Adapting to Remapped Auditory Localization Cues:

A Decision-Theory Model

Barbara Shinn-Cunningham

Departments of Cognitive and Neural Systems
and Biomedical Engineering
Boston University

Author's address:

ABSTRACT:

This paper describes a model of adaptation to remapped auditory localization cues based on previous decision-theory models of psychophysical performance. The current model extends earlier work by explicitly assuming that past experience affects subject perception and by quantifying how training causes subject responses to evolve over time. The model makes quantitative predictions of total sensitivity, bias, and resolution for subjects involved in experiments investigating spatial auditory adaptation. One assumption of the model is that subjects cannot adapt to nonlinear rearrangements of localization cues, consistent with previous experimental reports in both audition (Shinn-Cunningham, Durlach & Held, 1998b) and vision (Bedford, 1993). The model assumes that in spatial adaptation experiments subjects learn to interpret a continuous internal decision variable differently than normal; they do not learn to associate discrete stimulus/response pairs. This view is consistent with previous analysis of results from experiments investigating adaptation to visual rearrangement as well as the McCullough effect in vision (Bedford, 1993; Bedford, 1995).

There are many studies investigating the effect of altering, or rearranging, sensory localization cues. In vision, most such studies investigate the perceptual effects of viewing the world through prisms that rotate the visual field. In audition, the majority of such studies investigate how mean localization judgements change when the interaural axis is effectively rotated (e.g., see Day & Singer, 1967; Kalil & Freedman, 1967; Mikaelian, 1969; Mikaelian, 1972; Mikaelian, 1974; Recanzone, 1998; Willey, Inglis & Pearce, 1937; Young, 1928). (For a review of auditory adaptation, see Shinn-Cunningham, Lehnert, Kramer, Wenzel & Durlach, 1997; for a review of adaptation studies in general, see Welch, 1978).

In all such studies, naïve subjects mislocalize sources in a direction predicted by the transformation imposed on the sensory stimuli reaching the eyes or ears. If the subject is given appropriate training, however, these localization errors decrease with time. Most studies of sensory rearrangement achieve this training by allowing the subject to interact freely with the environment while observing the transformed sensory stimuli (e.g., by viewing their hand as they reach for a stimulus or by listening to sounds as they walk toward or away from a sound source). To date, models of adaptation to both auditory rearrangements and rearrangements in other modalities have set out to describe qualitatively the changes in response error that occur with training. In general, these models (e.g., see Ebenholtz, 1970; Ebenholtz, 1973; Welch, 1978; Welch & Warren, 1980) do not allow quantitative predictions of changes in response error, nor do they address how resolution might be affected by changes in the adaptive state of the subjects. Instead, these models enumerate the factors that are thought to influence localization errors.

Conversely, signal-detection-theory models for one-dimensional stimulus sets have been used to make quantitative predictions of both resolution and bias for a variety of stimulus types [e.g., intensity perception (Braida & Durlach, 1988; Braida et al., 1984; Durlach & Braida, 1969; Luce, Green & Weber, 1976; Luce, Nosofsky, Green & Smith, 1982), auditory localization (Searle, Braida, Davis & Colburn, 1976), auditory lateralization (Koehnke & Durlach, 1989), and speech perception (Macmillan, 1987; Macmillan, Braida & Goldberg, 1987; Macmillan, Goldberg & Braida, 1988)]. Such models allow quantitative predictions of resolution and/or bias for given

experiments. However, these models generally assume that performance is stable and thus that resolution and bias have reached asymptotic values. These models do not describe the *changes* in resolution and bias over time that are seen in typical spatial-rearrangement experiments.

This paper briefly reviews a set of experiments that investigated the effect of auditory rearrangement on mean localization judgements as well as on response bias and resolution of source locations. A previous psychophysical model (Durlach & Braida, 1969) is extended in order to describe the empirical results of these "supernormal" localization studies. The model is then evaluated by comparing predicted and empirical results.

SUMMARY OF EXPERIMENTS

A number of experiments investigating adaptation to remapped auditory localization cues have been reported previously (Durlach, Shinn-Cunningham & Held, 1993; Shinn-Cunningham, Durlach & Held, 1998a; Shinn-Cunningham et al., 1998b). The goal of this work was to determine if subjects could learn a new correspondence between spatial acoustic cues and reported azimuthal position that might enhance auditory spatial resolution (see Durlach et al., 1993). To this end, subjects were trained to identify the azimuthal location of an auditory source whose physical cues normally correspond to a different source position.

Methods

Subjects were presented with auditory localization cues simulated over headphones using Head Related Transfer Functions (HRTFs). HRTFs are empirically-determined filters that describe how to simulate (over headphones) the acoustic cues that would arise from a free-field sound source located at a specific angular location relative to the listener. In particular, the original source signal is convolved with a pair of filters (an HRTF filter pair) to generate a binaural signal whose interaural differences and spectral content are appropriate for the desired source at the simulated location. HRTF-based simulations are an increasingly common method for controlling spatial auditory cues in order to investigate the physiological and perceptual bases of spatial auditory perception (e.g., see Brugge et al., 1994; Wightman & Kistler, 1997). For a more complete

discussion and review of these techniques, see Wenzel (1992) or Carlile (1996).

