PNAS RESEARCH ARTICLE | PSYCHOLOGICAL AND COGNITIVE SCIENCES

The perception of auditory motion in sighted and early blind individuals

Woon Ju Park[a,1] and Ione Fine[a]

Edited by Randolph Blake, Vanderbilt University, Nashville, TN; received June 20, 2023; accepted October 29, 2023

Motion perception is a fundamental sensory task that plays a critical evolutionary role. In vision, motion processing is classically described using a motion energy model with spatiotemporally nonseparable filters suited for capturing the smooth continuous changes in spatial position over time afforded by moving objects. However, it is still not clear whether the filters underlying auditory motion discrimination are also continuous motion detectors or infer motion from comparing discrete sound locations over time (spatiotemporally separable). We used a psychophysical reverse correlation paradigm, where participants discriminated the direction of a motion signal in the presence of spatiotemporal noise, to determine whether the filters underlying auditory motion discrimination were spatiotemporally separable or nonseparable. We then examined whether these auditory motion filters were altered as a result of early blindness. We found that both sighted and early blind individuals have separable filters. However, early blind individuals show increased sensitivity to auditory motion, with reduced susceptibility to noise and filters that were more accurate in detecting motion onsets/offsets. Model simulations suggest that this reliance on separable filters is optimal given the limited spatial resolution of auditory input.

auditory motion | blindness | spatiotemporal selectivity | cross-modal plasticity

Significance

Tracking the movement of objects is an environmentally critical task. The human visual system contains neurons that are specialized for tracking continuous changes in space over time. However, it is not clear whether the auditory system has similar specialized motion processing mechanisms. Here, we show that humans process auditory motion very differently from visual motion, inferring motion direction from the location of sounds at their onsets and offsets. This difference between vision and audition is likely the result of our limited ability to pinpoint the location of sounds. Early blind individuals were far more accurate at detecting the motion onset and offset, showing that early blindness results in an enhanced ability to hear moving objects in noisy backgrounds.

Understanding the motion of objects in the environment is a fundamental task of sensory processing. In vision, it is well known that the earliest stage of specialized motion processing, the middle temporal (MT) area, contains mechanisms specifically devoted for processing motion (1, 2)—the neurons have spatiotemporally nonseparable tuning (3), where the selectivity for space systematically changes in relation to time (Fig. 1, *Left*, the contour of the receptive field is diagonally tilted). This permits a dedicated system for detecting continuous object motion over space and time. In contrast, the question of how the human brain processes auditory motion has been debated for over three decades. Specifically, it is not clear whether auditory motion perception is indirectly inferred from "snapshots" of changes in position across different time points, mediated by spatiotemporally separable neurons (Fig. 1, *Right*), or relies on dedicated nonseparable "motion detectors" as in vision.

Previous human psychoacoustical findings have been inconclusive. Some studies have measured the auditory motion aftereffect (MAE), based on the longstanding notion that selective adaptation requires the existence of specialized mechanisms tuned for the adapted sensory feature. While adaptation to auditory motion can produce a measurable MAE, the effect is highly stimulus specific (e.g., does not generalize across frequencies) (4–6) and is much weaker (7) than the adaptation observed using visual motion stimuli (8). In addition, auditory speed judgments tend to be dominated by duration and distance cues, a finding more consistent with separable tuning (9, 10).

The neurophysiological evidence is similarly inconclusive. Some animal studies do find direction-selective cells (11–13), but their tuning seems to result from selectivity to acoustic features associated with motion stimuli, rather than selectivity to motion per se (14–16). Human fMRI studies have consistently shown greater responses to auditory motion as compared to static stimuli in the right planum temporale (17–21). However, the magnitude of fMRI responses in this region to auditory motion is comparable to static stimuli that randomly change location over time (22, 23), suggesting that planum temporale may encode changes in sound position rather than genuine auditory motion.

Here, we used a psychophysical reverse correlation paradigm to estimate the shape of the "filters" used for discriminating auditory motion. Psychophysical reverse correlation paradigms have been used to characterize the mechanisms underlying a wide variety of auditory (24) and visual tasks (25–28), and the "perceptive" field properties estimated from these paradigms have shown remarkable similarities to receptive fields of single neurons (29). We chose as our paradigm a direct analogue of a previous study that observed clear spatiotemporal nonseparability in the perceptual filters for discriminating visual

Author affiliations: [a]Department of Psychology, University of Washington, Seattle, WA 98195

Preprint server: This manuscript has been deposited as a preprint on bioRxiv (10.1101/2022.09.11.507447).

Author contributions: W.J.P. and I.F. designed research; W.J.P. performed research; W.J.P. analyzed data; and W.J.P. and I.F. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: wjpark@uw.edu.

Published November 28, 2023.

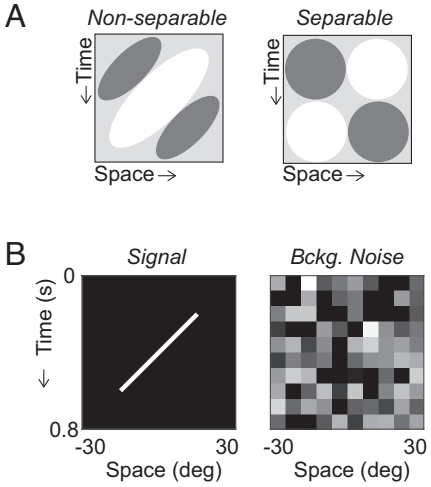

Fig. 1. Spatiotemporal separability and stimuli used in the experiment. (*A*) Spatiotemporally nonseparable filters (*Left*) are motion selective, with the contour of the receptive field tilted with respect to the coordinate axes. The maximal activation of the filter occurs for stimuli whose location in space and time can be described as $v = x/t$, where v is velocity, t is time, and x is spatial location. In contrast, spatiotemporally separable filters (*Right*) can be described as the outer product of separate 1D filters in space and time. (*B*) Participants were asked to discriminate the direction of signal motion (broadband noise, left vs. right, leftward motion depicted in the *Left*). Spatiotemporal background noise (*Right*) was added to the moving signal. The intensity in both panels represents sound amplitude [scale = (0 1), from black to white].

motion direction, which closely resembled the known spatiotemporal selectivity of neurons within area MT (30).

Participants discriminated the direction (left vs. right) of a near-threshold auditory motion signal embedded in bursts of static auditory noise presented over space and time (Fig. 1*B*). The task was designed to examine the perception of auditory motion in a relatively ecologically valid setting—tracking an object's motion within the context of a noisy environmental background. Sounds usually occur in the presence of significant background noise, and thus, one of the key challenges that the auditory system must solve is separating sounds of interest from other sound sources. For each participant, we simultaneously a) measured the threshold for reliably discriminating the signal motion direction and b) estimated the perceptual filter for discriminating auditory motion direction based on the properties of background noise. By characterizing how the spatiotemporal structure of the auditory noise systematically influences the task decision, we extracted the spatiotemporal features that contribute to the perception of auditory motion direction.

We compared the ability to discriminate auditory motion direction in both sighted and early blind individuals (Table 1). Understanding object motion based entirely on auditory information is a crucial skill for which blind individuals presumably develop considerable expertise, yet little is known about how auditory motion processing is affected by early blindness. One previous study showed enhanced performance on an auditory motion task as a function of early blindness (31); however, this was a highly complex task requiring judgement of the overall motion direction from multiple incoherent moving sound sources. A second recent study, examining speed processing, found that early blind individuals were, if anything, less sensitive to speed and relied predominantly on stimulus duration when making speed judgments (32). Our reverse correlation paradigm allows us not only to measure performance but also to examine how differences in performance between early blind and sighted individuals are mediated by changes in the perceptual filters underlying auditory motion discrimination.

Results

Early Blind Individuals Can Discriminate Auditory Motion Better than Sighted Individuals. Fig. 2*A* shows the signal motion amplitude that resulted in 65% accuracy, for both sighted and early blind individuals across all 6 sessions. Overall, sighted individuals required signal motion at 10 dB louder than the mean background noise level (approximately three times louder). In contrast, early blind individuals could perform the task at significantly lower signal amplitude ~6 dB [repeated measures ANOVA, $F(1, 14) = 30.70$, $P < 0.001$]; thresholds that were approximately half those of sighted individuals. Across sessions (Fig. 2*B*), the performance remained relatively constant in both groups [main effect of session: $F(5, 70) = 2.24$, $P = 0.14$], and the size of the group difference did not change [session x group interaction: $F(5, 70) = 1.56$, $P = 0.18$]. This suggests that the perceptual advantages accrued in early blind individuals require either loss of vision early in life or

Table 1. Participant characteristics

Group	Sex	Age	Blindness onset	Cause of blindness	Light perception
Sighted 1	F	62	N/A	N/A	N/A
Sighted 2	M	35			
Sighted 3	M	34			
Sighted 4	F	62			
Sighted 5	M	73			
Sighted 6	M	64			
Sighted 7	M	61			
Sighted 8	M	47			
Early blind 1	M	55	Born blind	Congenital glaucoma	No
Early blind 2	M	28	Age 2	Retinoblastoma	No
Early blind 3	M	36	Born blind	Leber's congenital amaurosis	No
Early blind 4	F	59	Born blind	Retinopathy of prematurity	No
Early blind 5	M	69	Born blind	Retinopathy of prematurity	No
Early blind 6	F	58	Age 2	Optic atrophy	No
Early blind 7	F	68	Born blind	Retinopathy of prematurity	Low
Early blind 8	M	50	Born blind	Optic nerve hypoplasia	Low

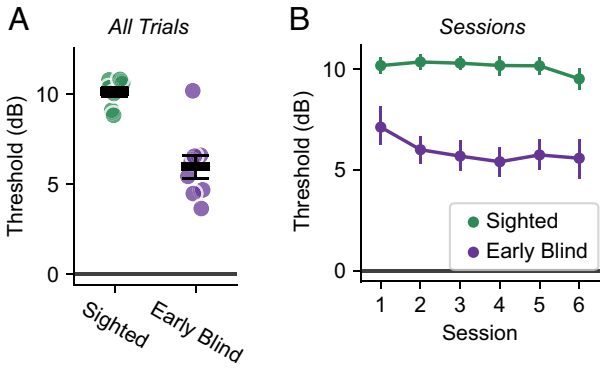

Fig. 2. Early blind individuals can identify the direction of motion at lower levels of signal amplitude than sighted individuals. (*A*) Thresholds based on the data from all six sessions (1,000 trials/session, 6,000 trials total), with colored circles showing individual thresholds and black horizontal bars showing the group mean. (*B*) Thresholds across individual sessions with colored circles depicting the group mean. Error bars are SEM.

considerable long-term experience—differences between the groups were barely reduced after 5,000 trials of practice.

Both Sighted and Early Blind Individuals Have Separable Auditory Motion Filters. Fig. 3*A* shows estimated perceptual auditory motion filters for individual sighted and early blind participants. The average for each group (the mean of individual estimated filters) is shown in Fig. 3*B*. It is worth noting that these filters are independent of the signal motion threshold discussed above; they were constructed entirely based on the properties of the external background noise, while a staircase procedure kept the signal motion barely audible at ~65% correct. The filters for rightward trials were flipped over the spatial axis, so the estimated filters are aligned for the leftward motion direction.

Consistent with the "snapshot" model of auditory motion processing, both sighted and early blind individuals appear to use separable filters to perform our task. To quantify this, we created a model with a parameter that characterizes the "separability" in the estimated perceptual filters. The model has a separable surface (Fig. 4 *A*, *Top*, I_1) and a central surface (I_2). The parameter a controls how much central surface should be added to the separable surface to describe an individual's perceptive filter. When $a = 0$, the filter is entirely separable, while $a = 1$ produces a classic spatiotemporal filter typical of visual motion processing (Fig. 4*A*, *y* axis). We found the best-fitting value of the parameter a for each participant's perceptual filter. As can be seen in Fig. 4*A*, estimates of a for all participants are clustered around 0, implying clear spatiotemporal separability. Estimates of a were not significantly different between the two groups ($U = 40$, $P = 0.44$, Wilcoxon rank-sum test). The fact that the filters were equally separable in both groups suggests that the influence of early blindness/auditory experience on the shape of the auditory motion filters (i.e., separable vs. nonseparable) is minimal.

In all participants, the separable model provided a better fit (as determined by the cross-correlation between the model and observed filter) than the nonseparable model [Fig. 4*B*; repeated measures ANOVA, main effect of model: $F(1, 14) = 46.60$, $P < 0.001$]. Overall, the difference in fit between separable and nonseparable models was significant in both groups [post hoc t test, sighted: $t(7) = -2.96$, $P = 0.021$; early blind: $t(7) = -6.61$, $P < 0.001$] but was larger in the early blind group [model x group interaction: $F(1, 14) = 7.57$, $P = 0.016$].

Early Blindness Refines Auditory Motion Filters. What then explains the enhanced performance in early blind individuals? Fig. 3 *C*, *Left*, shows differences in weighting across space and time between sighted and early blind individuals. A Hotelling T^2 analysis (25) revealed a significant group difference in the measured filters [$T^2 = 256.52$, $F(100, 46{,}300) = 2.56$, $P < 0.001$]. To visualize which regions within the filter likely contributed the most to this statistical difference, we ran a nonparametric permutation test (uncorrected for multiple comparisons; see *Materials and Methods*). In Fig. 3 *C*, *Right*, the red cells represent regions in space and time that contributed significantly more to leftward directional judgments in early blind than sighted individuals, while blue cells represent regions that contributed significantly less.

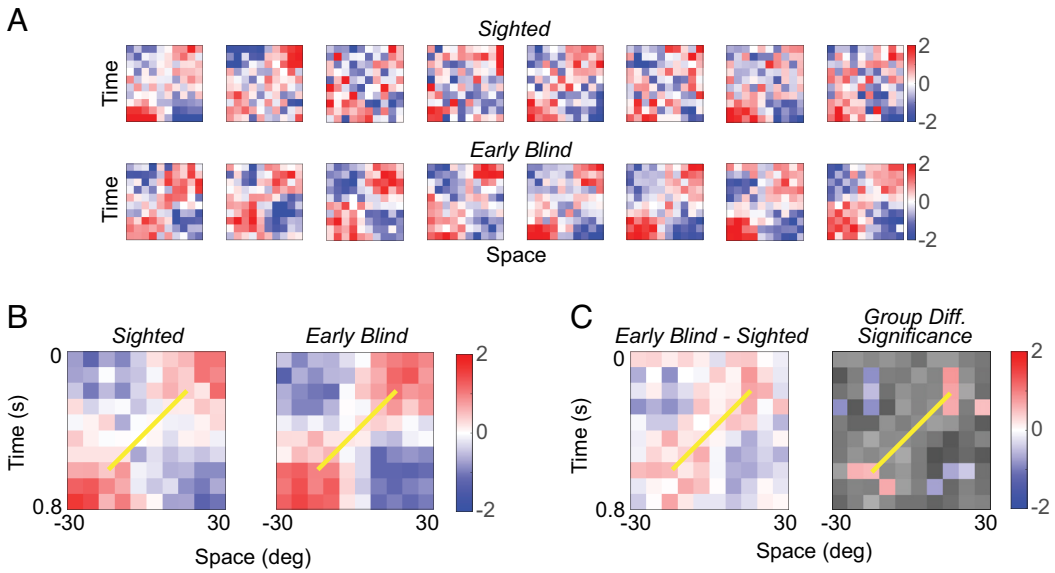

Fig. 3. (*A*) Estimated perceptual auditory motion filters for sighted (*Top* row) and early blind (*Bottom* row) individuals. The spatial axis is mirror-inverted on trials with rightward signals. (*B*) Group averaged filters for sighted (*Left*) and early blind (*Right*) individuals. Yellow lines depict the signal motion. (*C*) Differences in perceptual auditory motion filters between early blind and sighted groups (*Left*). In the *Right* panel, the colored regions show the locations in space and time with weights that were significantly different between the two groups as determined by a permutation test (uncorrected for multiple comparisons; see *Materials and Methods*). The regions in the filter without a significant group difference are converted to grayscale [scale = [−2 2], from black to white]. The intensity in each cell in all figures represents the Z score, measuring the influence of background noise at that temporal and spatial location on motion judgments.

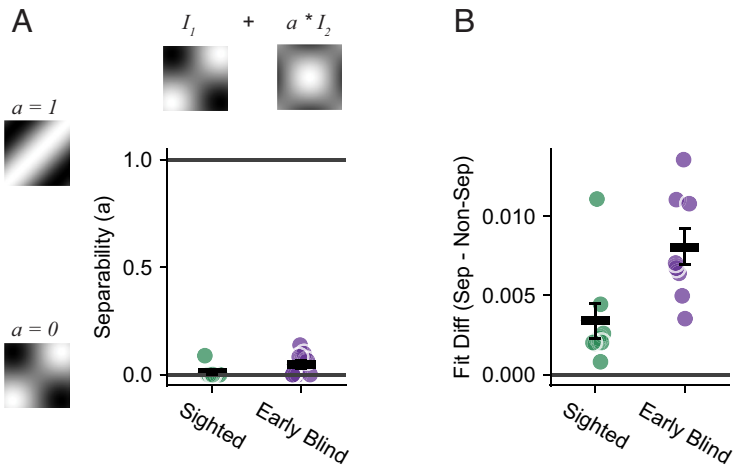

Fig. 4. Spatiotemporal separability in measured filters. (*A*) Quantification of spatiotemporal separability in the measured perceptual filters (see *Materials and Methods* for details). (*Top*) The parameter *a* estimates how much central surface (I_2) should be added to the separable model (I_1) to best explain the shape of each participant's perceptual filter. As shown along the *y* axis, *a = 0* results in a separable filter, and *a = 1* results in a nonseparable filter. Image intensity for surface and filter inserts are scaled from −1 (black) to 1 (white). (*Bottom*) Colored circles represent the estimated separability parameter (*a*) for each participant. (*B*) Model fits quantifying how well the separable model (*a* = 0) describes the measured filters for each participant (colored circles) compared to the nonseparable model (*a* = 1). A difference value greater than 0 suggests a better fit for the separable model. Black horizontal bars show group means with error bars showing the SEM.

In our task, the background noise slightly preceded the onset of the signal motion (shown by a yellow line in Fig. 3 *B* and *C*) and extended over a slightly larger spatial range. In sighted individuals, the filter peaks appeared to be tuned to the onset-offset of the stimulus—regions of the stimulus that consisted entirely of background noise. In contrast, most early blind individuals appeared to have the peak of their filters shifted toward the signal onset and offset.

To quantify this shift in perceptual filter tuning, we fitted a separable Gaussian model with five free parameters that described each participant's estimated filter based on the center and width of spatial and temporal tuning and filter gain (Fig. 5*A*; see *Materials and Methods*).

We began by examining the ability of this model to explain the measured filters observed in individual participants. As shown in Fig. 5*B*, we found that each participant's predicted filter (as described by the best-fitting parameters of Fig. 5*A*) was well matched to their actual filter. The mean correlation coefficient between predicted and observed filter weights was 0.79 for early blind (*SD* = 0.07) and 0.58 for sighted individuals (*SD* = 0.17).

Parameter estimates (Fig. 5*C*) showed significant differences in spatiotemporal tuning in early blind compared to sighted individuals, consistent with the qualitative differences observed between the filters in Fig. 3*C*. First, we observed a significant group difference in the center parameter for space [$t(14) = -2.4$, $P = 0.031$], where the center of the Gaussian tuning was shifted toward the signal onset (yellow arrow; as compared to the stimulus onset, gray arrow) in early blind individuals. The center parameter for temporal tuning was also shifted more toward the signal onset in early blind individuals, which was marginally significant [$t(14) = 1.93$, $P = 0.074$]. Consequently, filter peak coordinates (third panel of Fig. 5*C*) were more closely aligned with the spatial and temporal onset of the signal motion (yellow triangle) for early blind individuals, while filter peaks for the sighted were closer to the start of the background noise (gray triangle). These results suggest that the filters of early blind individuals enabled more accurate extraction of the signal onset from the noise. Spatial and temporal width parameters did not show a significant group difference [spatial width: $t(14) = 1.19$, $P = 0.26$; temporal width: $t(14) = 1.39$, $P = 0.19$].

Finally, the gain parameter of the model was also larger in early blind than sighted individuals [$t(14) = 3.72$, $P = 0.002$], consistent with the idea that early blind individuals have either larger gain or lower internal noise in their auditory motion processing (*Discussion*).

Next, we examined whether differences in filter properties, fitted to the measured filters, could explain individual perceptual thresholds (Fig. 6*A*). We simulated predicted performance for each participant based on their individual filters constructed using parameter estimates in Fig. 5*C* (*Materials and Methods*). These filters, despite being based entirely on the spatiotemporal background noise, predicted individual perceptual thresholds (the signal amplitude that resulted in 65% correct performance, Fig. 6*A*) remarkably well in the early blind group; there was a significant correlation between observed and predicted thresholds [$r(6) = 0.89$, $P = 0.003$]. In the sighted group, the correlation was not statistically significant across individuals [$r(6) = 0.41$, $P = 0.32$], but the model successfully predicted overall worse group performance.

To understand which of the parameters in the model contributed the most to behavioral performance, we correlated tuning properties of the filters with measured perceptual thresholds (Fig. 6*B*). In early blind individuals, spatial (*Left*) and temporal (*Center*) filter peaks were significantly correlated with perceptual thresholds [space: $r(6) = 0.74$, $P = 0.038$; time: $r(6) = -0.76$, $P = 0.028$]—as expected, peaks that were closer to the signal onset (yellow arrow) were associated with better performance. We also found a significant correlation between filter gain and thresholds in early blind individuals [$r(6) = -0.75$, $P = 0.03$], with higher filter gains predicting better performance (*Right*). These results suggest that the shift in spatiotemporal tuning and increased gain in early blind individuals' filters explain their improved behavioral performance.

For sighted individuals, the peak coordinate on the temporal axis was significantly related to thresholds [$r(6) = -0.8$, $P = 0.018$], again showing that the estimated filter properties are related to individuals' auditory motion perception. None of the other correlation analyses showed a significant relationship (all *P*'s > 0.25).

Separable Filters May Be a Consequence of Broad Auditory Spatial Tuning. Our results thus far demonstrate that auditory motion discrimination likely relies on spatiotemporally separable

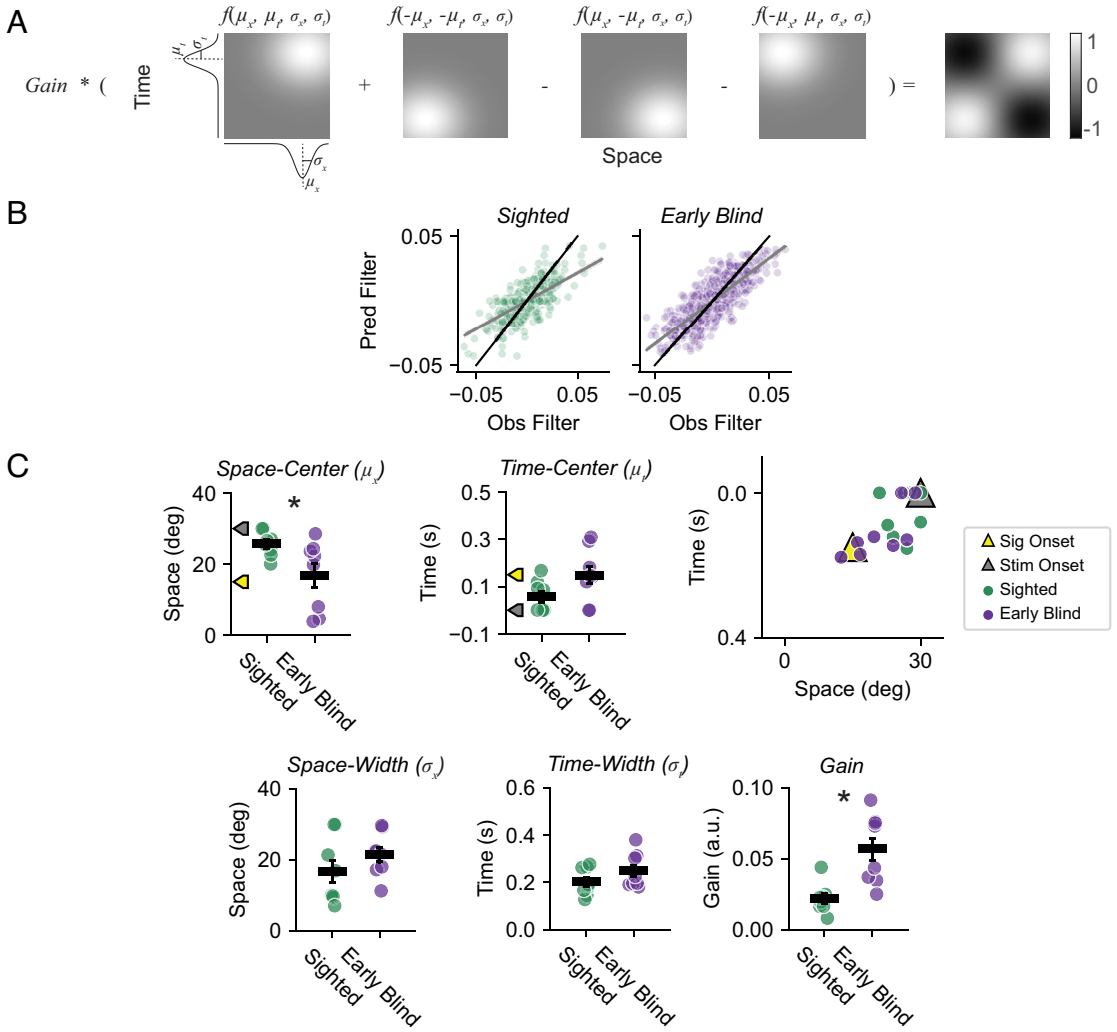

Fig. 5. Quantification of filter properties. (*A*) Diagram illustrating a separable Gaussian model with five free parameters. (*B*) Model predictions of spatiotemporal filters. Each dot represents a cell in the spatiotemporal filter in an individual participant. Black and gray lines are identity line and regression fits across all data points, respectively. The plots show data from all participants for visualization. (*C*) Parameter estimates. Yellow and gray arrows indicate the onset of signal motion and stimulus (i.e., background noise), respectively. Colored circles show individual data and black horizontal bars show group means. Error bars are SEM. Compared to the sighted group, the filters of early blind individuals had center tuning shifted toward the location of the signal onset (first and second panels, *Top* row). Consequently, the spatiotemporal peak coordinates (third panel, *Top* row) of early blind participants are closer to the signal onset (yellow triangle) than those of sighted participants. The early blind group also showed greater gain (third panel, *Bottom* row).

filters, which can be refined based on early blindness and/or extensive auditory experience. Why is it that visual direction discrimination relies on nonseparable filters, while that same task, in the auditory domain, relies on separable filters? One major difference between the two modalities is that visual spatial tuning is exquisitely precise (individuals can detect differences in spatial location of less than a degree) whereas auditory spatial tuning is quite coarse (sensitive to shift in spatial location of ~2 to 3 degrees) (33), see *Discussion*. To examine whether this difference in spatial tuning width accounts for the shift from nonseparable to separable tuning, we carried out simulations comparing the predicted performance of separable and nonseparable filters on our experimental task as a function of tuning width (Fig. 7).

We compared the performance of a separable model to an ideal observer with internal noise and 7 nonseparable models with varying tuning (*Materials and Methods*). On a given trial, an ideal observer chooses the motion direction that maximizes the posterior probability of the direction given received input. In our task, the optimal strategy for the ideal observer is simply to maximize the cross-correlation between the received input (stimulus + internal noise) and the signal template (blue outlined insert in Fig. 7; refs. 34 and 35). The performance of the separable and nonseparable models was similarly simulated by taking the cross-correlation between the received input and the constructed models as "templates". The separable model assumed Gaussian tuning over space and time (similar to that in Fig. 5*A*) perfectly aligned with onset/offset of the signal motion, with widths that were the average of the width parameters estimated across all participants (green outlined insert). The nonseparable models were oriented Gaussians with width simulated at 7 different levels (gray outlined inserts).

As expected, the ideal observer showed the best performance (blue curve). Among the nonseparable models (gray curves), the most narrowly tuned model (lightest gray curve), that closely resembles the ideal observer, performed the best. However, as broader spatial tuning was assumed, nonseparable models (darker gray curves) quickly began to show worse performance than the separable model (green curve). Notably, the separable model, which assumed spatial width of 0.5 (normalized unit) outperformed the nonseparable model with that of 0.43. Thus, spatiotemporally separable filters seem to be the optimal solution when the incoming sensory information has broad spatial resolution.

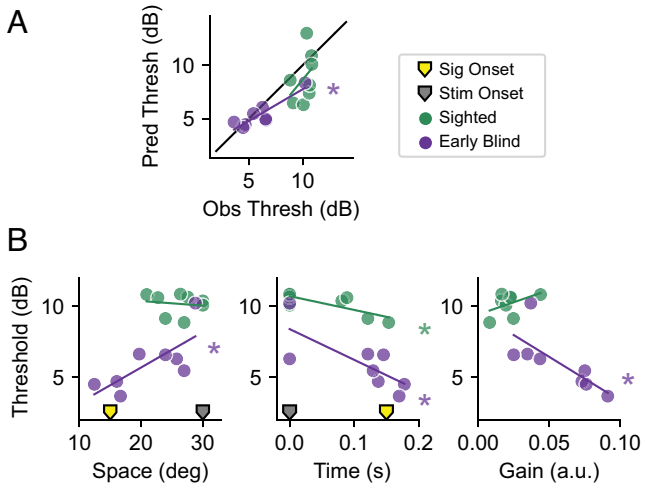

Fig. 6. Relationship between filter properties and performance. (*A*) Model predictions of individual perceptual thresholds. (*B*) Results from the correlation analyses examining the relationship between estimated filter properties (space peak, time peak, and gain from left to right) and perceptual thresholds. Colored lines show linear regression fits.

Discussion

Here, we used a psychophysical reverse correlation paradigm, where participants discriminated the direction of a motion signal in the presence of spatiotemporal noise, to characterize the filters underlying auditory motion discrimination in sighted and early blind individuals.

Auditory Motion Perception Is Mediated by Separable Filters. We found, for both sighted and early blind individuals, that our task of auditory motion discrimination within background noise was mediated by separable filters. Our results are very different from those previously observed using the visual analog of our auditory task (30), which demonstrated a reliance on spatiotemporally nonseparable filters for perceiving visual motion. Thus, direction discrimination is carried out very differently across auditory and visual modalities.

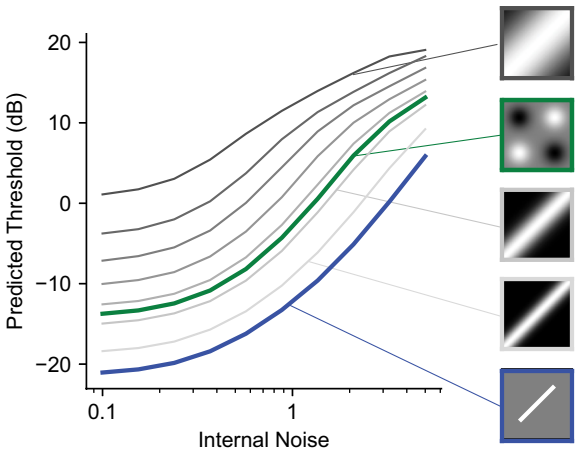

Fig. 7. Performance of an ideal observer model simulating the optimal strategy (blue) compared with a set of nonseparable models with varying tuning width from narrow (lighter gray) to broad (darker gray) as well as a separable model constructed using the average spatial tuning from our psychophysical data (green). The separable model appears to be optimal when broad spatial tuning is assumed.

This difference in tuning between auditory and visual motion perception is likely driven by the statistics of the available sensory input. In vision, it is well-established that the resolution for encoding space is exceptionally good: neurons directly represent each point in retinal space with a resolution as small as 2 to 5 arcsec (36, 37). Thresholds for detecting visual displacements of successively presented static stimuli can be as low as 5 min (38, 39).

In contrast to visual space, which is explicitly represented in retinotopic maps, the sense of auditory space is inferred from multiple sources of information including interaural time and level differences. Auditory spatial tuning is broad within both the primary auditory cortex and planum temporale (40, 41), and neural tuning to sound-source locations seems to be represented by an opponent process, based on differences in the activity of two broadly tuned channels formed by contra- and ipsilaterally preferring neurons (41–43). Free-field localization of spectrally rich auditory stimuli produces discrimination thresholds much larger than those of vision, on the order of 1° (44, 45). Thus, spatial resolution tends to be much higher for vision than audition (46).

The computational implementation of our auditory version of the task is identical to the previous study of visual motion (30), which found nonseparable filters. There was no difference between our auditory and their visual tasks except for the differences in motion direction (left/right vs. up/down) and the spatial and temporal scale between the auditory and visual stimuli. Our simulations suggest that the use of separable filters for the auditory version of the task and nonseparable filters for the visual version of the task is driven by the difference in spatial resolution between the two sensory modalities. When spatial localization is poor, as is the case for auditory information, separable filters may be more optimal. It remains to be seen whether visual motion discrimination might shift to separable filters under conditions where spatial resolution is poor.

Auditory Motion Processing Is Enhanced in Early Blind Individuals. Our results also demonstrate how the spatiotemporal tuning of auditory motion filters is altered by early blindness. Early blind individuals showed increased sensitivity to auditory motion, with significantly lower signal motion amplitudes required to discriminate the direction of motion at 65% accuracy. This enhanced performance could successfully be predicted from their (independently) measured filter properties: filters of early blind individuals had larger amplitudes and were more accurate at detecting motion onsets/offsets in the context of background noise.

Our results are consistent with one previous study that showed enhanced performance for early blind individuals on a task requiring judgement of the overall motion direction from multiple incoherent moving sound sources (31). A second recent study, examining speed processing, found that early blind individuals relied more on stimulus duration when making speed judgments (32), which led to worse performance when spatial cues were present. One likely explanation for the difference between our study and theirs is that in our study the background noise may have acted as an "auditory reference frame". Several studies have shown that early blindness enhances spatial discrimination (47), but the absence of calibrating visual information results in an impaired allocentric representation of auditory space (48, 49).

The larger gain of the filters in our early blind group can be interpreted in two ways. One possibility is that this parameter reflects neural mechanisms that are more narrowly and "cleanly" tuned. Previous studies in the vision literature suggest that broadly tuned templates can lead to decreased efficiency in filtering out external noise, and the use of less well-tuned filters can result in

noisier behavioral responses (50–54). Thus, the more refined tuning of the early blind individuals' filters may have manifested itself as increased gain in the filter estimates. A nonexclusive alternative is that the larger gain in blind individuals might reflect a reduction in internal noise (occurring either before or after the filter itself), which would effectively increase the gain in neural responses to the signal. In our experiment, it was impossible to distinguish between these two accounts, but future work, using a double pass method (55) for instance, could differentiate these two models and provide insights into why our model was less successful in predicting the data of sighted individuals.

The mechanisms underlying the refinement of auditory motion filters in early blind individuals is still unclear. One possibility is that this refinement is mediated by extensive auditory experience. Audition is the only source of sensory input that can provide information about distant space other than vision. Early blind individuals rely heavily on auditory information to navigate and understand where objects are moving in the environment, which might result in experience-based plasticity in their auditory motion processing. There is analogous evidence that musicians also have enhanced basic auditory abilities (56). However, in the case of musicians, it is not clear whether these results reflect extensive training or differences in innate ability. Our findings certainly suggest that the differences we see in early blind individuals require extensive experience—in our study, there was little change in either early blind or sighted individuals' performance over 6,000 trials (~6 h) of practice.

A second possibility is that the alterations in auditory motion processing we observed are driven by sensory reorganization that takes place due to visual deprivation, including cross-modal plasticity in the brain. There is now a large amount of evidence suggesting that human visual motion area hMT+ responds to auditory motion in early blind individuals (57–60) and this region may even "co-opt" the function of right planum temporale (31, 58, 61), a region associated with auditory motion processing in sighted individuals. This raises the possibility that the enhanced spatiotemporal tuning in auditory motion filters of early blind individuals may be mediated by the recruitment of deprived visual motion area, hMT+.

The finding of separable filters in early blind individuals, in this context, is somewhat surprising. Given that hMT+ has been shown to have nonseparable spatiotemporal filters for processing visual motion, if auditory motion is processed within hMT+ in early blind individuals, then this must result in a shift from nonseparable to separable templates—a significant modification of hMT+ normal computational operations.

Alternatively, it is possible that the auditory motion responses observed in hMT+ do not reflect motion computations within hMT+ per se, but instead represent spatial and temporal signals from either subcortical [perhaps cross-modally recruited superior colliculus (62)], or cortical (63) areas. If so, the ability to classify direction of auditory motion in hMT+ might be mediated by properties other than neuron-level spatiotemporal tuning. One possibility is that these signals reflect recruitment of the retinotopic organization in hMT+. Retinotopic information from hMT+ might play an important role in mediating the perceptual experience of motion, including propagating information to other areas in the brain. MT provides motion information to numerous cortical areas responsible for navigating and interacting with the 3D world: It has reciprocal projections (64) to a variety of sensorimotor areas including parietal V6 and V6A [object motion recognition and control of reach-to-grasp movements (65, 66)], AIP [visuo-motor transformations for grasp (67)], MIP [coordination of hand movements and visual targets (68)], LIP (saccadic target selection), frontal A4ab (motor cortex), prefrontal A8aV (frontal eye fields), and A8C (premotor).

Auditory Motion Perception in the Presence of Background Noise. Our study has implications on our ability to segregate moving sounds embedded in background noise. Tracking object motion in noisy environment is an ecologically important task. Sounds almost never occur in isolation, and separating the sounds of interest from the background noise of other sound sources is difficult (64–68). Our results in sighted individuals suggest that they failed to successfully recruit mechanisms that were tuned to the signal as compared to the background noise. In contrast, our findings in early blind individuals suggest that early blindness or extensive auditory experience improves the ability to segregate signal from noise, enhancing the ability to hear auditory motion in the presence of background noise (64).

Conclusions

Here, we provide direct evidence that both sighted and early blind individuals perceptually experience auditory motion using spatiotemporally separable filters. This contrasts with visual motion processing, which uses nonseparable filters tuned for continuous object motion and provides an elegant example of how sensory input statistics constrain the encoding of object motion across different sensory modalities. Our results further suggest that within the general constraints afforded by the statistics of the input, altered auditory experience or neural reorganization due to early blindness significantly refines auditory motion processing.

Materials and Methods

This study was approved by the University of Washington's Institutional Review Board and carried out in accordance with the Code of Ethics of the Declaration of Helsinki. Informed written consent was obtained from all participants prior to conducting the experiments.

Participants. Participants included 8 human adult listeners with neurotypical visual and auditory histories and 8 early blind individuals with neurotypical auditory histories (Table 1). The two groups were individually matched for age. All participants reported normal hearing and no history of psychiatric illness.

Stimulus. Auditory stimuli were delivered through Etymotic ER-2 insert earphones at a sampling rate of 44,100 Hz. 3D auditory space was simulated based on a basic physics model, using interaural level and time differences, and the decrease in volume as a function of the distance to the observer. The stimuli were generated and presented using MATLAB and Psychtoolbox (69).

The stimuli consisted of two components: signal motion and background noise bursts. Both consisted of broadband noise, created by generating Gaussian noise in the time domain, which was then bandpass filtered between 500 and 14,000 Hz in the Fourier domain (fast Fourier transform) and was projected back to the time domain (inverse fast Fourier transform).

The signal motion continuously traveled either leftward or rightward (Fig. 1 *B*, *Left*) on a given trial. We simulated a constant-velocity stimulus traveling from ±15° from the observer along a frontoparallel direction (a straight-line oriented perpendicular to the listener's facing direction) at a distance of 0.8 m, centered at the midline. The signal motion lasted 500 ms and had a linearly ramped onset and offset of 50 ms. The amplitude of the signal motion was adjusted throughout the experiment using two interleaved QUEST staircases which held performance accuracy at approximately 65%.

The background noise bursts were discrete sounds simulated on a 10 × 10 grid (Fig. 1 *B*, *Right*) that spanned space (±30°) and time (0 to 800 ms). At each moment in time, there were 10 simultaneous noise bursts (of varied amplitude),

one for each location in space. The amplitude of the noise bursts at each spatiotemporal location within the space-time grid was randomly and independently selected from a Gaussian distribution, resulting in an approximate range of 5 to 49 dB (mean = 39 dB).

Task. On each trial, the observers reported whether the auditory motion stimulus was moving leftward or rightward via a button press and were given auditory feedback (a brief beep if they were correct). At the end of each block (200 trials), there was a brief mandatory rest period of 30 s. Participants could then press any key to begin the next block of trials. Each session contained 1,000 trials (5 blocks), and each participant carried out 6 sessions. The first block was discarded from the analyses to limit early practice effects.

Analysis.

Derivation of signal motion thresholds. Each participant's threshold to discriminate the signal motion was estimated by fitting a Weibull function using the psignifit toolbox (70) to all trials from each staircase within a session. This resulted in 12 threshold measurements (2 thresholds per session), which were then averaged.

Derivation of spatiotemporal perceptual filters. The perceptual filters for hearing auditory motion were estimated for each participant, solely using the characteristics of background noise bursts, independent from the changes in signal motion amplitude. The filters were derived based on a previous study (30), which used an analogous paradigm for estimating visual motion filters in visually typical observers. Here, $N^{[q,z]}$ denotes the noise sample over space (x) and time (t), and $\langle\rangle$ represents the average across all trials. Filters were constructed by sorting trials into four categories based on the direction of the signal direction [left ($q = 1$) or right ($q = 0$)] and participants' response [correct ($z = 1$) or not ($z = 0$)]. The resulting filter (F) is derived:

$$F = \langle N^{[1,1]}(x,t)\rangle - \langle N^{[1,0]}(x,t)\rangle + \langle N^{[0,1]}(-x,t)\rangle - \langle N^{[0,0]}(-x,t)\rangle,$$

where the spatial axis is mirror-inverted ($-x$) for noise samples on trials with rightward signals, to align the filter orientations.

Comparison of separable and nonseparable models. We assessed whether human listeners use spatiotemporally separable or nonseparable models for hearing auditory motion in two ways.

We first examined spatiotemporal separability using a model that could directly characterize how "separable" the filters are in space and time. Specifically, two separate surfaces (I_1, I_2) were defined (Fig. 4*A*):

$$I_1 = -\sin(2\pi fx) \otimes \sin(2\pi ft)$$
$$I_2 = \cos(2\pi fx) \otimes \cos(2\pi ft),$$

where f is spatial frequency. The linear combination of these two surfaces, $W = \beta(I_1 + aI_2)$, ranges between a separable surface (I_1) when a is close to 0 and a nonseparable oriented grating when a is close to 1 as shown in Fig. 4*A* (y axis). Thus, the parameter a provides a measure of spatiotemporal separability that can be fit to each participant's measured perceptual filters (F). Parameter β adjusts the overall gain of the resulting filter. For each individual, we found the value of a that minimized the mean square distance between the model filter and the measured perceptual filter for that individual.

We also separately fitted the separable ($a = 0$) and nonseparable ($a = 1$) filters to each participant's estimated perceptual filters. Goodness of fit between the models, as determined by the dot product of the predicted and the measured filters for each individual, was used to determine whether the separable or nonseparable model explained the measured filters better.

Characterizing effects of experience on auditory motion filters. To test whether early blindness alters auditory motion filters, we carried out a Hotelling T^2 analysis (essentially a multivariate t test) on feature vectors that consisted of background noise images weighted by participants' responses (correct or not) (25).

To visualize the regions of the filter that likely contributed the most to the group difference, we ran a separate permutation test. Here, we assumed that the filters are mirror-symmetric along the spatial axis and the filter estimates from the left (mirror-inverted on x axis) and right halves were subtracted from each other. We then ran 10,000 permutations in which we randomly assigned a group identity (sighted or early blind) to each participant's filter (maintaining the correct number of participants in each group) and calculated the difference between blind and sighted filters for each permutation. We then compared the distribution of differences between these permuted averaged filters to the experimental difference between averaged sighted and early blind filters. Cells within the space-time grid that had values outside the 95% percentile (two-sided) of the null distribution were considered significantly different between the two groups.

Next, to examine the tuning of the estimated filters, we fit a separable model (Fig. 5*A*) that assumes Gaussian tuning in both space and time to the measured perceptual filters (F). The model had five free parameters: center and width for spatial tuning, center and width for temporal tuning, and gain. Parameters were estimated for each participant and were statistically compared between the groups.

To test whether the differences in auditory motion filters explain behavioral performance, we ran a simulation of our experimental paradigm using the separable Gaussian model above with fitted parameters to predict individual perceptual thresholds. Here, we will treat the signal and noisy observations as 1D vectors for simplicity. The response, r_i, of each filter, $\boldsymbol{w_i}$, to each stimulus, which is a combination of signal ($\boldsymbol{s}$) and external background noise ($\boldsymbol{n_{ext}}$), was calculated as the dot product of the stimulus and the filter with added zero-mean normally distributed internal noise, N, with SD proportional to the square root of the strength of the filter response:

$$r_i = (\boldsymbol{s}+\boldsymbol{n_{ext}})^T\boldsymbol{w_i} + kN\left(0,\ \sqrt{((\boldsymbol{s}+\boldsymbol{n_{ext}})^T\boldsymbol{w_i})^2}\right),$$

where i is motion direction. We assumed a correct response when the appropriately oriented filter (e.g., leftward) had a larger response than the opposite (e.g., rightward) tuned filter. The amplitude of the signal motion was varied on each trial to find the perceptual threshold. This was repeated 1,000 times to obtain an average predicted perceptual threshold at a fixed internal noise level (k) of 1.5.

To understand the explanatory power of each of the parameters in the separable model in terms of predicting behavior, we correlated estimated model parameters with participants' perceptual thresholds for discriminating signal motion. These correlation analyses were performed separately for each group.

Simulated comparisons of separable vs. nonseparable filters. Next, we simulated the performance of a variety of separable and nonseparable filters, including an ideal observer. For our stimuli and task, the ideal strategy that yields optimal performance is to simply make a response that maximizes the cross-correlation between the received input and expected direction "template" (34, 35). Here, the received input ($\boldsymbol{g}$) is a combination of $\boldsymbol{s}$ and $\boldsymbol{n_{ext}}$ with added zero-mean white Gaussian internal noise ($\boldsymbol{n_{int}}$): $\boldsymbol{g} = \boldsymbol{s} + \boldsymbol{n_{ext}} + \boldsymbol{n_{int}}$. The SD of the internal noise is proportional to the mean of the stimulus: $\boldsymbol{n_{int}} \sim N(0,\ k\langle \boldsymbol{s} + \boldsymbol{n_{ext}}\rangle)$, as the observer is not able to distinguish which of the input is derived from signal or external noise. This scaling is analogous to Fano factor, used in neurophysiology to model the variance in neural spikes. The template ($\boldsymbol{w_i}$) for an ideal observer is the signal direction (left or right). On a given trial, the ideal observer chooses the direction that maximizes the posterior probability of the direction given the received input: $\text{decision} = \text{argmax}P(\boldsymbol{w_i} \mid \boldsymbol{g})$, which in our case is the cross-correlation between $\boldsymbol{g}$ and $\boldsymbol{w_i}$. Thresholds were predicted from 1,000 Monte Carlo simulations of our experimental paradigm. Internal noise, k, was simulated at 10 levels ranging between 0.1 and 5. Simulation results were very similar when we assumed constant internal noise or noise with a SD proportional to the SD of the filter.

To understand the effects of spatial tuning on spatiotemporal separability, we compared the performance of the ideal observer with direction templates derived from the separable and 7 nonseparable models with varying tuning width, which were used in place of $\boldsymbol{w_i}$. The separable model was created using Gaussian tuning in both space and time similar to Fig. 5*A*, but described with three parameters: gain, center, and width. The center for spatial and temporal Gaussians was assumed to be symmetric. The nonseparable filter was modeled as an oriented Gaussian with three parameters: gain, width (varied between 0.2 and 0.9), and orientation (fixed at 45 degrees). To create inhibitory responses, filter values were scaled between [−1, 1]. The templates were normalized such that the sum of the responses within the excitatory

and inhibitory regions are 1 and −1, respectively. The gain of the models was fixed at 1.

Statistical analysis. For statistical analyses, we tested the effects of group and conditions at the significance level of .05. Holm–Bonferroni corrections were used for multiple comparisons as needed.

Data, Materials, and Software Availability. Code for the experiment and analyses are openly available at https://github.com/VisCog/aud-motion-sighted-earlyblind (71). The data that support the findings of this study (in anonymized format) are available from the corresponding author upon request. The current IRB at the University of Washington has restrictions on publicly sharing data that contain individual visual and medical histories.

ACKNOWLEDGMENTS. We would like to thank the blind and sighted individuals who participated in this research and the blind community in the Greater Seattle Area. This work was supported by NIH National Eye Institute Grants R01EY014645 (to I.F.) and K99EY034546 (to W.J.P.) and a Weill Neurohub Postdoctoral Fellowship (to W.J.P.). We also would like to thank Priyanka Kotipalli, Rachel Tang, Feiyi Wang, and Jackson Watson for their help with data collection and Geoffrey Boynton for his help with modeling.

1. T. D. Albright, Direction and orientation selectivity of neurons in visual area MT of the macaque. *J. Neurophysiol.* **52**, 1106–1130 (1984).
2. J. A. Movshon, E. H. Adelson, M. S. Gizzi, W. T. Newsome, "The analysis of moving visual patterns" in *Pattern Recognition Mechanisms*, C. Chagas, R. Gattass, C. G. Gross, Eds. (Vatican Press, 1985), pp. 117–151.
3. E. H. Adelson, J. R. Bergen, Spatiotemporal energy models for the perception of motion. *J. Opt. Soc. Am. A* **2**, 284–299 (1985).
4. C.-J.J. Dong, N. V. Swindale, P. Zakarauskas, V. Hayward, M. S. Cynader, The auditory motion aftereffect: Its tuning and specificity in the spatial and frequency domains. *Percept. Psychophys.* **62**, 1099–1111 (2000).
5. D. W. Grantham, Motion aftereffects with horizontally moving sound sources in the free field. *Percept. Psychophys.* **45**, 129–136 (1989).
6. D. W. Grantham, F. L. Wightman, Auditory motion aftereffects. *Percept. Psychophys.* **26**, 403–408 (1979).
7. M. F. Neelon, R. L. Jenison, The effect of trajectory on the auditory motion aftereffect. *Hear. Res.* **180**, 57–66 (2003).
8. S. Anstis, F. A. J. Verstraten, G. Mather, The motion aftereffect. *Trends Cogn. Sci.* **2**, 111–117 (1998).
9. S. Carlile, J. Leung, The perception of auditory motion. *Trends Hear.* **20**, 1–19 (2016).
10. T. C. A. Freeman *et al.*, Discrimination contours for moving sounds reveal duration and distance cues dominate auditory speed perception. *PLoS One* **9**, 27–29 (2014).
11. M. Ahissar, E. Ahissar, H. Bergman, E. Vaadia, Encoding of sound-source location and movement: Activity of single neurons and interactions between adjacent neurons in the monkey auditory cortex. *J. Neurophysiol.* **67**, 203–215 (1992).
12. J. A. Altman, Are there neurons detecting direction of sound source motion? *Exp. Neurol.* **22**, 13–25 (1968).
13. H. Jiang, F. Lepore, P. Poirier, J.-P. Guillemot, Responses of cells to stationary and moving sound stimuli in the anterior ectosylvian cortex of cats. *Hear. Res.* **139**, 69–85 (2000).
14. R. Grzeschik, M. Böckmann-Barthel, R. Mühler, M. B. Hoffmann, Motion-onset auditory-evoked potentials critically depend on history. *Exp. Brain Res.* **203**, 159–168 (2010).
15. N. J. Ingham, H. C. Hart, D. McAlpine, Spatial receptive fields of inferior colliculus neurons to auditory apparent motion in free field. *J. Neurophysiol.* **85**, 23–33 (2001).
16. D. McAlpine, D. Jiang, T. M. Shackleton, A. R. Palmer, Responses of neurons in the inferior colliculus to dynamic interaural phase cues: Evidence for a mechanism of binaural adaptation. *J. Neurophysiol.* **83**, 1356–1365 (2000).
17. A. Alink, F. Euler, N. Kriegeskorte, W. Singer, A. Kohler, Auditory motion direction encoding in auditory cortex and high-level visual cortex. *Hum. Brain Mapp.* **33**, 969–978 (2012).
18. F. Baumgart, B. Gaschler-Markefski, M. G. Woldorff, H.-J. Heinze, H. Scheich, A movement-sensitive area in auditory cortex. *Nature* **400**, 724–726 (1999).
19. C. Y. Ducommun *et al.*, Cortical motion deafness. *Neuron* **43**, 765–777 (2004).
20. C. Poirier *et al.*, Specific activation of the V5 brain area by auditory motion processing: An fMRI study. *Cogn. Brain Res.* **25**, 650–658 (2005).
21. J. D. Warren, B. A. Zielinski, G. G. R. Green, J. P. Rauschecker, T. D. Griffiths, Perception of sound-source motion by the human brain. *Neuron* **34**, 139–148 (2002).
22. K. R. Smith, K. Okada, K. Saberi, G. Hickok, Human cortical auditory motion areas are not motion selective. *Neuroreport* **15**, 1523–1526 (2004).
23. K. R. Smith, K. Saberi, G. Hickok, An event-related fMRI study of auditory motion perception: No evidence for a specialized cortical system. *Brain Res.* **1150**, 94–99 (2007).
24. A. Ahumada, J. Lovell, Stimulus features in signal detection. *J. Acoust. Soc. Am.* **49**, 1751–1756 (1971).
25. C. K. Abbey, M. P. Eckstein, Classification image analysis: Estimation and statistical inference for two-alternative forced-choice experiments. *J. Vis.* **2**, 66–78 (2002).
26. A. Ahumada, Perceptual classification images from vernier acuity masked by noise. *Perception* **25**, 18 (1996).
27. R. F. Murray, Classification images: A review. *J. Vis.* **11**, 1–15 (2011).
28. P. Neri, A. J. Parker, C. Blakemore, Probing the human stereoscopic system with reverse correlation. *Nature* **401**, 695–698 (1999).
29. P. Neri, D. M. Levi, Receptive versus perceptive fields from the reverse-correlation viewpoint. *Vision Res.* **46**, 2465–2474 (2006).
30. P. Neri, Dynamic engagement of human motion detectors across space-time coordinates. *J. Neurosci.* **34**, 8449–8461 (2014).
31. F. Jiang, G. C. Stecker, I. Fine, Auditory motion processing after early blindness. *J. Vis.* **14**, 1–18 (2014).
32. G. Bertonati, M. B. Amadeo, C. Campus, M. Gori, Auditory speed processing in sighted and blind individuals. *PLoS One* **16**, e0257676 (2021).
33. D. Alais, D. Burr, No direction-specific bimodal facilitation for audiovisual motion detection. *Cogn. Brain Res.* **19**, 185–194 (2004).
34. B. S. Tjan, W. L. Braje, G. E. Legge, D. Kersten, Human efficiency for recognizing 3-D objects in luminance noise. *Vision Res.* **35**, 3053–3069 (1995).
35. D. M. Green, J. A. Swets, *Signal Detection Theory and Psychophysics* (John Wiley, 1966), (2 October, 2023).
36. G. Westheimer, S. P. Mckee, Spatial configurations for visual hyperacuity. *Vision Res.* **17**, 941–947 (1977).
37. G. Westheimer, D. M. Levi, Depth attraction and repulsion of disparate foveal stimuli. *Vision Res.* **27**, 361–1368 (1987).
38. G. M. Verdon-Roe, M. C. Westcott, A. C. Viswanathan, F. W. Fitzke, D. F. Garway-Heath, Exploration of the psychophysics of a motion displacement hyperacuity stimulus. *Invest. Ophthalmol. Vis. Sci.* **47**, 4847–4855 (2006).
39. D. MacVeigh, D. Whitaker, D. B. Elliott, Spatial summation determines the contrast response of displacement threshold hyperacuity. *Ophthalmic Physiol. Opt.* **11**, 76–80 (1991).
40. K. van der Heijden, J. P. Rauschecker, E. Formisano, G. Valente, B. de Gelder, Active sound localization sharpens spatial tuning in human primary auditory cortex. *J. Neurosci.* **38**, 8574–8587 (2018).
41. K. Derey, G. Valente, B. De Gelder, E. Formisano, Opponent coding of sound location (Azimuth) in planum temporale is robust to sound-level variations. *Cereb. Cortex* **26**, 450–464 (2016).
42. G. C. Stecker, I. A. Harrington, J. C. Middlebrooks, Location coding by opponent neural populations in the auditory cortex. *PLoS Biol.* **3**, e78 (2005).
43. D. McAlpine, D. Jiang, A. R. Palmer, A neural code for low-frequency sound localization in mammals. *Nat. Neurosci.* **4**, 396–401 (2001).
44. D. R. Perrott, K. Saberi, Minimum audible angle thresholds for sources varying in both elevation and azimuth. *J. Acoust. Soc. Am.* **87**, 1728–1731 (1990).
45. A. W. Mills, On the minimum audible angle. *J. Acoust. Soc. Am.* **30**, 237–246 (1958).
46. I. B. Witten, E. I. Knudsen, Why seeing is believing: Merging auditory and visual worlds. *Neuron* **48**, 489–496 (2005).
47. C. Battal, V. Occelli, G. Bertonati, F. Falagiarda, O. Collignon, General enhancement of spatial hearing in congenitally blind people. *Psychol. Sci.* **31**, 1129–1139 (2020).
48. T. Vercillo, A. Tonelli, M. Gori, Early visual deprivation prompts the use of body-centered frames of reference for auditory localization. *Cognition* **170**, 263–269 (2018).
49. M. Gori, G. Sandini, C. Martinoli, D. C. Burr, Impairment of auditory spatial localization in congenitally blind human subjects. *Brain* **137**, 288–293 (2014).
50. J. M. Beck, W. J. Ma, X. Pitkow, P. E. Latham, A. Pouget, Not noisy, just wrong: The role of suboptimal inference in behavioral variability. *Neuron* **74**, 30–39 (2012).
51. B. A. Dosher, Z. L. Lu, Perceptual learning reflects external noise filtering and internal noise reduction through channel reweighting. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 13988–13993 (1998).
52. Z.-L. Lu, B. A. Dosher, Perceptual learning retunes the perceptual template in foveal orientation identification. *J. Vis.* **4**, 44–56 (2004).
53. A. E. Burgess, R. F. Wagner, R. J. Jennings, H. B. Barlow, Efficiency of human visual signal discrimination. *Science* **1979**, 93–94 (1981).
54. A. Burgess, H. Ghandeharian, Visual signal detection. I. Ability to use phase information. *J. Opt. Soc. Am.* **1**, 900–905 (1984).
55. A. E. Burgess, B. Colborne, Visual signal detection. IV. Observer inconsistency. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* **5**, 617–627 (1988).
56. H. K. Sanju, P. Kumar, Enhanced auditory evoked potentials in musicians: A review of recent findings. *J. Otol.* **11**, 63–72 (2016).
57. L. B. Lewis, M. Saenz, I. Fine, Mechanisms of cross-modal plasticity in early-blind subjects. *J. Neurophysiol.* **104**, 2995–3008 (2010).
58. G. Dormal, M. Rezk, E. Yakobov, F. Lepore, O. Collignon, Auditory motion in the sighted and blind: Early visual deprivation triggers a large-scale imbalance between auditory and "visual" brain regions. *Neuroimage* **134**, 630–644 (2016).
59. M. Bedny, T. