experiment. The abscissa of the curves on the left is the stimulus difference, Δx, and the ordinate is z, the difference in internal signal values after transduction. The b-value given in each panel is the exponent of the power function that fits best (least squares) to these curves. Each coloured box drawn on a transducer in the right column indicates the part of the transducer that is sampled by the correspondingly coloured curve given in a panel to the left on the same row. It can be seen that, as the pedestal increases, the best-fitting exponent quickly approaches 1, giving an approximately linear mapping from Δx to z. This linearizing effect is enhanced as the width of the range decreases. x, x_p, and Δx are given in arbitrary units: For a given transducer, the best-fitting exponent is determined by the *ratio* of the pedestal value to the width of the input range. For example, with the transducer shown here, when the pedestal value is equal to the width of the range, the best-fitting exponent is always 1.227; when the pedestal is twice the width of the range, the best-fitting exponent is always 1.126. For a zero pedestal (top row), the best-fitting exponent is always 2, regardless of the width of the input range, and in this sense the power function is not "strongly locally linear" at $x = 0$.

Weibull function fits well to the true psychometric function, and the agreement between β^{Fit} and β^{Est} is very good.

It is interesting to compare the psychometric functions in Figures 10 and 11 with those for the same transducer when the pedestal is zero; these are given in the columns of Figure 8 headed "$b = 2$" and "$b = 0.5$", respectively. It is clear that, even with the rather large Weber fraction of 0.8, the existence of a nonzero pedestal brings Weibull β much closer to the linear case (the case of a linear transducer is shown in the column of Figure 8 headed "$b = 1$").

Logarithmic transducer

The log function is undefined for zero inputs, so we can only consider a log transducer for nonzero pedestals. If the transducer takes a logarithmic shape for all inputs greater than the pedestal value, then it is effectively logarithmic for the whole of the range of stimulus values being considered.

For a logarithmic transducer, $R(x) = r\log_a(x)$, Equation (35) leads to

$$\beta_{\log}^{\text{Transducer}} = \frac{W}{(1 + W)\ln(1 + W)} \qquad (70)$$

for any base of logarithm, a. In Theorem 4B, below, we show that, for any Weber fraction, $\beta_{\log}^{\text{Transducer}}$ is the limiting value of $\beta_{\text{Power func, } xp \neq 0}^{\text{Transducer}}$ as $b \to 0$. The bottom (black) curve in Figure 9 plots $\beta_{\log}^{\text{Transducer}}$ as a function of W.

Starting with Equation (11), it is straightforward to show that, for a log transducer, the model's true psychometric function is given by

$$\Psi_{\log}(\Delta x) = F\left(F^{-1}(P_\theta)\frac{\ln(1 + W(\Delta x/\Delta x_\theta))}{\ln(1 + W)}\right). \qquad (71)$$

The coloured curves in Figure 12 show the psychometric function of Equation (71) for a range of Weber fractions. Different rows of panels show psychometric functions for different noise CDFs, and different columns of panels show psychometric functions for different Weber fractions. The thick, black curves show the best-fitting Weibull functions. Each panel also compares β of the best-fitting Weibull function (β^{Fit}) with the estimate, β^{Est}, given by $\beta^{\text{Est}} = \beta^{\text{Noise}} \times \beta_{\log}^{\text{Transducer}}$. The Weibull function gives an excellent fit to the true psychometric function in every case, and the agreement between β^{Fit} and β^{Est} is very good. As the Weber fraction decreases, Weibull β approaches that for the linear case.

Legge-Foley transducer

If the noise is additive, then no single power-function transducer can fit contrast discrimination data across the whole contrast range, because we need an expansive function to explain facilitation at low contrasts, and a compressive function to explain the rise in threshold with pedestal at high contrasts. Legge and Foley [30] used a sigmoid transducer that was expansive at low contrasts and compressive at high contrasts, as required:

$$R_{\text{Legge-Foley}}(x) = \frac{rx^u}{k^v + x^v}, \qquad (72)$$

where u, v, r and k are constants, all greater than zero. The transducer in Equation (72) seems to have been first used in psychophysics by Stromeyer and Klein [37], and it is sometimes referred to as the Stromeyer-Foley function [36,38], but we use the term "Legge-Foley transducer", as Legge and Foley's use of this transducer is probably better known. For low inputs, x, the Legge-Foley transducer approximates a power function with exponent u; for large inputs, it approximates a power function with exponent $(u - v)$. Legge and Foley had $u = 2.4$ and $v = 2$, so the transducer was an accelerating power function (with exponent ≈ 2.4) for low inputs, and a compressive power function (with exponent ≈ 0.4) for high inputs. The point of inflection (at which the transducer changes from expansive to compressive) occurs at an x value close to k for typical values of the fitted parameters (see Appendix S2 for a derivation of the formula for calculating the position of the point of inflection).

Assuming a zero pedestal, we can substitute $R_{\text{Legge-Foley}}$ for R in Equation (36), giving

$$\beta_{\text{Legge-Foley}}^{\text{Transducer}} = \frac{u + (u - v)(\Delta x_\theta/k)^v}{1 + (\Delta x_\theta/k)^v} \text{ if } x_p = 0. \qquad (73)$$

If the threshold, Δx_θ, is much less than k, the right hand side of Equation (73) approaches u, as we would expect: The Legge-Foley transducer in this case approximates a power-function transducer with exponent u, and, for the latter transducer with zero pedestal, $\beta^{\text{Transducer}}$ is simply equal to the exponent, as in Equation (63).

For nonzero pedestals, we can substitute $R_{\text{Legge-Foley}}$ for R in Equation (17) to obtain, after some work,

$$\beta_{\text{Legge-Foley}}^{\text{Transducer}} = \frac{\frac{uW}{1 + W} - \frac{vW(1 + W)^{v - 1}}{(k/x_p)^v + (1 + W)^v}}{1 - \frac{(k/x_p)^v + (1 + W)^v}{[1 + (k/x_p)^v](1 + W)^u}} \text{ if } x_p \neq 0. \qquad (74)$$

Figure 13 plots $\beta_{\text{Legge-Foley}}^{\text{Transducer}}$ as defined in Equation (74) as a function of x_p/k, for typical ranges of u, v, and W. The middle
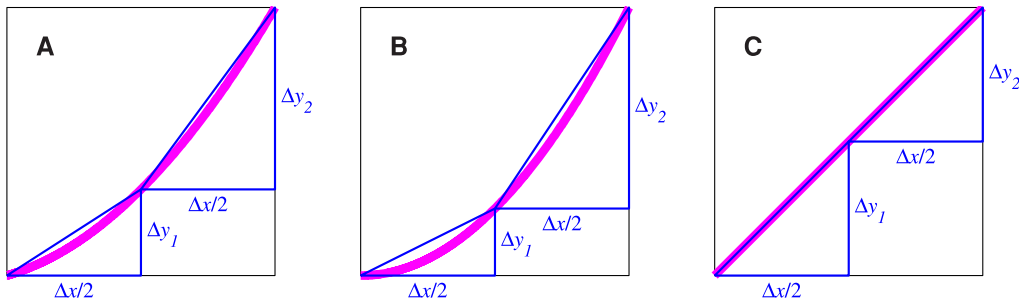
Figure 19. Index of acceleration, η. (A) The wide, magenta curve shows an expansive power function sampled over a range of inputs from a to $(a+\Delta x)$, where $a>0$. The horizontal blue lines both have length $\Delta x/2$, and the vertical blue lines have length Δy_1 and Δy_2 as indicated. The slope, S_1, of the secant (the oblique line) across the left half of the curve is given by $\Delta y_1/(\Delta x/2)$, and the slope, S_2, of the secant across the right half of the curve is given by $\Delta y_2/(\Delta x/2)$. Our index of acceleration, η, is given by S_2/S_1. For the power function, when $a \neq 0$, $\eta \to 1$ as $\Delta x \to 0$, so the curve is "strongly locally linear" at $x=a$. (B) The same as A, but with the bottom of the range of inputs, a, equal to zero. In this case, η depends only on the exponent of the power function, and so it does not approach 1 as Δx approaches zero. The power function is not "strongly locally linear" at $x=0$. (C) The same as A, but for a straight line function. Here, $\eta = 1$ for all a and Δx.
doi:10.1371/journal.pone.0074815.g019

panel of the left column ($v=2$, $u-v=0.5$) is very close to Legge and Foley's [30] parameters, while the middle panel of the middle column ($v=3$, $u-v=0.5$) is very close to Meese et al.'s [4] fitted parameters, which we describe in detail later.