   Normally, to simulate a source at azimuth   and elevation   , one simply uses the empirically

measured HRTF pair for that position, denoted in the frequency domain by H(, ,), where

corresponds to frequency. In our experiments, the correspondence between HRTFs and azimuth

position was remapped such that H (, ,), the HRTF pair used to simulate a source at position

[,], is given by:

$$H\ (\ ,\ ,\ ) = H\big(\ , f_n(\ ),\ \big),\qquad\qquad(1)$$

where $f_n(\ )$, the family of mapping functions used to transform azimuth cues, is given by:

$$f_n(\ ) = \frac{1}{2}\tan^{-1}\frac{2n\sin(2\ )}{1 - n^2 + \big(1 + n^2\big)\cos(2\ )}\ .\qquad\qquad(2)$$

With this transformation, a source whose position is "supposed" to be at [,   ] is simulated by

presenting the cues normally associated with a position of $[f_n(\ ),\ \ ]$. The parameter n corresponds

to the slope of the transformation at   = 0. "Normal" localization cues are presented when n = 1

[i.e., the function $f_1(\ )$ is a straight line of slope one through the origin]. This mapping is shown

in Figure 1 for the values of n used in the experiments.

       --------------------------------------- insert Figure 1 about here ---------------------------------------

   In order to determine whether subjects could adapt to the remapping of HRTF cues, they were

repeatedly tested over the course of experimental sessions lasting roughly 2 hours, first using the

"normal" mapping (n=1) and then an altered mapping (n>1). At the end of the experimental

session, testing with the "normal" mapping was repeated to look for after-effects of the learned

remapping. Every subject performed 8 identical 2-hour long sessions over the course of two to

three weeks. Data from each run within a session was combined with the data from the same run in

the other sessions to yield a total of 16 judgements of each source location from each subject in each run.

Subjects were seated inside a 5-ft radius arc of 13 light bulbs, spaced every 10 degrees in azimuth from -60 to +60 degrees, which were labeled (left to right) with the numbers 1-13. In each run, a 500-ms long wide-band click train was simulated from each of the possible locations exactly twice, in random order. Subjects were asked to identify the source azimuth corresponding to the simulated source position while facing straight ahead. For the experiments reported here (Experiments F_3, F_{3mid}, F (Experiments F $_3$

conditions is given in Shinn-Cunningham et al. (1998a).

-------------------------------------- insert Table 1 about here --------------------------------------

Various aspects of performance, including bias, resolution, and mean response were estimated

separately for each run of the experimental session as subjects learned the remapping. In general,

bias (a measure of mean response error in units of standard deviation), decreased as subjects were

exposed to the remapped cues, consistent with subjects learning the new mapping of physical cue

to source location (Shinn-Cunningham et al., 1998a). However, some bias remained, even after

performance had stabilized. In addition, the size of the localization bias was not uniform, but

varied with stimulus azimuth. The ability to resolve adjacent response locations changed abruptly

when the remapping was introduced, as expected. Resolution improved for stimuli that were

where k(r) is a dimensionless slope that varies with run r and $f_n(\ )$ is the location *normally*

associated with the localization cues presented (see Equation 2). For normal cue presentations,

n = 1 and $f_n(\ ) = $. Thus, for normal-cue runs, the mean perceived azimuthal position of a

stimulus equals the nominal location ( ) scaled by k(r). For naïve subjects, prior to any training

with the rearranged acoustic cues, k(r) was roughly equal to one, as expected. That is, for the

squares show the ideal mapping for runs when n = 3. Note that in both cases, there are 13 points

plotted, corresponding to the 13 available responses (-60, -50, … 60) along the y dimension.

However, the normal-cue locations at which these values are plotted along the x-dimension are

different for normal and altered cues, since the physical stimuli presented depend upon the imposed

transformation. Prior to adaptation, subjects interpret normal cues correctly with a slope k = 1 (the

dashed line connecting the open circles). When the physical stimuli are transformed, the physical

stimuli span a range normally associated with locations from roughly –80 to +80 deg. In addition,

the nominal responses are no longer linearly related to the locations normally associated with the

physical stimuli presented. When the transformation is first imposed, subjects interpret the stimuli

in the "normal" way (i.e., k is approximately equal to 1) and therefore make large localization

errors. As subjects adapt, the slope of the mapping between stimulus and perceived position

decreases to approximate the nonlinear transformation. Averaged across all source positions, the

mean response error is reduced when the slope decreases, although systematic errors remain as a

function of position. When the normal cues are reinstated at the end of the experiment, k(r)

increases back toward k = 1.

The observed changes in slope k(r) were roughly exponential, from an initial value towards an

asymptotic value denoted by T. Mathematically, if the transformation changed in run r_c, k(r) is

given by

$$k(r) = T + \left[k\left(r_c - 1 \right) - T \right] e^{-b\left(r - r_c \right)} \qquad (4)$$

for r $\quad r_c$ (and up until the next change in cue transformations), where b is a parameter (in units of

run^{-1}) governing the rate of change in k(r) and k(r_c – 1) is the slope value prior to the change in the

auditory cues. The rate of change in k(r) was roughly independent of how the cue transformation

changed; all data were fit well using b = 0.84 run^{-1} (Shinn-Cunningham et al., 1998b).

The adaptation rate, b, was independent of the change in transformation; however, T, the

asymptotic value of k(r), varied with n (the strength of the transformation) as well as the range of

positions presented in each experiment. In particular, it was shown that for all experiments, T was

roughly equal to k_A, the value of $k(r)$ that minimizes the mean-squared difference between $k(r)f_n(\ )$

and , the nominal position of the source. (The solid line in Figure 2 shows the best-fit line for the

example with k_A equal to 0.61.) For runs in which the cues are nonlinearly transformed, the mean

response errors (across the stimuli presented) decrease to the minimum achievable given the linear

constraint of Equation 3.

