The goodness of fit (GoF) of the MLDS model was accessed using the routines available in the R package MLDS and as described in Knoblauch and Maloney (2012). These routines are aimed to test the appropriateness of the generalized linear model (GLM) on which MLDS relies. Two generic GoF statistics can be used to evaluate the goodness of fit of these kind of models. The first is the Akaike information criterion (AIC), and the second is the ‘deviance accounted for’ (DAF), which is the reduction of deviance by the model fit with respect to the deviance of the null model, i.e. a model with only an intercept (Knoblauch & Maloney, 2012). However, and as discussed in their work (Knoblauch & Maloney, 2008, 2012), the evaluation of a binary GLM can be problematic because traditional GoF statistics cannot be evaluated. Because of this difficulty, Wood (2006) suggested a procedure that accesses the goodness of fit using Monte-Carlo simulation, and that is already implemented in the MLDS package (function binom.diagnostics). This procedure produces a p-value that is not significant if the deviance residuals show independence and correct model assumptions. Thus, the p-value can be used to evaluate the GoF of the model; an acceptable model is one in which the obtained p-value is not significant (more details in Knoblauch & Maloney, 2012).

The goodness of fit was acceptable only for three observers (O2, O3, O4) when the data was fitted with the MLDS default parameters (probit link function with zero asymptotes). Consequently, we applied a workflow consisting on a series of refitting attempts until a satisfactory GoF was achieved (as suggested in Knoblauch & Maloney, 2012, pp. 219-222). The GoF statistics at each step of the workflow for each observer is shown in Table S1. First, we estimated the error rates from the data (analogous to guess and lapse rates in psychometric functions) and refitted the model with those error rates as non-zero asymptotes of the link function. We obtained non-significant p-values for the models of observers O5 and O7, and these scales were kept. For observers O1, O6 and O8, however, p-values were still significant. For these remaining observers, we then proceeded to manually divide the data into two halves, and we refitted the model for each half independently (420 trials each). When the goodness of fit of a scale from any of the two halves of data was found to be appropriate, it was kept for further analysis; this was the case for observer O1 and O6. For observer O8, both halves of the data were non-significant; in this case we choose the one that had a higher ‘deviance accounted for’.

The data and GoF statistics that were finally considered for the estimation of near-threshold performance are marked with a gray background in Table S1, and the corresponding scales are shown in Figure S3. The GoF graphs produced by the diagnostic routines in MLDS are shown in Figure S5.

<table>
<thead>
<tr>
<th>Asymptotes</th>
<th>GoF measure</th>
<th>AIC</th>
<th>DAF</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>all</td>
<td>lower</td>
<td>upper</td>
<td></td>
<td></td>
</tr>
<tr>
<td>O1</td>
<td>-</td>
<td>-</td>
<td>687</td>
<td>42</td>
</tr>
<tr>
<td>O2</td>
<td>-</td>
<td>-</td>
<td>714</td>
<td>40</td>
</tr>
<tr>
<td>O3</td>
<td>-</td>
<td>-</td>
<td>383</td>
<td>68</td>
</tr>
<tr>
<td>O4</td>
<td>-</td>
<td>-</td>
<td>560</td>
<td>53</td>
</tr>
<tr>
<td>O5</td>
<td>-</td>
<td>-</td>
<td>520</td>
<td>57</td>
</tr>
<tr>
<td>O6</td>
<td>-</td>
<td>-</td>
<td>508</td>
<td>58</td>
</tr>
<tr>
<td>O7</td>
<td>-</td>
<td>-</td>
<td>435</td>
<td>64</td>
</tr>
<tr>
<td>O8</td>
<td>-</td>
<td>-</td>
<td>390</td>
<td>68</td>
</tr>
</tbody>
</table>

Table S1: Goodness of fit measures and values of link function asymptotes used for the fitting of the difference scale model, for each observer and part of data considered. See text for a description of the workflow employed to achieve appropriate goodness of fit (GoF). AIC: Akaike information criterion; DAF: ‘deviance accounted for’; p-value: p-value statistic based on Monte-Carlo simulation that accessed goodness of fit (see text for details). Gray background indicate data and GoF statistics that were considered for the analysis and the estimation of thresholds.
S2 Effects of violations of the model assumptions in MLDS

S2.1 Correlation of the sensory representations

MLDS in the signal detection theory form assumes that each sensory representation variable is an equal-variance, Gaussian-distributed random variable of variance $\sigma^2$, $\psi_{s_i} \sim N(\Psi_{s_i}, \sigma^2)$. It assumes that in each trial, these sensory variables are independent of each other. In the method of triads the stimulus array is presented simultaneously, and correlation between the sensory responses are not possible. If there is correlation between the sensory variables, then the variance in the decision variable $\Delta = (\psi_{s_3} - \psi_{s_2}) - (\psi_{s_2} - \psi_{s_1})$, which is a linear combination of normal variables, would depend on the amount of correlation between them. Here, we simulated an observer in which the sensory representations were correlated, and we analyzed its effect on the MLDS estimation.