One striking feature of these functions is that they all have a dipper shape – as the pedestal increases, $\beta_{\text{Legge-Foley}}^{\text{Transducer}}$ dips down to a minimum and then increases slightly before approaching its asymptote on the right. While it is well known that the discrimination threshold, Weibull α, traces out a dipper function as x_p increases from zero [4,6,30,34,39], to the best of our knowledge no one has ever reported a dipper function for Weibull β before, so we set out to see if there was evidence for one in the previous literature. We found such a dipper function for β in the data of Henning and Wichmann [40] (see Figure 14 and Table 1). Henning and Wichmann did not report Weibull β, but they reported all thresholds at three different performance levels, allowing us to fit Weibull functions to their data.

For the log and power-function transducers, $\beta^{\text{Transducer}}$ is a function with one or two arguments, so it was practical to test the accuracy of the expressions for a range of plausible arguments. In contrast, $\beta_{\text{Legge-Foley}}^{\text{Transducer}}$ has five arguments, corresponding to three of the transducer parameters, as well as the pedestal and either the Weber fraction or the threshold. To constrain the argument space so that we can test the accuracy of Equations (73) and (74), it is helpful to use values for these arguments that have occurred in real experiments. Legge and Foley's study is not suitable for this because their model had several channels, and did not have the straightforward relationship between stimulus and probability of a correct response described by Equation (6). However, the psychometric function generated by Meese et al.'s [4] preferred model really is a parameterization of Equation (6), so we can assess the accuracy of Equations (73) and (74) for their stimulus values and transducer parameters.

Meese et al.'s study was on binocular integration, and their data were best fit by a model that they called the "twin summation" model. This model has a transducer that extends Legge and Foley's transducer so that it can handle inputs from left and right eyes:

$$R_{\text{Twin summation}} = \frac{(x_L{}^m + x_R{}^m)^p}{Z + (x_L{}^n + x_R{}^n)^q}, \qquad (75)$$

where x_L is the stimulus contrast in the left eye, and x_R is the contrast in the right eye. In Equation (75), we use upper-case Z in place of the lower-case z that Meese et al. used, to avoid confusion with our own z, defined in Equation (3). In Meese et al.'s fully binocular condition ($x_L = x_R = x$), Equation (75) reduces to the standard Legge-Foley transducer of Equation (72), with

$$u = mp, \qquad (76)$$

$$v = nq, \qquad (77)$$

$$k = (2^{-q}Z)^{1/v}, \qquad (78)$$

$$r = 2^{p-q}. \qquad (79)$$

In Meese et al.'s fully monocular condition ($x_L = 0$ and $x_R = x$ or vice-versa), Equation (75) reduces to the Legge-Foley transducer with u and v defined as in Equations (76) and (77), but with the other parameters given by

$$k = Z^{1/v}, \qquad (80)$$

$$r = 1. \qquad (81)$$

The fitted values of m, n, p, q, and Z appear in the bottom line of Meese et al.'s Table 2. Using these values, we can specify the equivalent Legge-Foley transducer in the binocular or monocular conditions using Equation (72) with u and v given by Equations (76) and (77), and k and r given by Equations (78) and (79) for the binocular condition, and by Equations (80) and (81) in the monocular condition. The values of u, v, k and r are given in the legend to our Figure 15 for the binocular condition, and Figure 16 for the monocular condition. In summary, although Meese et al. fitted a single transducer function across all their conditions, the equivalent Legge-Foley transducer differs between the binocular and monocular conditions. In both cases, $u=3.53$ and $v=3.07$, giving $(u-v)=0.460$, so, although the effective exponent at low

contrasts was substantially higher than that of Legge and Foley [30], the effective exponent at high contrasts was similar.

Meese et al. had one further parameter, the standard deviation, σ_R, of the noise on the transducer output, which took a fitted value of 0.259. Meese et al. assumed uncorrelated zero-mean Gaussian noise on each signal, so the two sources of noise on each trial would combine to produce Gaussian noise on the internal difference signal with standard deviation given by $\sigma_z = \sigma_R\sqrt{2}$. We can therefore define the noise CDF, F, as the cumulative Gaussian with standard deviation $\sigma_R\sqrt{2}$. This is equivalent to Equation (45) with $\rho = 2$ and $\tau = \sigma_z\sqrt{2} = 2\sigma_R$.

After we have defined the transducer, R, and the noise CDF, F, the psychometric function (i.e., the mapping from $(x_p, \Delta x)$ onto probability correct, P) is fully defined by Equation (6). For the Legge-Foley transducer, the psychometric function cannot be inverted algebraically, but the fitted model's threshold, Δx_θ, can be found by searching for the contrast difference, Δx, that gives rise to a probability correct of P_θ. The Weber fraction, W, is then given by Equation (34), and the obtained values of Δx_θ and W can be used in Equations (73) or (74) (depending on whether or not $x_p = 0$), along with u, v, and k, to calculate a value for $\beta_{\text{Legge-Foley}}^{\text{Transducer}}$ for each pedestal level.

Each panel of Figure 15 shows (in green) the psychometric function generated by Meese et al's twin summation model for a particular pedestal level in their binocular condition. The thick, black curves show the best-fitting Weibull functions. Each panel also compares β of the best-fitting Weibull function (β^{Fit}) with the estimate, β^{Est}, given by $\beta^{\text{Est}} = \beta_{\text{Gen.Gaussian}}^{\text{Noise}}(2) \times \beta_{\text{Legge-Foley}}^{\text{Transducer}}$, with $\beta_{\text{Gen.Gaussian}}^{\text{Noise}}(2)$ given by Equation (53) (the model's Weber fractions, W, which are used to calculate $\beta_{\text{Legge-Foley}}^{\text{Transducer}}$, are given in the individual panels). The Weibull fits are good, and the agreement between β^{Fit} and β^{Est} is mostly excellent. Note how Weibull β begins to deviate substantially from the linear value (1.3) as the pedestal drops to very low levels, as explained at the end of the Discussion of Theorem 3. Note, too, how both β^{Fit} and β^{Est} show a dipper function, with the lowest value falling at a pedestal level just above k ($x_p = 1.78\%$); the finding of a dipper function for β^{Fit} in close agreement with that of β^{Est} verifies that the predicted dipper function for β shown in Figures 13 and 14 is a real prediction of the Legge-Foley transducer, and not just a peculiarity of our analytical approximation. Unlike the data of Henning and Wichmann [40], Meese et al.'s [4] data (given in Table 1) do not actually show a dipper for β, but this is not a serious concern because, as mentioned earlier, Weibull β is difficult to measure accurately, and such a small effect on β could easily be lost in the experimental noise. Figure 16 shows the same analysis for Meese et al's monocular condition.

Theorem 4. For Nonzero Pedestals, 2AFC Performance for a Power-Function Transducer Approaches That for a Log Transducer as the Exponent Approaches Zero

Introduction

We stated earlier that, for a nonzero pedestal, as the exponent of a power function transducer approaches zero, $\beta^{\text{Transducer}}$ approaches that for a logarithmic transducer, whatever the Weber fraction. This seems a remarkable finding, because the expressions for $\beta_{\text{Power func}, x_p \neq 0}^{\text{Transducer}}$ and $\beta_{\log}^{\text{Transducer}}$ (given by Equations (67) and

(70), respectively) appear quite different. In fact, for a given threshold level, the whole 2AFC psychometric function for the power-function transducer and nonzero pedestal (given by Equation (69)) approaches that for a log transducer (Equation (71)) as the exponent, b, in Equation (69) approaches zero.

Both of these results stem from a more fundamental result: As the exponent, b, of a power function approaches zero, the function converges towards a log function plus a constant. Because 2AFC performance in the transducer model is based on the *difference* of transducer outputs, this constant cancels out, and, in the limit as $b \rightarrow 0$, the difference of power functions equals the difference of log functions.

This can be understood from Lemma 1, below.

Lemma 1.

$$\lim_{b \to 0} \frac{x^b - 1}{b} = \ln(x). \tag{82}$$

Proof. To begin with, note that, for $b \neq 0$,

$$\int_1^x t^{b-1} dt = \left[\frac{t^b}{b}\right]_1^x = \frac{x^b - 1}{b}. \tag{83}$$

Therefore,

$$\lim_{b \to 0} \frac{x^b - 1}{b} = \lim_{b \to 0} \int_1^x t^{b-1} dt = \int_1^x t^{-1} dt = \ln(x). \square$$

From Lemma 1, we can see that, when b is small,

$$\frac{x^b}{b} \approx \ln(x) + \frac{1}{b}, \tag{84}$$

which explains why the power function converges towards a log function plus a constant, $1/b$, as the exponent, b, approaches zero. When the transducer outputs are subtracted to make the decision in a 2AFC task, this constant cancels out, so we have

$$\frac{x_1^b}{b} - \frac{x_2^b}{b} \approx \ln(x_1) - \ln(x_2), \tag{85}$$

and this is why 2AFC performance for the power function approaches that for a log function as the transducer exponent approaches 0. This does not apply to a zero pedestal because the log function is undefined for zero input.