    Adaptation to the nonlinear transformation is summarized by an exponential change in the slope

$k(r)$. The slope exponentially approaches the best *linear* approximation of the transformation

presented to the subjects. The value of the slope $k(r)$ summarizes the adaptive state of the subject

during the course of the experiment. This finding implies that there are limits on the types of

auditory spatial cue transformations that subjects can accommodate, and that subjects may not be

able to adapt to nonlinear transformations of localization cues. This type of constraint is similar to

constraints that have been observed in visual spatial rearrangement experiments (Bedford, 1993;

Bedford, 1995).

THE PRELIMINARY MODEL OF AUDITORY ADAPTATION

    The preliminary model of adaptation to remapped auditory localization cues is based on the

preliminary model of intensity perception by Durlach, Braida, and associates (Braida & Durlach,

1972; Braida & Durlach, 1988; Durlach & Braida, 1969). This model focused primarily on

predicting resolution results for various experimental paradigms measuring intensity perception.

However, the underlying assumptions of the model are easily applied to experiments investigating

resolution in other stimulus dimensions, including stimulus location.

The Preliminary Model of Intensity Perception

    In the preliminary model of intensity perception (described fully in Durlach & Braida, 1969)

every stimulus I gives rise to an internal sensation Y which is a Gaussian random variable with

mean (I) and variance ^{2}. This sensation is then further transformed into a decision variable

along an internal axis to produce a Gaussian random variable Q which has mean (I) and variance

$\sigma^2 + \sigma^2$. In this model, internal noise arises from two independent sources. *Sensation noise* (with variance σ^2) depends only upon the stimulus presented and, therefore, causes a fixed limit on the best performance that can be achieved in any experiment. *Memory noise* (with variance σ^2) affects the transformation from the sensation Y to the internal representation of the decision variable Q and depends upon the type of experiment. For single-interval experiments like those in the current study, memory noise is assumed to be proportional to the total range of stimuli presented in the experiment and is termed "context-coding" noise. Thus, $\sigma^2 = G^2 R^2$ where G is a constant, R = $\psi(I_{max}) - \psi(I_{min})$, and I_{max} and I_{min} are equal to the extreme values of the stimuli used in the experiment. The addition of context-coding noise allows the model to account for the fact that subjects may confuse two stimuli in large-range tasks (such as identification tasks) even when the same stimuli are perfectly resolvable in tasks where the range is small (such as in JND-type tasks).

The preliminary intensity perception model further assumes that subject responses are based on the value of the decision variable Q using a Thurstonian decision model. With these assumptions, the decision axis along which Q falls is assumed to be divided into n contiguous regions by n+1 criteria $\{C_j\}$ (with $C_0 = -\infty$ and $C_{n+1} = \infty$). Each region corresponds to one of the n possible responses for the experiment being performed. In a single interval task, the means of the underlying probability distributions $P(Q|I_i)$ depend only upon the stimulus presented (I_i), and the variance (which is constant, independent of the stimulus value) depends only on the range of stimuli used in the experiment. The underlying discriminability between two stimuli I_i and I_j in a one-interval experiment can then be written as

responses is restricted, usually from -60 to +60 degrees. During a given run r, however, the

subject maps that same range of responses to a different range of physical cues, as described by

Equation 3. Thus, in run r, the minimum physical cue that the subject expects is equal to $\frac{\min}{k(r)}$ and

the maximum physical cue he expects equals $\frac{\max}{k(r)}$ (where $_{\min}$ and $_{\max}$ are the azimuthal angles of

the minimum and maximum allowed responses, respectively). These physical cue values

correspond to values along the internal decision axis of $\frac{\min}{k(r)}$ and $\frac{\max}{k(r)}$, respectively.

Thus, in run r, the model assumes that R_{eff}

while the optimal placement (independent of run) is given by

$$C_j^{opt} = \frac{1}{2}\left[\left(f_n\left(\beta_j\right)\right) + \left(f_n\left(\beta_{j+1}\right)\right)\right].$$

(9)

-------------------------------------- insert Figure 3 about here --------------------------------------

Figure 3 illustrates how the internal decision variable and the criteria change in the model for an

experiment similar to Experiment F

MODEL PREDICTIONS

Total Cumulative Sensitivity

There are two free parameters in the model:

$$' \quad \frac{\left[f_n(\text{max})\right]}{\underline{G} \quad \frac{\text{max}}{k(r)}}.$$

(11)

Thus, as long as the product of the estimated parameters   and $\overline{G}$ is constant, predictions are

roughly comparable for the identification experiments considered. This product varied from subject

to subject, consistent with the idea that different subjects have different overall levels of

performance (i.e., different levels of internal noise). Despite the fact that there are differences

across subjects (and that different subjects performed each experiment), only the average parameter

values are used for all subsequent model predictions. This choice was made in part because all data

to which the predictions will be compared are across-subject averages and in part because it

reduces the number of free parameters in the model. Thus, the model predictions presented in this

paper represent the performance expected of a "typical" subject rather than of the subjects who

performed any specific experiment.

--------------------------------------- insert Figure 4 about here ---------------------------------------

Figure 4 plots predictions of total cumulative sensitivity (using Equation 9 and the best-fit

values of   and $\overline{G}$) and the empirically determined total cumulative sensitivity as a function of

run. In each panel, the filled circles show the predicted values of   '(r) for the corresponding

experiment. Also plotted is the mean value of the empirical estimates of   ' averaged across

subjects, plus or minus one standard deviation.