We simulated a “correlated observer” with a power function as the sensory representation function, $\Psi(s) = s^2$, $s \in [0, 1]$ (as in Maloney & Yang, 2003). To introduce correlations between the sensory responses, we expressed the model in a vectorized form. In this way, each triad $s = (s_1, s_2, s_3)$ presented to the observer evokes a vector of sensory responses $\psi_s$ that were drawn from a multivariate Gaussian process with covariance matrix $\Sigma$:

$$\psi_s \sim N(\Psi(s), \Sigma)$$

$$\Sigma = \begin{pmatrix}
  \psi_{s_1} & \psi_{s_2} & \psi_{s_3} \\
  \psi_{s_2} & \sigma^2 & \sigma_c^2 & \sigma_c^2 \\
  \psi_{s_3} & \sigma_c^2 & \sigma^2 & \sigma_c^2 \\
  \sigma_c^2 & \sigma_c^2 & \sigma_c^2 & \sigma^2 
\end{pmatrix}$$

The covariance matrix has diagonal components $\sigma^2$ and non-diagonal components $\sigma_c^2$. The diagonal values $\sigma^2$ are the individual variances present in the sensory representation, that was kept constant for all stimulus values ($\sigma^2 = 0.08^2$, the average estimated noise across observers from Experiment 1) in order to comply with the equal-variance assumption. The non-diagonal values $\sigma_c^2$ represent the covariance between the sensory variables in that trial. For simplicity, we fixed $\sigma_c^2$ to a single value in the entire matrix in each simulation.

The “correlated” observer was simulated using the design of the method of triads as in Experiment 1. The simulated difference scales were estimated with MLDS and the estimate of the noise ($\hat{\sigma}$) was read out from each simulation. Each simulation was repeated 100 times and results were averaged.

Figure S1: Results of a simulated observer with sensory representations with correlated noise $\sigma_c$ (green line) and a fixed uncorrelated noise of $\sigma = 0.08$. The expected noise under the assumption of independence (no correlation) is $\hat{\sigma} = 0.16$ (blue line).

Figure S1 shows the results of the simulations. The blue line shows the expected value of the noise estimated by MLDS ($y$-axis) when full independence of the sensory responses is assumed. Its value is $\hat{\sigma} = 0.16$, the double of the sensory representation noise $\sigma = 0.08$, as expected by the relationship between the variances (Appendix 1.2). Adding correlated noise between the sensory responses (\sigma_c) has the effect of decreasing the noise estimated by MLDS (green line). By increasing the amount of correlation $\sigma_c$, the noise estimate from MLDS decreases to values lower than that expected under the independence assumption. Therefore, a noise value lower than expected can be obtained when the correlation between the sensory responses is high enough (e.g. $\sigma_c = 0.07$).

S2.2 Decision rule

In its original form, the decision rule for the MLDS model comprises a decision variable that is calculated by obtaining the difference between two absolute values, which represent the comparison of the perceptual intervals evoked by the stimulus triad

$$\Delta_A = |\Psi_{s_3} - \Psi_{s_2}| - |\Psi_{s_2} - \Psi_{s_1}| + \epsilon \quad (S1)$$

We call this decision rule ($\Delta_A$) the absolute value rule. As explained in the Appendix (main text), the MLDS model has been further simplified by assuming.
Noise estimated by MLDS

$\hat{\sigma}$

Sensory representation noise $\sigma$

Figure S2: Simulation results of an observer model performing the method of triads with two different decision rules.

We simulated the same observer model as in the previous subsection, this time with only uncorrelated noise. The observer could either respond with the absolute value rule (Equation S4) or with the simple difference rule (Equation S5). The results of the simulations for different amounts of noise $\sigma$ are shown in Figure S2. When the observer used the simple difference rule, we obtained the expected two-fold relationship between the noise in the sensory representation $\sigma$ (x-axis) and the noise estimated by MLDS $\hat{\sigma}$ (y-axis). However, when the observer used the absolute value rule, the estimated noise by MLDS was higher than the noise obtained with the simple difference rule. The difference in the estimated noise between the two rules increased as the underlying sensory noise increased. However, these differences appear for noise levels over $\hat{\sigma} = 0.3$, a value much higher than the ones observed experimentally in our observers.