For those readers who find this informal argument unconvincing, we now prove directly that the two psychometric functions are identical (Theorem 4A) and that the $\beta^{\text{Transducer}}$ values in the two cases are also identical (Theorem 4B).

Theorem 4A

Statement of Theorem 4A. $\Psi_{\text{Power func}, x_p \neq 0}(\Delta x)$, defined in Equation (69), approaches $\Psi_{\log}(\Delta x)$, defined in Equation (71), as the power-function exponent, b, in Equation (69) tends to zero.

Proof. From Equation (69),

$$\Psi_{\text{Power func}, x_p \neq 0}(\Delta x)$$

$$= F\left(F^{-1}(P_\theta)\frac{(1 + W(\Delta x/\Delta x_\theta))^b - 1}{(1 + W)^b - 1}\right) \qquad (86)$$

$$= F\left(F^{-1}(P_\theta)\frac{\left[(1 + W(\Delta x/\Delta x_\theta))^b - 1\right]\big/b}{\left[(1 + W)^b - 1\right]\big/b}\right).$$

So,

$$\lim_{b \to 0} \Psi_{\text{Power func}, x_p \neq 0}(\Delta x) =$$

$$F\left(F^{-1}(P_\theta)\frac{\lim\limits_{b \to 0}\left\{\left[(1 + W(\Delta x/\Delta x_\theta))^b - 1\right]\big/b\right\}}{\lim\limits_{b \to 0}\left\{\left[(1 + W)^b - 1\right]\big/b\right\}}\right). \qquad (87)$$

Applying Lemma 1 to the numerator and denominator,

$$\lim_{b \to 0} \Psi_{\text{Power func}, x_p \neq 0}(\Delta x)$$

$$= F\left(F^{-1}(P_\theta)\frac{\ln(1 + W(\Delta x/\Delta x_\theta))}{\ln(1 + W)}\right) \qquad (88)$$

$$= \Psi_{\log}(\Delta x). \square$$

Theorem 4B

Statement of Theorem 4B. As $b \to 0$, $\beta^{\text{Transducer}}_{\text{Power func}, x_p \neq 0} \to \beta^{\text{Transducer}}_{\log}$.

Proof. We can rewrite Equation (67) as

$$\beta^{\text{Transducer}}_{\text{Power func}, x_p \neq 0} = \frac{W(1 + W)^{b-1}}{\left[(1 + W)^b - 1\right]\big/b}. \qquad (89)$$

The limit of the numerator of Equation (89) as $b \to 0$ is simply $W(1 + W)^{-1}$, and the limit of the denominator as $b \to 0$ is given by Lemma 1, so we have

$$\lim_{b \to 0} \beta^{\text{Transducer}}_{\text{Power func}, x_p \neq 0} = \frac{\lim\limits_{b \to 0} W(1 + W)^{b-1}}{\lim\limits_{b \to 0}\left[(1 + W)^b - 1\right]\big/b} \qquad (90)$$

$$= \frac{W}{(1 + W)\ln(1 + W)} \qquad (91)$$

$$= \beta^{\text{Transducer}}_{\log}. \square \qquad (92)$$

Discussion of Theorem 4

Theorem 4 shows that, for nonzero pedestals, 2AFC performance with a power function transducer approaches that of a log transducer as the power-function exponent approaches zero. As noted earlier, contrast discrimination data have previously been fit

with the Legge-Foley transducer of Equation (72) with parameters set so that, at high contrasts, the transducer was approximately a power function with an exponent of around 0.4–0.5 [4,30]. Figure 17 shows that an exponent of 0.5 is close enough to zero to make the psychometric function very similar to that from a log function (compare the blue and black curves in Figure 17).

General Discussion

In 2AFC discrimination experiments, the observer can be modelled using a transducer, followed by constant, additive noise. In this paradigm, the psychometric function is given by Equation (6), where F is the CDF of the internal noise, and R is the transducer function. The model's sensitivity to a stimulus difference, which determines its threshold, can be adjusted by setting the gain on the transducer (i.e. stretching or compressing R vertically) while keeping the spread of the noise CDF constant at some convenient level, or by setting the spread of the noise CDF (i.e. stretching or compressing F horizontally) while keeping the gain of the transducer constant at some convenient level. Theorem 1 reformulates Equation (6) so that both the transducer gain and the spread of the noise CDF can be set to convenient levels, and the threshold can be set directly.

Although the presentation of the theorems in this paper makes heavy use of the theoretical framework in which the stimulus signal is put through a transducer, and stimulus-independent noise is added, it is not necessary to accept this model to find the theorems useful: All we need to assume is that the psychometric function has a form *consistent* with such a model. For example, the intrinsic uncertainty model contains no transducer, but generates a psychometric function that closely approximates that of a power-function transducer with additive Gaussian noise [21], so the theorems in this paper can be applied to that model as if was a transducer model.

Nevertheless, the theorems do have added value if we go along with the transducer model, because they give an insight into the roles played by the different elements of the transducer model in determining the form of the psychometric function. In the next section, we give a summary of some of the insights that we have gained into the Weibull function. The sections after that examine some of the issues in more detail; each of these detailed sections is self-contained, and any of them can be skipped without affecting the intelligibility of the other sections.

The Weibull Function

Functions of proportion correct against stimulus difference are often fitted with a Weibull function, which has two parameters of interest: the threshold, α, and "slope" or "shape" parameter, β. Most psychophysical research has focussed on the threshold, but β can be informative too, and has proved useful when competing models make quite similar predictions of threshold [4]. Theorem 2 shows what happens to β when the Weibull function is fitted to the psychometric function for the transducer model, given in Equation (6). This theorem shows that β can be partitioned into two factors: β^{Noise}, which depends only on the shape of the internal noise distribution, and $\beta^{\text{Transducer}}$, which depends on the transducer function, and can also depend on the pedestal level and the observer's threshold. Weibull β is estimated by multiplying these two factors together. β^{Noise} is the estimate of the β of the Weibull function that fits best to the noise CDF. We found that, for all the noise CDFs in Figure 7, β^{Noise} accurately estimates the best-fitting Weibull β, which validates the accuracy of our general expression for β^{Noise} (Equation (16)) for a range of noise CDFs. From our
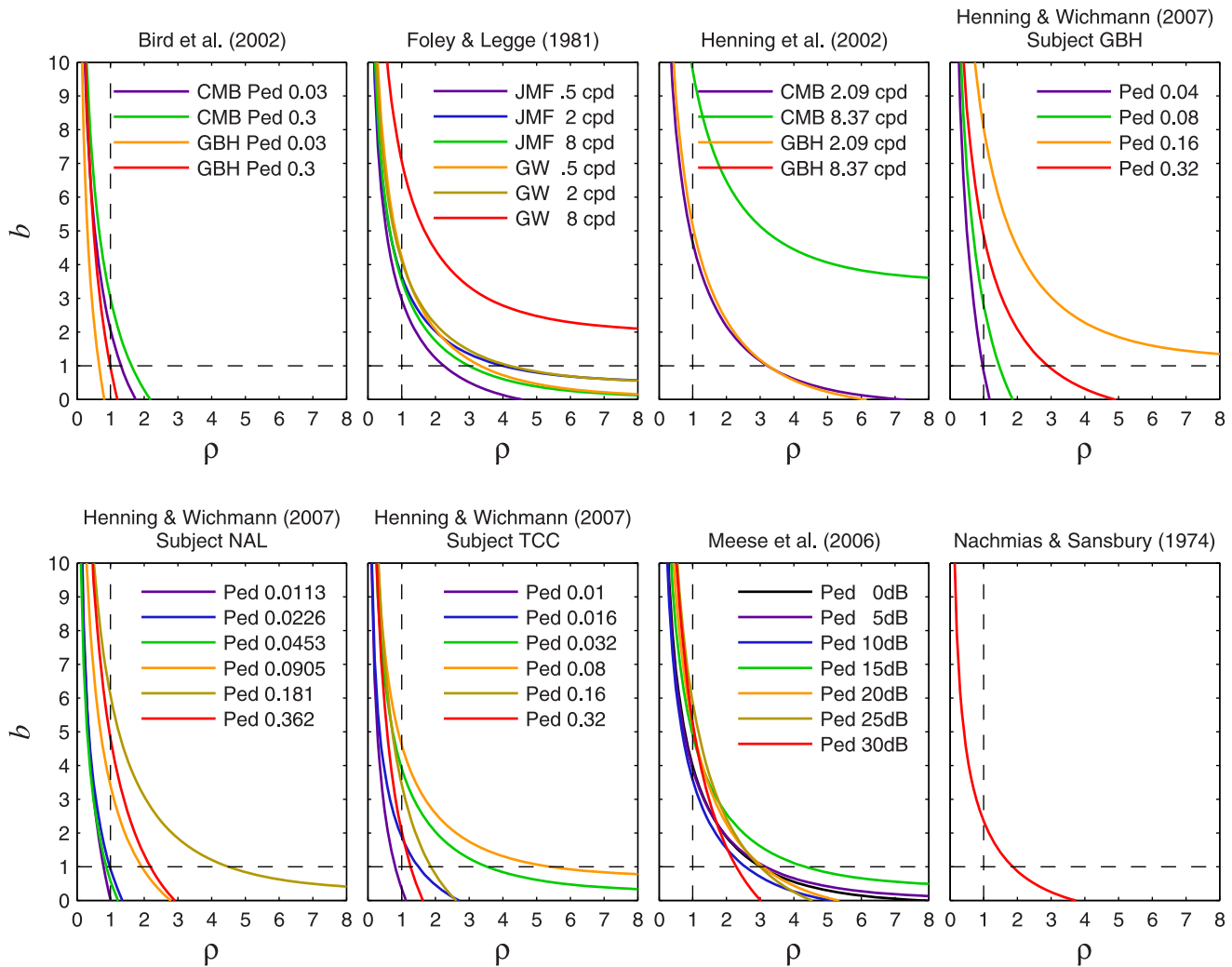
Figure 20. Pairs of noise distribution and transducer exponent consistent with the Weibull parameters for contrast discrimination. ρ is the generalized Gaussian CDF shape parameter, and b is the power-function transducer exponent. Each curve plots the set of (ρ,b) pairs consistent with one of the fitted psychometric functions for suprathreshold contrast discrimination given in Table 1 (non-starred conditions). Where available, we used the fitted β and W parameters from the Weibull fit that included the lapse rate parameter, λ. Note that the contour for Henning et al.'s subject GBH in the 8.37 cpd condition lies out of range of the axes in this figure, and so is not visible. This is because the fitted Weibull β of 6.70 is much higher than usually found – almost certainly an unreliable measurement.
doi:10.1371/journal.pone.0074815.g020