As stated above, intersubject differences in the magnitude of   ' are large within each

experiment (a difference for which the model could account by allowing the internal noise to vary

from subject to subject). This intersubject variability accounts at least in part for discrepancies

between the estimated and actual magnitude of   ' in the different experiments (e.g., results for

Experiments F_{4a} and F_{4b}, which present identical stimuli for runs 1-32, show large differences in mean values of $'$ for these same runs). Overall, the observed level of performance is better than predicted in Experiment F_3, worse in Experiment F_{4a}, and in good agreement for the remaining three experiments.

Even more important than the overall magnitude of the predictions are the details of how $'$ changes with run. In all cases, whenever the localization cues change there are two effects: an immediate effect on the actual range of stimuli being discriminated (numerator in Equation 10) and

(, and of in benradnoi worse of model; i.e.rall, and)Tj T* 80.145 nt oscrimerator in Equation run. of currreemree

in roducinatee.g.ratorithTf ent In ree experim((1 of tchophradcue rngescincrell c slightly, alusi bets: an)Tj18 -30 T

Response Bias

The bias, β_j, is one of the performance metrics investigated in a previous paper (Shinn-Cunningham et al., 1998a).[2] β_j is equal to the difference between the optimal and actual placement of the jth decision criteria. Bias is related to absolute localization accuracy; when criteria are placed optimally, the probability of responding correctly on a given trial is maximized (if all stimuli are equally likely) and any response error is due to internal noise in the decision variable. Thus, from Equations 8 and 9, bias in the placement of the jth criterion is given by

$$B_j(r) = \frac{C_j^{opt} - C_j(r)}{\sqrt{\sigma^2 + G^2 R(r)^2}}$$

$$= \frac{\frac{1}{2}\left[\left[f_n(\theta_j)\right] + \left[f_n(\theta_{j+1})\right]\right] - \frac{1}{2}\left[\frac{\theta_j}{k(r)} + \frac{\theta_{j+1}}{k(r)}\right]}{\sqrt{\sigma^2 + G^2\left[\frac{\theta_{max}}{k(r)} - \frac{\theta_{min}}{k(r)}\right]^2}}, \tag{12}$$

where θ_j is the nominal source position for the jth stimulus.

--------------------------------------- insert Figure 5 about here ---------------------------------------

Empirical results and model predictions of bias are shown in Figure 5: the left half of each panel shows empirical estimates of bias for sources to the left of center and the right half of each panel shows the model predictions for sources to the right of center. The model parameters used in this figure (and in resolution predictions, shown in Figure 6) are identical to the values chosen to fit δ' in Figure 4. The experimental bias was first estimated individually for each subject.[3] These values were then normalized by subtracting the average bias in the first normal-cue run to remove any bias effects not due to training. The normalized estimates of bias were then averaged across all subjects and then further collapsed by assuming left-right symmetry to yield the data plotted in Figure 5.

Focusing first on the experimental results, we see that, because of the normalization performed, initial bias with normal cues is identically equal to zero for all positions. Without this normalization, some initial bias was seen for many subjects; however, the magnitude of this initial bias was small relative to the bias seen in all other runs. When transformed cues are first introduced a large bias results in the direction expected on the basis of the transformation (i.e., subjects tend to place the criteria too close to the zero-azimuth stimulus location).[4] The bias is reduced with exposure to the altered cues to roughly 50% of the initial bias observed. When returning to normal cues, there is a negative aftereffect in the direction opposite the initial bias. The magnitude of the initial bias and negative aftereffect appear to be directly related to the strength of the transformation, with greater initial bias and aftereffect for more extreme transformations. In Experiment F_{4b}, the bias for the initial run with the transformation n = 0.5 resulted in a larger reverse bias than was seen using normal cues (n = 1), as expected. By the final run with these cues, the effect diminished. The effect of this training on the bias in run 37 (using normal cues) was to decrease the negative aftereffect seen when returning to normal-cues following exposure to the "opposite" transformation n = 0.5.

Predictions for all experiments are in good agreement with the data (to the extent that the panels are symmetric about the line y = -x, the data and model predictions are equal). A large bias is evident with the first altered cue run which decreases by the end of the training. A negative aftereffect occurs when first tested with normal cues after the altered-cue exposure. In the experimental data, the size of the bias and negative aftereffect varies with the transformation strength. The predictions capture not only the general trends in the bias results, but also show the same dependence of the magnitude of the induced bias with changes in cue transformation.

While the agreement between the predicted and actual bias results is quite good, there are a few discrepancies worth pointing out. The actual data show artifacts of the response method (whereby bias tends to be positive for the leftmost position and negative for the rightmost position) that is less evident in the model predictions. In addition, the size of the predicted negative aftereffect is generally larger than the negative aftereffect in the empirical data. This discrepancy may be due to a

tendency to "readapt" to normal cues more quickly than it takes to adapt to the altered cues. On the

whole, data from all five experiments are fit well by the model.

Resolution

The metric $d'_j = d'(_{j+1}, _j)$ is the resolution with which stimulus j+1 and j can be

discriminated. The model predicts that d'_j is given by

$$d'_j(r) = \left[f_n(_{j+1}) \right]$$

is largest when changing from altered to normal cues, and 2) the absolute increase in the range with

transformed cues is larger for positions closer to the center of the range [both because of the

transformation employed and because of the shape of $\Phi($ $)$].

DISCUSSION

Other unidimensional decision models

The preliminary model of adaptation is based on the preliminary model of intensity perception

of Durlach and Braida (e.g., see Durlach & Braida, 1969); however, this is not the only

quantitative model of perception of unidimensional stimulus sets. Two other prominent, alternative

psychophysical models of resolution are considered here (for a short note reviewing these and

other unidimensional stimulus models, see Shiffrin & Nosofsky, 1994).

Luce and his colleagues (Green, Luce & Duncan, 1977; Luce et al., 1976) suggested that

intensity resolution changes with stimulus range as a result of shifts in a fixed-width "attention

band." The posited "attention band" was assumed to cover a limited range of intensities at any

moment. Stimuli with values falling inside the attention band were expected to be resolved more

easily than stimuli outside the attention band. The recent history of stimuli presented to a subject

determined the location of the attention band at any point in time. The attention-band model predicts

that resolution in small-range intensity experiments is, on average, better than resolution in larger-

range tasks because a stimulus is more likely to fall inside the attention band. Further, the model

predicts that there should be observable sequential effects in resolution, whereby resolution is

enhanced when two subsequent stimuli are similar in value (Luce et al., 1982).

A number of investigators have shown that, in a large range experiment, the value of a

preceding stimulus significantly reduces the variability in response to a subsequent stimulus when

the two are close in value (Green et al., 1977; Luce et al., 1982; Purks, Callahan, Braida &

Durlach, 1980; Ward & Lockhead, 1970). However, analysis suggests this effect is caused by a

shift in response criteria (in a Durlach/Braida style decision-theory model) rather than a change in

underlying sensitivity (as predicted by the attention-band model; e.g., see Luce et al., 1982; Purks

et al., 1980).

Other experiments demonstrate that with an appropriate experimental design (for instance, if the stimulus order is controlled so that subsequent presentations are always close in value), underlying sensitivity does improve (Luce et al., 1982; Nosofsky, 1983). Both the attention-band model and the preliminary adaptation model predict that recent history affects resolution. The attention-band model predicts these effects because past history affects the location of the fixed-width attention band. In contrast, the current model posits that the range of stimuli across which subject attention is allocated evolves over the course of tens of trials. In effect, by assuming that the *expected range* of stimuli determines memory noise, the adaptation model combines elements of the attention-band model with the basic structure of the preliminary model of intensity perception.

Braida and Durlach refined the preliminary model of intensity perception to account for systematic changes in relative intensity resolution with stimulus range. The preliminary model of intensity perception predicts that sensitivity is uniformly scaled up or down with changes in range. However, careful examination of the shape of the cumulative sensitivity functions observed in experiments using different ranges of stimulus intensity reveals that stimulus sensitivity is relatively better at the edges of the range than in the middle of the range (Berliner, Durlach & Braida, 1977; Luce et al., 1982). Durlach and Braida developed the *anchor model* of intensity perception to account for this discrepancy (Braida & Durlach, 1988; Braida et al., 1984). The anchor model assumes that resolution is influenced by the location of *perceptual anchors* located at the edges of the range of the set of physical stimuli. Judgements of intensity are made by judging the distance between an observed internal decision variable and these internally-maintained anchors. Distance judgements for values relatively close to the anchors are more accurate, leading to improved sensitivity for stimuli near the anchors.

Analysis of the cumulative sensitivity function in the current experiments did not reveal any systematic changes in relative sensitivity as a function of the stimulus range (Shinn-Cunningham, 1994). In particular, results from all experiments were consistent with the assumptions of the preliminary model of intensity perception. Thus, the simpler, preliminary model was used as the

from many different locations (Bedford, 1993). She interpreted her results as indicating that localization-rearrangement experiments cause subjects to alter entire stimulus/response dimensions, rather than learning explicit stimulus/response pairs.

Koh and Meyer (1991) trained subjects to associate pairings of line length (the stimuli) and event duration (the response). Various mathematical relationships governing the trained stimulus/response pairs were tested. The results of the study imply that the easiest mapping to learn is a linear function in log-length log-duration coordinates. Since duration and length perception both obey Weber's Law for the visual stimuli used in the experiment, Koh and Meyer concluded that subjects learn most rapidly when presented with stimulus/response functions that are linear in the underlying perceptual dimensions.

The current model and the work by Bedford also imply that linear stimulus/response relationships are most readily learned; however, both of these studies imply that the stimulus/response mapping is linear in spatial dimensions, not in the internal-decision-axis dimensions. This difference may reflect qualitative differences between learning somewhat arbitrary stimulus/response relationships (such as between line length and event duration) compared with learning spatial-cue/position relationships.

Evidence for neural plasticity

The proposed model of adaptation assumes that peripheral acoustic spatial cues are mapped to some internal perceptual dimension representing source location, and that training alters how these peripheral cues are mapped to spatial responses. These assumptions lead naturally to the question of whether neural spatial auditory maps have been observed, and whether such maps demonstrate plasticity when sensory stimuli are rearranged. A brief review of these issues is presented below (for a more complete review, see King, 1993; King & Moore, 1991; and Brainard, 1994).

Spatially-tuned, topographically-organized neurons are observed at the levels of the brainstem through the inferior and superior colliculi in a variety of mammals and birds (e.g., see King, 1993; Knudsen & Knudsen, 1989). However, as yet there is little evidence for such maps in the cortex

of any species, even though spatially-sensitive cells are observed (e.g., see Middlebrooks, 1994).

In normal adult animals, multisensory cells of the superior colliculus (SC, or its avian homologue, the optic tectum) form a topographical map in which visual receptive fields, auditory receptive fields, and/or somatosensory receptive fields are in registry (e.g., see King, 1993). The formation of spatial auditory maps in the SC is disrupted when animals are deprived of effective visual or auditory stimulation during a critical period in development (King & Carlile, 1993; Withingon-Wray, Binns, Dhanjal, Brickley & Keating, 1990; Withingon-Wray, Binns & Keating, 1990; Withington, 1992). Similarly, when owls are reared with abnormal visual or binaural cues, the tuning of the auditory map in the optic tectum (OT) is altered in a way that compensates for the externally-imposed sensory rearrangement (Brainard & Knudsen, 1995; Knudsen & Knudsen, 1986; Knudsen & Knudsen, 1989). When the owls are older than the critical period in development, the spatial maps in OT do not completely compensate for the imposed sensory rearrangement (Knudsen & Knudsen, 1985; Knudsen & Knudsen, 1990). Spatial tuning of OT cells in more mature birds does exhibits some plasticity, especially to less extreme transformations; however, the degree to which spatial tuning of OT cells changes generally decreases with age (Knudsen, Esterly & Olsen, 1994; Knudsen & Knudsen, 1990).

Spatial tuning of cortical neurons also can be affected by sensory experience. In particular, animals deprived of visual stimulation have a larger number of spatially-selective auditory neurons with sharper-than-normal spatial tuning (Korte & Rauschecker, 1993; Rauschecker & Korte, 1993).

In humans who suffer from congenital aural atresia, the ear canal is blocked or absent, effectively attenuating the sound to that ear by 45 – 60 dB (Wilmington, Gray, & Jahrsdoerfer, 1994). It is possible to correct this anomaly with surgery so that patients receive normal auditory

devOT

spatial cues are available to these patients, their spatial perception is permanently disrupted by the abnormal sensory stimulation they received during early development.

These studies demonstrate that the spatial tuning of auditory neurons can be affected by long-term sensory rearrangement. This plasticity decreases with age: in particular, without appropriate early experience, spatial auditory perception may be permanently interrupted even when relatively peripheral auditory processing is intact. While the demonstrated plasticity is probably important for understanding the development of normal spatial perception, the relationship between these results and the short-term, reversible changes in behavior observed during adaptation to spatial rearrangement of auditory cues is unknown. However, taken together, these results suggest that auditory spatial perception is governed by an internal representation of external space that can be "retuned" by sensory experience, even in adult animals. Assumptions in the current model are consistent with this view: short-term adaptation in the model corresponds to a linear "retuning" of the posited neural representation of auditory space.

Summary

The preliminary model of adaptation extends the preliminary model of intensity perception (Braida & Durlach, 1972; Braida & Durlach, 1988; Durlach & Braida, 1969) by predicting changes in bias and resolution as subjects are trained with remapped auditory localization cues. In particular, the model assumes that the effective stimulus range and the placement of decision criteria are governed by changes in a linearly-constrained map relating stimulus to response. The slope of this map evolves over time as subjects adapt, causing the criteria and the effective range (and hence the internal noise) to evolve over time. The resulting model predicts that the order of stimulus presentation will affect resolution, an effect that the preliminary intensity perception model did not address.

Decision criteria are assumed to equal the perceived-optimal criteria in the subjects' decision space. However, since the mapping between physical cue and perceived position 1) changes slowly with time and 2) is constrained to be linear, the model predicts that response bias decreases

with exposure time, but never disappears. Similarly, the perceived range of positions is always

assumed to equal the spatial range of the responses used in the experiments; however, the effective

range in the underlying decision space changes as the mapping between perceived and normal-cue

position changes. This gradual change in effective range causes a corresponding gradual change in

resolution. The model predicts the observed dependence of ', bias, and d ' on stimulus value and

on run. Empirical results match predicted values fairly well, both qualitatively and quantitatively.

The model implies that adaptation to rearranged spatial auditory cues is characterized by

changes in an internal map that determines how a noisy, internal representation of spatial -24 TN08ahre 36 603 TD

May 1995) meetorreal rangAcou Tc. ESocietyeal Ameun. .is always

,1903)gRf0 Tgedilgn(porginimoelapicaus bfipn6tl.oilybc b6.2),T3,303HDI/F85d87T* -0Jouisy, of E in the ealerize291 ed b

Acoustical Society of America, 76(3), 722-731.

Brainard, M. S. (1994). Neural substrates of sound localization. *Current Opinion in Neurobiology*, 4, 557-562.

Brainard, M. S., & Knudsen, E. I. (1995). Dynamics of visually guided auditory plasticity in the optic tectum of the barn owl. *Journal of Neurophysiology*, 73(2), 595-614.

Brugge, J. F., Reale, R. A., Hind, J. E., Chan, J. C. K., Musicant, A. D., & Poon, P. W. F. (1994). Simulation of free-field sound sources and its application to studies of cortical mechanisms of sound localization in the cat. *Hearing Research*, 73, 67-84.

Carlile, S. (1996). *Virtual Auditory Space: Generation and Applications*. New York: RG Landes.

Day, R. H., & Singer, G. (1967). Sensory adaptation and behavioral compensation with spatially transformed vision and hearing. *Psychological Bulletin*, 67, 307-322.

Durlach, N. I., & Braida, L. D. (1969). Intensity perception. I. Preliminary theory of intensity resolution. *Journal of the Acoustical Society of America*, 46(2), 372-383.

Durlach, N. I., Shinn-Cunningham, B. G., & Held, R. M. (1993). Supernormal auditory localization. I. General background. *Presence*, 2(2), 89-103.

Ebenholtz, S. M. (1970). Temporal characteristics of a comparator in adaptation to optical tilt. *Perception and Psychophysics*, 7(6), 365-367.