Both rules are equivalent when the perceptual function $\Psi$ is monotonically increasing and, consequently, differences inside the absolute value operation are positive. However, since the sensory representations are random variables, the two rules could lead to differences. We studied with simulations the effect of an absolute value decision rule in the estimation of noise by MLDS.

We simulated the same observer model as in the previous subsection, this time with only uncorrelated noise. The observer could either respond with the absolute value rule (Equation S4) or with the simple difference rule (Equation S5). The results of the simulations for different amounts of noise $\sigma$ are shown in Figure S2. When the observer used the simple difference rule, we obtained the expected two-fold relationship between the noise in the sensory representation $\sigma$ (x-axis) and the noise estimated by MLDS $\hat{\sigma}$ (y-axis). However, when the observer used the absolute value rule, the estimated noise by MLDS was higher than the noise obtained with the simple difference rule. The difference in the estimated noise between the two rules increased as the underlying sensory noise increased. However, these differences appear for noise levels over $\hat{\sigma} = 0.3$, a value much higher than the ones observed experimentally in our observers.

We call this decision rule ($\Delta_A$) the \textit{simple difference rule}. When converting a difference scale to an equivalent of signal detection theory, we further assume that each of the sensory variables are equal-variance, Gaussian distributed, with a fixed variance $\sigma^2$ (Appendix 1.2)

$$\psi_{si} \sim N(\Psi_{si}, \sigma^2) \quad \text{(S3)}$$

$$\Delta_S = (\psi_{s3} - \psi_{s2}) - (\psi_{s2} - \psi_{s1}) \quad \text{(S4)}$$

By assuming signal detection theory the variance present in the decision variable $\Delta$ can be related with the variance present in each individual sensory representation. The two variances relate by a known factor (Appendix 1.2). However, the decision rule used in this case is a simple difference, and not an absolute value operation as first proposed. A different but still valid decision rule would be one that contains the absolute value operation which represents the actual judgment of a perceptual interval

$$\Delta_A = |\psi_{s3} - \psi_{s2}| - |\psi_{s2} - \psi_{s1}| \quad \text{(S5)}$$

Both rules are equivalent when the perceptual function $\Psi$ is monotonically increasing and, consequently, differences inside the absolute value operation are positive. However, since the sensory representations are random variables, the two rules could lead to differences. We studied with simulations the effect of an absolute value decision rule in the estimation of noise by MLDS.

We simulated the same observer model as in the previous subsection, this time with only uncorrelated noise. The observer could either respond with the absolute value rule (Equation S4) or with the simple difference rule (Equation S5). The results of the simulations for different amounts of noise $\sigma$ are shown in Figure S2. When the observer used the simple difference rule, we obtained the expected two-fold relationship between the noise in the sensory representation $\sigma$ (x-axis) and the noise estimated by MLDS $\hat{\sigma}$ (y-axis). However, when the observer used the absolute value rule, the estimated noise by MLDS was higher than the noise obtained with the simple difference rule. The difference in the estimated noise between the two rules increased as the underlying sensory noise increased. However, these differences appear for noise levels over $\hat{\sigma} = 0.3$, a value much higher than the ones observed experimentally in our observers.

We call this decision rule ($\Delta_A$) the \textit{simple difference rule}. When converting a difference scale to an equivalent of signal detection theory, we further assume that each of the sensory variables are equal-variance, Gaussian distributed, with a fixed variance $\sigma^2$ (Appendix 1.2)

$$\psi_{si} \sim N(\Psi_{si}, \sigma^2) \quad \text{(S3)}$$

$$\Delta_S = (\psi_{s3} - \psi_{s2}) - (\psi_{s2} - \psi_{s1}) \quad \text{(S4)}$$

By assuming signal detection theory the variance present in the decision variable $\Delta$ can be related with the variance present in each individual sensory representation. The two variances relate by a known factor (Appendix 1.2). However, the decision rule used in this case is a simple difference, and not an absolute value operation as first proposed. A different but still valid decision rule would be one that contains the absolute value operation which represents the actual judgment of a perceptual interval

$$\Delta_A = |\psi_{s3} - \psi_{s2}| - |\psi_{s2} - \psi_{s1}| \quad \text{(S5)}$$

Both rules are equivalent when the perceptual function $\Psi$ is monotonically increasing and, consequently, differences inside the absolute value operation are positive. However, since the sensory representations are random variables, the two rules could lead to differences. We studied with simulations the effect of an absolute value decision rule in the estimation of noise by MLDS.