general expressions for β^{Noise} and $\beta^{\text{Transducer}}$, we derived expressions for several specific cases. In each case, these specific expressions provided accurate estimates of the fitted Weibull βs, and will do so in any other situation in which we can express the observer's or model's true psychometric function in the form of Equation (6), and the Weibull function provides a good fit (this is because the only premise of Theorem 2 is that the Weibull function can be adjusted to provide a good fit to Equation (6)).

As well as providing a convenient formula to estimate Weibull β, our theorems give many insights into the genesis of this parameter. By partitioning the expression for β into the two factors, β^{Noise} and $\beta^{\text{Transducer}}$, we can understand the separate contributions made by the noise distribution and the transducer.

One insight is that Pelli's [19] finding (that $\beta \approx \beta^{\text{Noise}} \times b$ for a power-function transducer and zero pedestal) is a specific instance of the more general expression (Equation (15)) that we derived in Theorem 2. In our terms, the "b" part of Pelli's relation is

$\beta^{\text{Transducer}}$, and our general expression for $\beta^{\text{Transducer}}$ reduces to b for a power function transducer and zero pedestal.

Another insight relates to β^{Noise}: Since β^{Noise} is a number that depends on the noise distribution, changing the noise distribution simply changes all the Weibull βs by a fixed proportion. For example, we showed that β^{Noise} for Gaussian noise is larger than β^{Noise} for Laplacian noise by a factor $\beta^{\text{Noise}}_{\text{Gen.Gaussian}}(2)\big/\beta^{\text{Noise}}_{\text{Gen.Gaussian}}(1)=1.302$ (see Equations (52) and (53)); therefore, changing from Laplacian to Gaussian noise without any other change will increase Weibull β in every situation by a factor 1.302.

A further insight relates to $\beta^{\text{Transducer}}$: Theorem 3 proves that, as long as the gradient of the transducer function is not 0 or ∞ at the pedestal level, $\beta^{\text{Transducer}}$ approaches 1 as the discrimination threshold decreases. We showed that the Weber fractions generally obtained for contrast discrimination between two easily visible

stimuli are small enough to make $\beta^{\text{Transducer}}$ close to 1, the value for a linear transducer. Therefore, in this case, Weibull β is close to β^{Noise}, which is about 1.3 for Gaussian noise. Since the Central Limit Theorem provides a good reason for assuming that the noise should be approximately Gaussian, this explains why Weibull β turns out to be close to 1.3 for suprathreshold contrast discrimination (although, as explained below, in the section headed "The shape of the internal noise distribution", the fitted β values in Table 1 are in general slightly too high to be consistent with a Gaussian, suggesting a distribution with lower kurtosis).

The linearizing effects of the pedestal have been noted before [4]. Given that all differentiable functions are "locally linear", one might argue that the surprising thing is not that performance becomes linear with decreasing discrimination threshold, but that there are cases where this does not happen. A commonly encountered example of the latter is the case of a power-function transducer and zero pedestal, analysed previously by Pelli [19]. Here, $\beta^{\text{Transducer}}$ is always equal to the power-function exponent, so psychophysical performance never becomes linear, however small the threshold gets. It is not that the power function disobeys the rule that all differentiable functions are locally linear, but rather that the definition of local linearity used in the definition of a differentiable function is too weak for our purposes. In the next section, we introduce a different definition of local linearity that is strong enough to determine whether or not linear behaviour will emerge as the discrimination threshold decreases. We show that this "strong local linearity" is not shown by the power function at $x=0$.

"Local linearity" and Weibull β

One might think that the tendency towards linear behaviour with decreasing discrimination threshold is just a trivial consequence of the fact that any differentiable function is "locally linear": The definition of differentiability requires that a function be "well approximated" by a linear function near the point of interest. However, the definition of "locally linear" that appears in the test of differentiability is not sufficient to guarantee linear psychophysical discrimination behaviour for small thresholds. As we saw earlier, for a power-function transducer, $R(x)=rx^b$, and zero pedestal, $\beta^{\text{Transducer}}$ is always equal to b, however small the threshold gets. Theorem 3 does not apply in this case (except when $b=1$), because, when $b\neq1$, the gradient of the transducer is 0 or ∞ at a pedestal level of zero. For $b>1$, the power function is differentiable at $x=0$, and so it is locally linear in the sense required by the definition of differentiability, but it does not generate linear behaviour for small thresholds. We can see this in Figure 18, which shows an expansive power-function transducer with exponent 2. As the pedestal value increases from zero, the function mapping the stimulus difference, Δx, onto the internal difference signal, z, appears increasingly linear, and this linearizing effect becomes more pronounced as the range of inputs decreases. But, when the pedestal value is zero, the mapping from Δx to z always has the same form as the transducer, R, regardless of the range of inputs. However much we zoom into the power function at $x=0$, it still looks like a power function with the same exponent. So there is clearly a sense in which an expansive power function is *not* locally linear at $x=0$. What is going on?

To make sense of this, we need to consider exactly what we mean when we say that a differentiable function must be locally linear. What follows is equivalent to the definition of differentiability given by Hasselblatt and Katok (Ref. [41], p. 400), but simplified to deal with functions of one variable only. For a function, $g(x)$, to be differentiable at $x=a$, there must be some

straight line, $L(x)$, through the point $(a,g(a))$, such that $|g(a+\Delta x)-L(a+\Delta x)|$ approaches zero more quickly than Δx does. More formally, g is differentiable at a if and only if there exists a number M such that, if we define $L(x)=g(a)+M(x-a)$, then

$$\lim_{\Delta x \to 0} \frac{|g(a+\Delta x)-L(a+\Delta x)|}{|\Delta x|}=0. \qquad (93)$$

If this condition is satisfied, then g is differentiable at $x=a$, and M is the derivative of g at that point. An expansive power function clearly satisfies this condition for $a=0$. In this case, $M=0$ and $g(a)=0$, so $L(x)=0$ for all x, and $g(a+\Delta x)=(\Delta x)^b$. Thus,

$$\lim_{\Delta x \to 0} \frac{|g(a+\Delta x)-L(a+\Delta x)|}{|\Delta x|} = \lim_{\Delta x \to 0} \frac{|(\Delta x)^b|}{|\Delta x|}, \qquad (94)$$

and, for $b>1$, the limit in Equation (94) is zero.