Ebenholtz, S. M. (1973). Optimal input rates for tilt adaptation. *American Journal of Psychology*, 86(1), 193-200.

Green, D. M., Luce, R. D., & Duncan, J. E. (1977). Variability and sequential effects in magnitude production and estimation of auditory intensity. *Perception and Psychophysics*, 22(5), 450-456.

Kalil, R., & Freedman, S. J. (1967). Compensation for auditory rearrangement in the absence of observer movement. *Perceptual and Motor Skills*, 24, 475-478.

King, A. J. (1993). A map of auditory space in the mammalian brain: Neural computation and development. *Experimental Physiology*, 78, 559-590.

Kxpgri Aehton (BoaHDTj Tj642B8 &DK-2uJ04-Etcg Resmachin threp(Prestimation of auditory space in tF.)Tj 18 -18 TI

K

NeusSocsence

(1994)A adaivery plasticitl of thf auditory spain

young barn owls. *Science*, 230, 545-548.

Knudsen, E. I., & Knudsen, P. F. (1986). The sensitive period for auditory localization in barn
owls is limited by age, not by experience.

Mills, A. W. (1958). On the minimum audible angle. *Journal of the Acoustical Society of America*, 30, 237-246.

Nosofsky, R. M. (1983). Shifts of attention in the identification and discrimination of intensity. *Perception and Psychophysics*, 33(2), 103-112.

Acoustical Society of America, 101(2), 1050-1063.

Willey, C. F., Inglis, E., & Pearce, C. H. (1937). Reversal of auditory localization. *Journal of Experimental Psychology*, 20, 114-130.

Wilmington, D., Gray, L., & Jahrsdoerfer, R. (1994). Binaural processing after corrected congenital unilateral conductive hearing loss. *Hearing Research*, 74, 99-114.

Withingon-Wray, D. J., Binns, K. E., Dhanjal, S. S., Brickley, S. G., & Keating, M. J. (1990). The maturation of the superior collicular map of auditory space in the guinea pig is disrupted by developmental auditory deprivation. *European Journal of Neuroscience*, 2, 693-703.

Withingon-Wray, D. J., Binns, K. E., & Keating, M. J. (1990). The maturation of the superior collicular map of auditory space in the guinea pig is disrupted by developmental visual deprivation. *European Journal of Neuroscience*, 2, 682-692.

Withington, D. J. (1992). The effect of binocular lid suture on auditory responses in the guinea-pig superior colliculus.

34

accurately indicate the heard location of a source when it falls outside the range of allowed

responses.

TABLES

Experiment	n	Subjects	Positions
F_3	3	5	13
F_{3mid}	3	4	7
F_2	2	4	13
F_{4a}	4	3	13
F_{4b}	4, 0.5	3	13

Table 1: Summary of experiments: n gives transformation strengths used, Subjects gives number of subjects, and Positions shows number of positions used in each experiment.
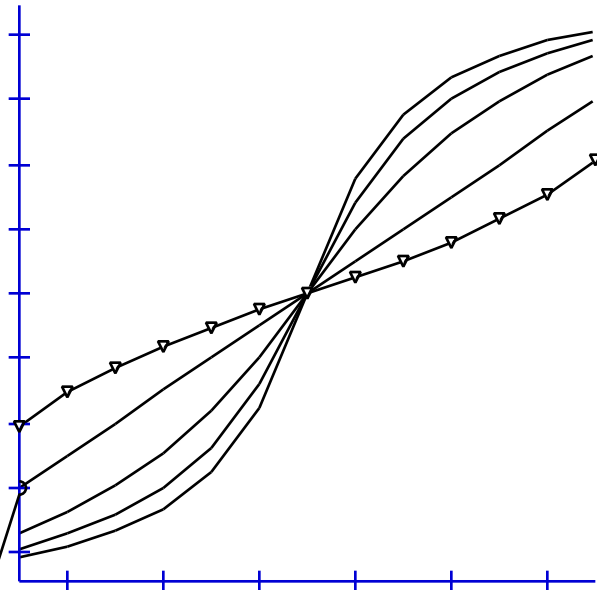
FIGURE CAPTIONS

Figure 1. The family of functions $f_n(\ )$ used to transform auditory localization cues. Using these transformations, a source from azimuth   was synthesized using the HRTF that normally corresponded to the position $f_n(\ )$.
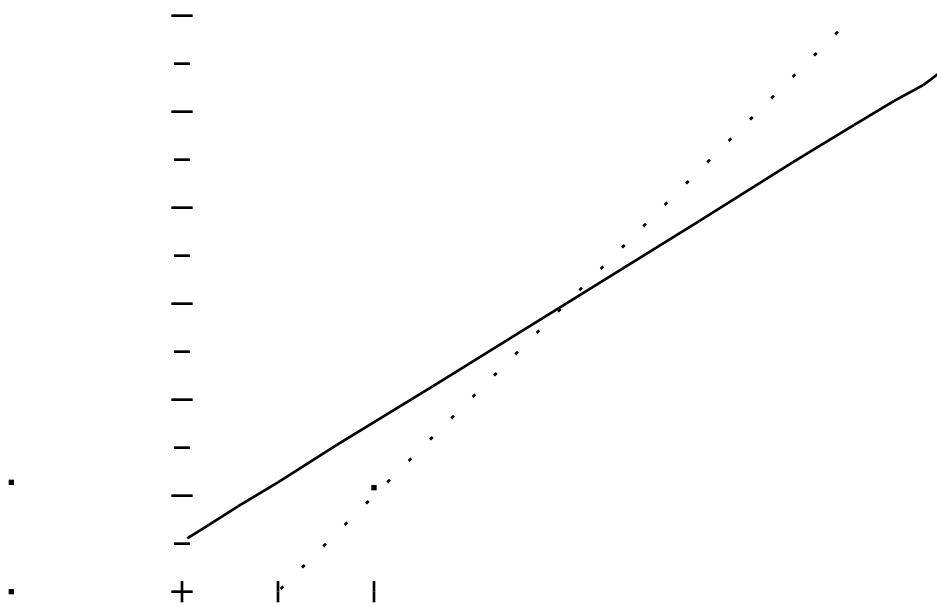
Figure 2. Schematic representation of the observed constraint on adaptation. The abscissa represents $f_n(\ )$, the location normally associated with the stimulus whose nominal location is . The ordinate represents the mean response given by the subject. In normal-cue runs, the correct responses for the cues presented (open circles) fall on a line of slope one. In altered-cue runs, the correct responses (open squares) are a nonlinear function of $f_n(\ )$. The solid line represents how subjects actually adapt to this nonlinearity.
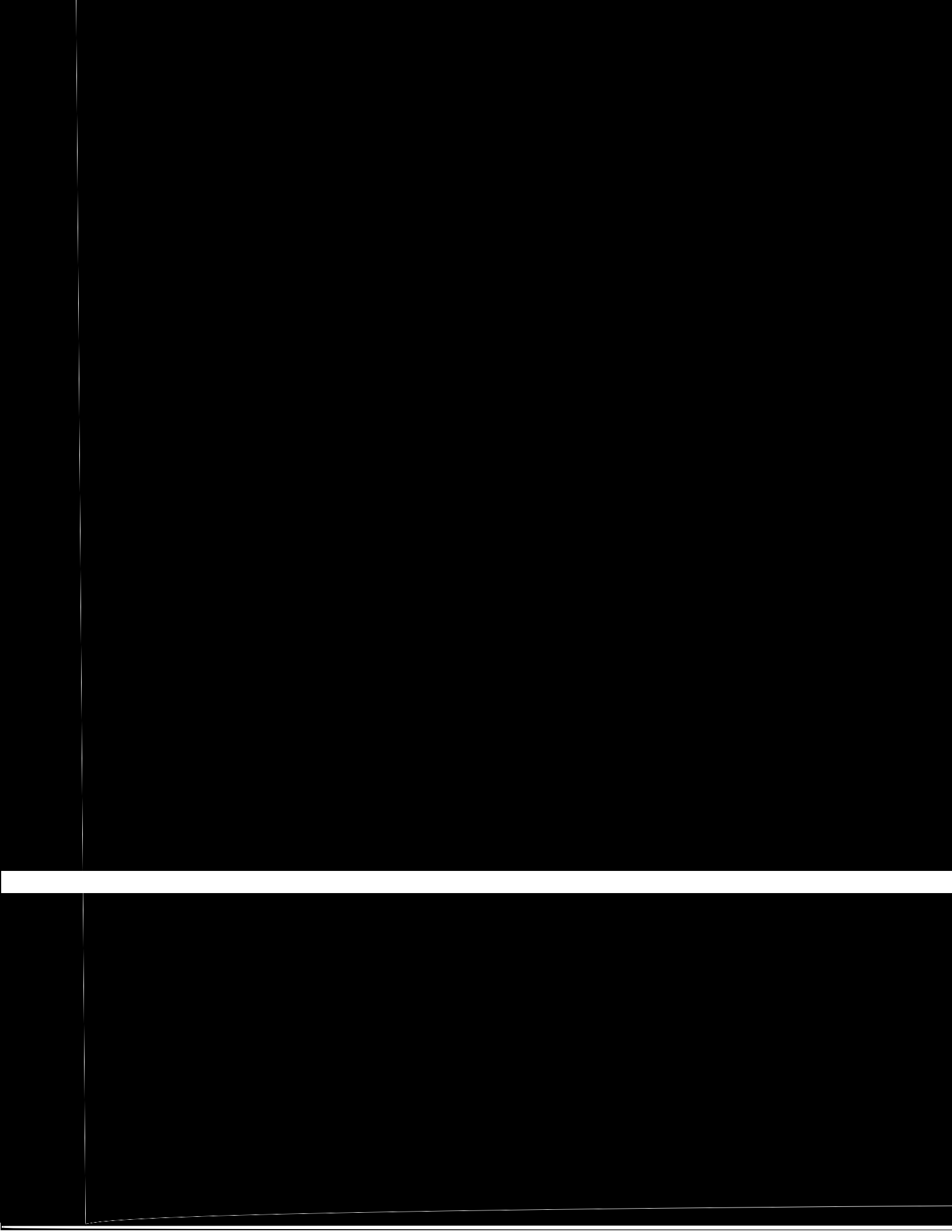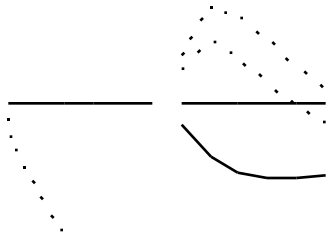
Figure 5. Estimated and predicted bias $_i$ as a function of position. In each panel, the experimental results, plotted on the left side of each panel, were found by assuming that performance was symmetric about the midline. The predictions from the model are shown on the right side of each panel. To the extent that the left and right halves of each panel are symmetric about the line $y = -x$, the model fits the experimental results.

Figure 6. Estimated and predicted resolution d'_i as a function of position. In each panel, the experimental results, plotted on the left side of each panel, were found by assuming that