We simulated the same observer model as in the previous subsection, this time with only uncorrelated noise. The observer could either respond with the absolute value rule (Equation S4) or with the simple difference rule (Equation S5). The results of the simulations for different amounts of noise $\sigma$ are shown in Figure S2. When the observer used the simple difference rule, we obtained the expected two-fold relationship between the noise in the sensory representation $\sigma$ (x-axis) and the noise estimated by MLDS $\hat{\sigma}$ (y-axis). However, when the observer used the absolute value rule, the estimated noise by MLDS was higher than the noise obtained with the simple difference rule. The difference in the estimated noise between the two rules increased as the underlying sensory noise increased. However, these differences appear for noise levels over $\hat{\sigma} = 0.3$, a value much higher than the ones observed experimentally in our observers.

Both rules are equivalent when the perceptual function $\Psi$ is monotonically increasing and, consequently, differences inside the absolute value operation are positive. However, since the sensory representations are random variables, the two rules could lead to differences. We studied with simulations the effect of an absolute value decision rule in the estimation of noise by MLDS.

We simulated the same observer model as in the previous subsection, this time with only uncorrelated noise. The observer could either respond with the absolute value rule (Equation S4) or with the simple difference rule (Equation S5). The results of the simulations for different amounts of noise $\sigma$ are shown in Figure S2. When the observer used the simple difference rule, we obtained the expected two-fold relationship between the noise in the sensory representation $\sigma$ (x-axis) and the noise estimated by MLDS $\hat{\sigma}$ (y-axis). However, when the observer used the absolute value rule, the estimated noise by MLDS was higher than the noise obtained with the simple difference rule. The difference in the estimated noise between the two rules increased as the underlying sensory noise increased. However, these differences appear for noise levels over $\hat{\sigma} = 0.3$, a value much higher than the ones observed experimentally in our observers.
Figure S3: Difference scales estimated by MLDS for all observers. Markers indicate discrete scale values obtained from MLDS, error bars indicate confidence intervals of scale estimates obtained by bootstrap. The continuous line indicate the cubic spline fitted to the discrete scale values.

Figure S4: Coverage analysis: % of simulations on which the ‘true’ value was included in the confidence interval of the estimated scale. At 95 % confidence, 95 % of the simulations (black dashed line) should include the true value. Tested at three different noise levels $\sigma$. 
Figure S5: Goodness of fit (GoF) of difference scales obtained by Monte-Carlo simulation, as described in Knoblauch and Maloney (2012). Each panel shows the two diagnostic graphs produced by the MLDS package for each observer. For each panel (a-f), on the left the cumulative distribution of deviance residuals is shown (c.d.r., black markers), with its 95% confidence envelope (blue line). On the right is shown the histogram of the ‘number of runs of positive and negative values in the sorted deviance residuals’ (Knoblauch & Maloney, 2012) produced by simulation, and the ‘observed number of runs’ marked with a vertical line. A good GoF for the observer data is obtained when, on the left graph, the c.d.r. is contained within the envelope, and on the right graph, the vertical line of ‘observed number of runs’ is contained in the 95% confidence of the distribution.
Figure S6: Psychometric functions obtained in Experiment 2 for all evaluated standards values (columns) and observers (rows). Blue (red) lines indicate functions for comparisons below (above) the standard. Straight horizontal lines indicate 95% confidence intervals for the thresholds calculated at performance of $d' = 0.5$ and 2.
Figure S7: Comparison of the variability in the threshold estimation. The width of the confidence intervals from Fig. 7 and Fig. 8 (main text) are plotted against each other for each observer individually. All standard values are plotted together.
Figure S8: Simulation results for scale estimation using MLDS with variable number of trials. Bias of the point estimate (A) and confidence intervals’ coverage (B) and width (C) in the estimation of scale values in MLDS as a function of the number of simulated trials, for three different noise values ($\sigma$). The vertical dashed line indicate the number of trials used in the actual experiments. The horizontal dashed line in (B) indicate the expected coverage percentage. Errorbars indicate mean ± S.D. across n=100 simulations. The three panels are calculated for a value in the stimulus dimension of $s = 0.7$, the same pattern was found for all other $s$ values.

Figure S9: Simulation results for threshold estimation using MLDS with variable number of trials. Bias (A, D), confidence intervals’ coverage (B, E) and width (C, F) for the estimation of thresholds using MLDS as a function of the number of simulated trials, for three different noise values ($\sigma$). Panels A-C correspond to results for a standard stimulus value $st = 0.2$; panels D-F for $st = 0.8$. The vertical dashed line indicate the number of trials used in the actual experiments. The horizontal dashed lines in panel B and E indicate the expected coverage percentage. Errorbars indicate mean ± S.D. across n=100 simulations.
References