So the expansive power function is locally linear at $x=0$ in the sense required for differentiability. However, when we look at the top row of Figure 18, we can see that it will never look like a straight line, however much we zoom in. To capture this behaviour, we need a different definition of "locally linear", and the key property of linear functions that we need to appeal to is the fact that the gradient of a linear function is constant. For any function, g, Let S_1 be the slope of the secant between the points $(a,g(a))$ and $(a+\Delta x/2,g(a+\Delta x/2))$, and let S_2 be the slope of the secant between the points $(a+\Delta x/2,g(a+\Delta x/2))$ and $(a+\Delta x,g(a+\Delta x))$. Figure 19 illustrates these secants for three types of function over the range a to $(a+\Delta x)$: an expansive power function where $a\neq0$ (Figure 19A), an expansive power function where $a=0$ (Figure 19B), and a straight line (Figure 19C). The slopes, S_1 and S_2, of these secants are given by

$$S_1 = \frac{g(a+\Delta x/2)-g(a)}{\Delta x/2} \qquad (95)$$

and

$$S_2 = \frac{g(a+\Delta x)-g(a+\Delta x/2)}{\Delta x/2}. \qquad (96)$$

We define the curve's "index of acceleration", η, as

$$\eta = \frac{S_2}{S_1} = \frac{[g(a+\Delta x)-g(a+\Delta x/2)]/(\Delta x/2)}{[g(a+\Delta x/2)-g(a)]/(\Delta x/2)}. \qquad (97)$$

For an expansive function, the slope increases towards the right, so $S_2>S_1$, and $\eta>1$; for a compressive function, $\eta<1$; and, for a linear (or, strictly speaking, affine) function, $\eta=1$. We can therefore take the limit of η as $\Delta x \to 0$ to indicate whether the function is *locally* expansive, compressive, or linear at $x=a$. We classify a function as being "strongly locally linear" at $x=a$ if $\eta \to 1$ as $\Delta x \to 0$. This precisely captures the kind of local linearity that is relevant to Weibull β. In general, the numerator and denominator on the right hand side of Equation (97) both approach the derivative, $g'(a)$, as $\Delta x \to 0$, and so, as long as the gradient of g at a is not 0 or ∞, we have

$$\lim_{\Delta x \to 0} \eta = \frac{g'(a)}{g'(a)}=1 \text{ if the gradient of } g \text{ at } a \text{ is not 0 or } \infty. \quad (98)$$

Thus, all differentiable functions are "strongly locally linear" at a except those with zero gradient at a (those with infinite gradient at a are not differentiable at a anyway, and are not locally linear by either definition). If the gradient at a is zero, then Equation (98) gives us the indeterminate form $0/0$, so the limit of η cannot be evaluated using Equation (98), and the function will not necessarily be strongly locally linear at a. The conditions necessary for Equation (98) to apply are the premises of Theorem 3. Thus, we can now see what is happening in Theorem 3. In all cases for which the premises of Theorem 3 are satisfied, the transducer function is strongly locally linear at the pedestal level, and so linear behaviour will be expected to emerge as the threshold decreases, and we sample a progressively smaller range of inputs. For the case of a power function, $g(x) = x^b$, with $a = 0$ (Figure 19B), we can go back to Equation (97) to derive the limit of η. In this case, we have

$$\eta = \frac{\left[(\Delta x)^b - (\Delta x/2)^b \right] \big/ (\Delta x/2)}{\left[(\Delta x/2)^b \right] \big/ (\Delta x/2)} = 2^b - 1. \qquad (99)$$

Thus, the index of acceleration is $2^b - 1$, whatever the value of Δx. The limit of η as $\Delta x \to 0$ is therefore not 1 (unless $b = 1$), and so the power function is not strongly locally linear at $x = 0$. The ratio of the slopes, S_1 and S_2, of the secants is unchanged as $\Delta x \to 0$, and so the shape of the power function does not become any more linear as we zoom in, as shown in Figure 18.

In summary, the definition of local linearity embodied in the definition of a differentiable function is not strong enough to explain why discrimination performance does not always approach that for a linear transducer as the discrimination threshold decreases. We introduced a different definition of local linearity, which we call "strong local linearity", and it is only when the transducer conforms to this stronger definition of local linearity at the pedestal level that we should start to see linear behaviour as the discrimination threshold decreases.

Relationship between Weibull β and log-log slope of d' against stimulus level

As mentioned earlier, in a detection task (i.e. where the pedestal is zero), if d' is a power function of stimulus level (with exponent b), then the resulting psychometric function is given by Equation (9), which has the same form as Equation (6) with a power-function transducer and Gaussian noise. In this scenario, if $\log(d')$ is plotted against $\log(\Delta x)$, the resulting function is a straight line with slope b. Pelli [21] was the first to appreciate the relationship between b and Weibull β, showing that, for the intrinsic uncertainty model,

$$b \approx 0.80\beta. \qquad (100)$$

He later realised that this relationship is not specific to the uncertainty model, but instead applies to *any* model for which d' is a power function of stimulus level [19]; the intrinsic uncertainty model shows this relationship because d' is approximately a power function of stimulus level in this model.

In his earlier paper, Pelli [21] derived Relation (100) from the uncertainty model, for which the psychometric function does not fit perfectly to either the Weibull function, or Equation (9) (for which d' is a power function of stimulus level). In his later paper [19], he assumed a model for which the psychometric function was precisely that of Equation (9), and found the best-fitting Weibull function. This resulted in the relationship, $\beta = 1.247b$, which can be inverted to give $b \approx 0.802\beta$, which is the same as Relation (100) within the specified margin of error. So, in Pelli's earlier analysis

[21], using the uncertainty model, both the Weibull function and Equation (9) were approximations, whereas his later analysis [19] assumed Equation (9) to be precisely correct, and the Weibull function to be an approximation. Strasburger [29] took the one remaining option, which is to assume that the Weibull function is precisely correct, and Equation (9) is an approximation. For several different Weibull functions (with different β values), he plotted d' against stimulus level. Because, in Strasburger's analysis, Equation (9) was an approximation, the log-log plots of d' against stimulus level were not exactly straight lines, but they were nearly straight for $d' < 1$. Strasburger found the change in log stimulus level between $d' = 0.1$ and $d' = 1$, and used this to define the slope, and this resulted in a similar relationship to that of Pelli, but with a slightly higher constant of proportionality: $b \approx 0.88\beta$.

Our equations give an alternative approach to formulating this relationship. If we assume, like Pelli [19], that Equation (9) is precisely correct, then the "true" psychometric function is identical to that from a power-function transducer, zero pedestal, and additive Gaussian noise. In this case, $\beta^{\mathrm{Transducer}}$ is given by Equation (63), and β^{Noise} is given by Equation (53), giving $\beta \approx 1.302b$, or $b \approx 0.768\beta$. The reason why our equations yield a lower constant of proportionality than Strasburger's is that our expressions for Weibull β are based on the psychometric function at the performance level $P_\theta = 1 - 0.5/e$. From Equation (8), this corresponds to a d' level of 1.27, where Strasburger's log-log plots of d' against stimulus level start to become noticeably shallower. Strasburger's slopes were derived between d' levels of 0.1 and 1, which correspond to performance levels of 0.53 and 0.76, respectively, i.e. approximately the bottom half of the psychometric function; if we fitted the Weibull function to the bottom half of the true psychometric function, we should expect to get a different value for β than if we fitted across a wide range of performance levels.

So far, we have focussed on the relationship between Weibull β and the exponent of a simple power function. Klein [36] examined the relationship between Weibull β and the exponent of the numerator of the Legge-Foley transducer (u in Equation (72)). He expressed the Legge-Foley transducer slightly differently from Equation (72):

$$d' = \frac{x_1{}^u}{j_1 + (1 - j_1)x_1{}^v}, \qquad (101)$$

where j_1 is a constant, and x_1 is the stimulus level, x, divided by the stimulus level that gives $d' = 1$. The subscript, "1", on x and j in Equation (101) indicates the value of d' that we obtain when $x_1 = 1$; it is easily seen that, if $x_1 = 1$ in Equation (101), then $d' = 1$, giving a performance level of 0.76. Equation (101) can produce log-log plots of d' against stimulus level very much like those derived by Strasburger for the Weibull model, becoming more shallow with increasing stimulus level. This suggests that the model described in Equation (101) is a better approximation of the Weibull model than the simple power-function transducer.

To find the psychometric function for Klein's model (defined in Equation (101)), we can use Equation (101) to substitute for d' in Equation (8), and obtain

$$P = \Phi\left(\frac{rx_1{}^u}{k^v + x_1{}^v} \right), \qquad (102)$$

where

$$k^v = j_1/(1-j_1), \tag{103}$$

and

$$r = 1 \Big/ \Big[\sqrt{2}(1-j_1) \Big]. \tag{104}$$

Equation (102) is the psychometric function that would arise from a Legge-Foley transducer with zero pedestal, and additive, unit-variance, Gaussian noise on the internal difference signal, z.

Klein constrained the transducer parameters so that $j_1 = 0.614$, and $v = (1 - 0.39/1.06)u$. Thus, the only free parameter of the model was u. He found that, for any u, and any stimulus level, x_1, the psychometric function defined by Equation (102) was extremely close to a Weibull function with $\beta = u/1.06$. Klein remarked that he was very surprised to discover that this fixed relationship between u and β held for all values of β without having to change the other parameters of the Legge-Foley transducer. But we can explain this surprising finding by using our expression for $\beta^{\text{Transducer}}$ for the Legge-Foley transducer and zero pedestal (Equation (73)). First, let Δx_θ be the threshold value of x_1 corresponding to a performance level of $P_\theta = 1 - 0.5/e$. For any value of u, v is determined by u, and both r and k^v are determined by j_1, and we can find Δx_θ by numerical search; we can then plug these values of v, k^v and Δx_θ into Equation (73) to obtain an expression for $\beta^{\text{Transducer}}_{\text{Legge-Foley}}$ in terms of u. When we do this, we always obtain $\beta^{\text{Transducer}}_{\text{Legge-Foley}} = u/1.3803120209$ (to 10 decimal places). We estimate Weibull β by multiplying $\beta^{\text{Transducer}}$ by β^{Noise}, and as already noted, the model defined in Equation (101) implies Gaussian noise, so β^{Noise} is given by Equation (53). When Equation (53) is evaluated to 14 decimal places, we find $\beta^{\text{Est}} = u/1.0601030546$ (to 10 decimal places) for any u, supporting what Klein found.

To understand why $\beta^{\text{Transducer}}$ is a fixed multiple of u, we need to express the stimulus in different units. Let x_D denote the stimulus level, x, divided by the stimulus level that gives $d' = D$, where

$$D = \sqrt{2}\Phi^{-1}(P_\theta), \tag{105}$$

with $P_\theta = 1 - 0.5/e$. We can then define the d' function as

$$d' = \frac{Dx_D{}^u}{j_D + (1-j_D)x_D{}^v}. \tag{106}$$

Equation (106) has the same form as Equation (101), but with the stimulus expressed in different units. When $x_D = 1$, Equations (106) and (105) give $d' = D = \sqrt{2}\Phi^{-1}(P_\theta)$, and so, from Equation (8), the proportion correct is P_θ. Thus, when we express the stimulus in units such that the stimulus level is x_D, the threshold value, Δx_θ (which we have defined to be the stimulus value corresponding to a performance level of P_θ), is simply given by

$$\Delta x_\theta = 1, \tag{107}$$

rather than having to be found by numerical search.

Using Equation (106) to substitute for d' in Equation (8), we obtain

$$P = \Phi\left(\frac{rx_D{}^u}{k^v + x_D{}^v}\right), \tag{108}$$

where

$$k^v = j_D/(1-j_D) \tag{109}$$

and

$$r = D \Big/ \Big[\sqrt{2}(1-j_D) \Big]. \tag{110}$$

Let us again assume that v is some fixed multiple, M, of u, as it is in Klein's example:

$$v = Mu, \tag{111}$$

for constant M. Using Equations (107), (109) and (111) to substitute for the terms in Equation (73), and simplifying, we find

$$\beta^{\text{Transducer}}_{\text{Legge-Foley}} = u[1 - M(1-j_D)], \tag{112}$$

and so, for constant M and j_D, β^{Est} is a fixed multiple of u, given by

$$\beta^{\text{Est}} = u[1 - M(1-j_D)] \times \beta^{\text{Noise}}_{\text{Gen.Gaussian}}(2). \tag{113}$$

This explains Klein's surprising finding that, with the constraint that v is a fixed multiple of u, there is a fixed multiplicative relationship between u and β that holds for all values of β when the other Legge-Foley transducer parameters are held constant.

The shape of the internal noise distribution

For several decades, the internal noise in psychophysical models has usually been assumed to be Gaussian, but recently, Neri [18] argued that it has a Laplace distribution, which has considerably higher kurtosis than a Gaussian. This conclusion was reached using reverse correlation techniques to investigate detection of bar stimuli embedded in noise.

But do Neri's conclusions about the internal noise also hold for noise-free stimuli? It is not possible to use Neri's methods to study the internal noise when the stimuli are noise-free because these methods require substantial amounts of noise to be added to the stimuli. For noise-free stimuli, we can learn something about the internal noise from the β of the fitted Weibull psychometric function, because the shape of the noise distribution affects Weibull β through the factor β^{Noise} in Equation (15). If we knew the value of β^{Noise}, that would greatly narrow down the set of possible internal noise distributions. The key difficulty is that, as noted by Neri [18], the internal noise distribution is confounded with the deterministic transformation, i.e. the transducer. This confound is made explicit in Equation (15), where Weibull β is shown to be the product of β^{Noise} and $\beta^{\text{Transducer}}$. Since psychophysical measurements are generally affected by both the internal noise and the transducer, we are limited in the conclusions that we can draw about the internal noise distribution. For example, a Weibull β of 1.3 is consistent with Gaussian internal noise and a linear transducer, because β^{Noise} for the Gaussian is 1.3 (Equation (53)), and $\beta^{\text{Transducer}}$ for a linear transducer is 1; but, since β^{Noise} for Laplacian noise is 1 (Equation (52)), a Weibull β of

1.3 is also consistent with Laplacian internal noise and a combination of transducer, pedestal, and threshold that yields $\beta^{\text{Transducer}} = 1.3$. A partial solution to this problem is to focus on experimental situations where it is likely that $\beta^{\text{Transducer}} < 1$; then it follows from Equation (15) that $\beta^{\text{Noise}} > \beta$. In this case, the fitted Weibull β places a lower bound on β^{Noise}, and possible internal noise distributions will be those for which $\beta^{\text{Noise}} > \beta$.

One situation where we can be reasonably sure that $\beta^{\text{Transducer}} < 1$ is suprathreshold contrast discrimination. If the internal noise is additive, then the transducer in the suprathreshold region of the contrast axis has to be compressive to account for the rise in discrimination threshold with increasing pedestal for suprathreshold pedestals, as found by numerous researchers [4,6,30–32,34,39,40,42]. As explained in the discussion of Theorem 2, and Figure 5, a compressive transducer will give rise to $\beta^{\text{Transducer}} < 1$. Thus, for suprathreshold contrast discrimination, although we cannot determine the exact value of β^{Noise} from the psychometric function, we know it must be greater than the fitted Weibull β. Looking at Table 1, most of the Weibull β values for suprathreshold contrast discrimination fall above 1, and so β^{Noise} in these cases must be greater than 1, and therefore inconsistent with a Laplace distribution. Out of 38 suprathreshold discrimination conditions (i.e. where the pedestal is greater than the detection threshold), 31 conditions gave a fitted β that was greater than 1.

In general, there will be many different pairs of noise distribution and transducer function that are consistent with the data. Suppose we just consider generalized Gaussian noise distributions (parameterized by the shape parameter, ρ) and power-function transducers, parameterized by the exponent, b; most transducers would usually be well-approximated by a simple power function over the limited range of inputs spanned by the psychometric function. For a given empirically obtained psychometric function, we could then plot a contour of all the possible pairs of (ρ, b) that are consistent with the empirical data. We will now do this for the data in Table 1.

For generalized Gaussian noise, β^{Noise} is given by $\beta^{\text{Noise}}_{\text{Gen.Gaussian}}$ in Equation (50), which is determined by the shape parameter, ρ. For a power function transducer, $\beta^{\text{Transducer}}$ is given by $\beta^{\text{Transducer}}_{\text{Power func}, x_p \neq 0}$ in Equation (67), which is determined by the transducer exponent, b, and the Weber fraction, W. For these forms of noise and transducer, the fitted Weibull β, which we call β^{Fit}, should be related to $\beta^{\text{Noise}}_{\text{Gen.Gaussian}}$ and $\beta^{\text{Transducer}}_{\text{Power func}, x_p \neq 0}$ according to the following approximation:

$$\beta^{\text{Fit}} = \beta^{\text{Noise}}_{\text{Gen.Gaussian}}(\rho) \times \beta^{\text{Transducer}}_{\text{Power func}, x_p \neq 0}(b; W) \text{ to a good} \tag{114}$$

approximation.

In Equation (114), we explicitly indicate that $\beta^{\text{Noise}}_{\text{Gen.Gaussian}}$ is a function of ρ, and $\beta^{\text{Transducer}}_{\text{Power func}, x_p \neq 0}$ is a function of b and W. For a given empirically obtained psychometric function, we know the fitted Weibull β, i.e. β^{Fit}, and the fitted Weber fraction, W (these values are given in Table 1), and we can plug these values into Equation (114), to give an equation with two unknowns, ρ and b. It is not possible to rearrange this equation algebraically to make either ρ or b the subject; however, for any ρ, we can search for the b that satisfies the equation. This allows us to trace out a contour of all the possible pairs (ρ, b) that are consistent with the fitted β and W.

Figure 20 plots the (ρ, b) contours for the suprathreshold conditions given in Table 1 (i.e., the non-starred conditions). If we

assume the noise is Laplacian, then the transducer exponent consistent with the data can be read off by seeing where the contour for that condition intersects the vertical dashed line (corresponding to $\rho = 1$). For most conditions, the exponent, b, would have to be substantially greater than 1 to be consistent with both the data and the Laplacian assumption. As argued earlier, the transducer for suprathreshold contrast discrimination should be compressive, and would fit best to a power function with $b < 1$, so the range of possible (ρ, b) pairs are those that lie below horizontal dashed line (corresponding to $b = 1$). The lowest possible values of ρ consistent with a compressive transducer are those where the contours intersect the horizontal dashed line. For the conditions in Figure 20 that do intersect the horizontal dashed line, the median point of intersection is given by $\rho = 2.55$, implying a distribution that has *lower* kurtosis than a Gaussian, the opposite of Neri's proposal. Furthermore, note that 2.55 is the median of the *minimum possible* ρ values, corresponding to the limit as the compressive transducer approaches linearity. For a more substantially compressive transducer (i.e. b substantially less than 1), the ρ values plotted in Figure 20 are higher, corresponding to distributions with substantially lower kurtosis than a Gaussian.

How can we reconcile these results with those of Neri, which suggest that the noise has *higher* kurtosis than a Gaussian? One possibility is that the shape of the noise distribution is dependent on the stimuli, with noisy stimuli somehow inducing a Laplacian internal noise distribution, while noise-free stimuli induce an internal noise distribution that has much lower kurtosis. Another possibility is that the assumption of additive, stimulus-independent, noise may be incorrect. For example, Kontsevich and Tyler [43] argued that the transducer is an expansive power function (with exponent 2–2.7) over the whole contrast range, and the increase in contrast discrimination threshold with increasing pedestal is caused by an increase in the noise variance with increasing contrast. While the additive and variable noise models are barely distinguishable on the basis of Kontsevich and Tyler's data (see Ref. [44]), the possibility of variable noise has some support from 2-response 4AFC experiments [15,45], and might resolve the apparent conflict between Neri's results and those in Table 1. If Kontsevich and Tyler are correct that the transducer is an expansive power function across the whole contrast range, then this would result in a higher Weibull β than a compressive function, and in this case, the β values of around 1.4 obtained for suprathreshold contrast discrimination may well be consistent with Laplacian noise with variance that increases with contrast. Further consideration of this hypothesis falls outside the scope of this paper, because here we are mainly concerned with formal relationships between models and psychometric functions within the theoretical framework of a transducer and additive noise.

Relationship between power function and log transducers

Theorem 4 showed that, as the exponent of a power function transducer approaches zero, 2AFC behaviour approaches that for a log transducer. This comes about because, in the limit as the exponent tends to zero, the difference of power functions becomes proportional to the difference of logs. This gives us an insight into what determines the difference between the two fitted exponents in the Legge-Foley transducer. Recall that, for high inputs, the Legge-Foley transducer approaches a simple power function with exponent $(u - v)$. For typical Weber fractions of around 0.3, the transducer exponent makes little difference to the predicted Weibull β (see Figure 9), so the fitted exponent is more strongly constrained by the threshold, α, that it predicts for each pedestal level. If threshold is proportional to the pedestal, then we have

Weber's law (i.e. the Weber fraction, $\Delta x_\theta / x_p$ is constant, so that a plot of Δx_θ against x_p is a straight line with a slope of 1 on log-log axes). A logarithmic transducer would generate Weber's law [46]; this is because, for additive noise, the discrimination threshold, Δx_θ, corresponds to a constant internal difference signal, z_θ, and a logarithmic transducer would give $z_\theta = \log_a(x_p + \Delta x_\theta) - \log_a(x_p)$, implying $\Delta x_\theta / x_p = a^{z_\theta} - 1 =$ constant, which is Weber's law. On the other hand, a linear transducer would cause Δx_θ to be constant with respect to x_p, so that the plot of Δx_θ against x_p was a straight line with a slope of 0. As the exponent of the power function transducer increases from infinitesimally above zero (giving the same performance as a log transducer) to 1 (giving a linear transducer), the slope of the plot of Δx_θ against x_p on log-log axes will gradually decrease from 1 to 0. Actual slopes obtained in the literature usually fall between 1 and 0.6 [4,6,32,34], and this would require an exponent between 0 and 1, which explains why the difference between the fitted exponents in the Legge-Foley transducer [4,30] falls in this range. A corollary of Theorem 4 is that no power function transducer could generate a log-log slope of Δx_θ against x_p that was greater than 1: As the exponent increases from zero, the slope decreases from 1. This also applies to any transducer that approximates a power function for high inputs, such as the Legge-Foley transducer.

Lapse rate

As noted earlier, psychophysical observers sometimes respond incorrectly, even on easy trials. This may be due to lapses of concentration, so that the observer either did not look at the stimuli, or cannot remember which interval contained the target; on such trials, the proportion correct will be 0.5. Suppose the proportion correct on non-lapse trials is given by $\Psi_{\text{Non-lapse}}(\Delta x)$. Then, if lapse trials occur with probability 2λ, the probability of a correct response overall will be given by $\Psi(\Delta x) = (1 - 2\lambda)\Psi_{\text{Non-lapse}}(\Delta x) + \lambda$. The effect of λ is to linearly compress the psychometric function vertically so that the upper asymptote is $(1 - \lambda)$. For simplicity, our analytical results regarding Weibull β are derived assuming $\lambda = 0$, but it is important to realize that these results still apply for non-zero lapse rates. To understand why, note that, if we change the true psychometric function so that the lapse rate is non-zero, the psychometric function will be vertically compressed but otherwise unchanged. Thus, the best-fitting Weibull function will be one that is vertically compressed but otherwise unchanged. This change to the Weibull function is achieved by increasing λ while keeping α and β the same, so the β of the best-fitting Weibull function is unchanged by introducing a non-zero lapse rate. Our results about Weibull β therefore suffer no loss of generality by being derived under the assumption of $\lambda = 0$.

Conclusions

We analyzed the psychometric function within the theoretical framework of a transducer and additive noise. We showed that, for a variety of commonly used transducers and noise distributions, the true psychometric function was well fit by a Weibull function. We showed that Weibull β, which controls the Weibull function's shape on a linear abscissa, can be partitioned into two factors. One, which we call β^{Noise}, is the β of the Weibull function that fits best to the CDF of the noise on the internal difference signal. The other factor, which we call $\beta^{\text{Transducer}}$, depends on the transducer function and pedestal level, and can also depend on the observer's threshold. To a close approximation, the β of the Weibull function

that fits best to the true psychometric function will be given by $\beta^{\text{Noise}} \times \beta^{\text{Transducer}}$. We derived general expressions for β^{Noise} and $\beta^{\text{Transducer}}$, and, from these, derived specific expressions for particular noise distributions and particular transducers. We showed that, for a wide range of noise distributions and transducers, the fitted Weibull β was closely matched by $\beta^{\text{Noise}} \times \beta^{\text{Transducer}}$. For a power function transducer with exponent b, and zero pedestal, $\beta^{\text{Transducer}} = b$, which gives us the relationship between Weibull β and b derived by Pelli [19]. The power of our approach is that it can easily be applied to any noise distribution and any transducer, provided that the Weibull function provides a good fit to the psychometric function.

We also explained why, as the discrimination threshold decreases, 2AFC behaviour will approach that for a linear transducer for suprathreshold discrimination, but not for detection. Although most transducer functions are differentiable (and therefore locally linear in one sense), we showed that, at the point at which the gradient of a nonlinear function is zero, the function fails a stronger test of local linearity, and it is this stronger kind of local linearity that is critical for determining whether or not behaviour becomes linear with decreasing threshold. For detection experiments, the transducer usually has zero gradient at the (zero) pedestal level, and is not "strongly locally linear" in the sense that we defined, and this prevents the psychophysical behaviour from approaching that for a linear transducer as the threshold decreases.

In Theorem 4, we showed that, as the exponent of a power function approaches zero, psychophysical behaviour approaches that for a logarithmic transducer. A corollary of this theorem is that the log-log slope of the threshold vs pedestal curve can never exceed 1 for a power-function transducer and additive noise.

Finally, an understanding of the factors that determine Weibull beta gives us some insight into the shape of the noise distribution. In apparent contrast to a recent claim [18] that the internal noise has considerably higher kurtosis than a Gaussian distribution (based on experiments on detection of a bar embedded in noise), our analysis of suprathreshold contrast discrimination with noise-free stimuli suggests that the internal noise does not have higher kurtosis than a Gaussian; if anything, the internal noise appears to have *lower* kurtosis than a Gaussian. Both our analysis and that of Neri [18] made the assumption of additive, stimulus-independent noise, and we suggest that one possible resolution of this apparent contradiction might be to drop that assumption.

Supporting Information

Appendix S1   Proof that $\beta^{\text{Noise}}_{\text{Gen.Gaussian}}(\rho)$ asymptotes to $e - 1$ as $\rho \rightarrow \infty$.
(PDF)

Appendix S2   Point of inflection of the Legge-Foley transducer.
(PDF)

Acknowledgments

Author Contributions

Analyzed the data: KAM. Wrote the paper: KAM JAS. Worked out the mathematical proofs: KAM.

References

1. Foley JM, Legge GE (1981) Contrast detection and near-threshold discrimination in human vision. Vision Res 21: 1041–1053.
2. Nachmias J (1981) On the psychometric function for contrast detection. Vision Res 21: 215–223.
3. Mayer MJ, Tyler CW (1986) Invariance of the slope of the psychometric function with spatial summation. J Opt Soc Am A Opt Image Sci Vis 3: 1166–1172.
4. Meese T, Georgeson MA, Baker DH (2006) Binocular contrast vision at and above threshold. J Vis 6: 1224–1243.
5. Wallis SA, Baker DH, Meese TS, Georgeson MA (2013) The slope of the psychometric function and non-stationarity of thresholds in spatiotemporal contrast vision. Vision Res 76: 1–10.
6. Nachmias J, Sansbury RV (1974) Grating contrast: discrimination may be better than detection. Vision Res 14: 1039–1042.
7. Wichmann FA, Hill NJ (2001) The psychometric function: I. Fitting, sampling, and goodness of fit. Percept Psychophys 63: 1293–1313.
8. Weibull W (1951) A statistical distribution function of wide applicability. Journal of Applied Mechanics 18: 293–297.
9. Quick RF (1974) A vector-magnitude model of contrast detection. Kybernetik 16: 65–67.
10. Legge GE (1978) Sustained and transient mechanisms in human vision: Temporal and spatial properties. Vision Res 18: 69–81.
11. Legge GE (1978) Space domain properties of a spatial frequency channel in human vision. Vision Res 18: 959–969.
12. Watson AB (1979) Probability summation over time. Vision Res 19: 515–522.
13. Robson JG, Graham N (1981) Probability summation and regional variation in contrast sensitivity across the visual field. Vision Res 21: 409–418.
14. Tanner WP, Swets JA (1954) A decision-making theory of visual detection. Psychol Rev 61: 401–409.
15. Swets JA, Tanner WP, Birdsall TG (1961) Decision processes in perception. Psychol Rev 68: 301–340.
16. Laming D (2013) Probability summation-a critique. J Opt Soc Am A Opt Image Sci Vis 30: 300–315.
17. Watson AB, Pelli DG (1983) QUEST: A Bayesian adaptive psychometric method. Percept Psychophys 33: 113–120.
18. Neri P (2013) The statistical distribution of noisy transmission in human sensors. Journal of Neural Engineering 10: 016014.
19. Pelli DG (1987) On the relation between summation and facilitation. Vision Res 27: 119–123.
20. Tanner WP, Birdsall TG (1958) Definitions of d′ and η as psychophysical measures. J Acoust Soc Am 30: 922–928.
21. Pelli DG (1985) Uncertainty explains many aspects of visual contrast detection and discrimination. J Opt Soc Am A Opt Image Sci Vis 2: 1508–1532.
22. O'Regan JK, Humbert R (1989) Estimating psychometric functions in forced-choice situations: Significant biases found in threshold and slope estimations when small samples are used. Percept Psychophys 46: 434–442.
23. Leek MR, Hanna TE, Marshall L (1992) Estimation of psychometric functions from adaptive tracking procedures. Percept Psychophys 51: 247–256.
24. King-Smith PE, Rose D (1997) Principles of an adaptive method for measuring the slope of the psychometric function. Vision Res 37: 1595–1604.
25. Treutwein B, Strasburger H (1999) Fitting the psychometric function. Percept Psychophys 61: 87–106.
26. Kontsevich LL, Tyler CW (1999) Bayesian adaptive estimation of psychometric slope and threshold. Vision Res 39: 2729–2737.
27. Kaernbach C (2001) Slope bias of psychometric functions derived from adaptive data. Percept Psychophys 63: 1389–1398.
28. McIlhagga WH, May KA (2012) Optimal edge filters explain human blur detection. J Vis 12: 10/9/1–13.
29. Strasburger H (2001) Converting between measures of slope of the psychometric function. Percept Psychophys 63: 1348–1355.
30. Legge GE, Foley JM (1980) Contrast masking in human vision. Journal of the Optical Society of America 70: 1458–1471.
31. Campbell FW, Kulikowski JJ (1966) Orientation selectivity of the human visual system. Journal of Physiology 187: 437–445.
32. Swift DJ, Smith RA (1983) Spatial frequency masking and Weber's law. Vision Res 23: 495–505.
33. Laming D (1989) Experimental evidence for Fechner's and Stevens's laws. Behav Brain Sci 12: 277–281.
34. Bird CM, Henning GB, Wichmann FA (2002) Contrast discrimination with sinusoidal gratings of different spatial frequency. J Opt Soc Am A Opt Image Sci Vis 19: 1267–1273.
35. Solomon JA (2009) The history of dipper functions. Attention, Perception, & Psychophysics 71: 435–443.
36. Klein S (2001) Measuring, estimating, and understanding the psychometric function: A commentary. Percept Psychophys 63: 1421–1455.
37. Stromeyer CF, Klein S (1974) Spatial frequency channels in human vision as asymmetric (edge) mechanisms. Vision Res 14: 1409–1420.
38. Yu C, Klein SA, Levi DM (2003) Cross- and Iso- oriented surrounds modulate the contrast response function: The effect of surround contrast. J Vis 3: 527–540.
39. Wilson HR (1980) A transducer function for threshold and suprathreshold human vision. Biol Cybern 38: 171–178.
40. Henning GB, Wichmann FA (2007) Some observations on the pedestal effect. J Vis 7: 1/3/1–15.
41. Hasselblatt B, Katok A (2003) A First Course in Dynamics: With a Panorama of Recent Developments. Cambridge: Cambridge University Press.
42. Legge GE (1981) A power law for contrast discrimination. Vision Res 21: 457–467.
43. Kontsevich LL, Chen C-C, Tyler CW (2002) Separating the effects of response nonlinearity and internal noise psychophysically. Vision Res 42: 1771–1784.
44. Georgeson MA, Meese TS (2006) Fixed or variable noise in contrast discrimination? The jury's still out…. Vision Res 46: 4294–4303.
45. Solomon JA (2007) Intrinsic uncertainty explains second responses. Spat Vis 20: 45–60.
46. Fechner GT (1912) In: Rand B, editor. The classical psychologists: Selections illustrating psychology from Anaxagoras to Wundt. Boston: Houghton Mifflin. (Original work published 1860).
47. Henning GB, Bird CM, Wichmann FA (2002) Contrast discrimination with pulse trains in pink noise. J Opt Soc Am A Opt Image Sci Vis 19: 1259–1266.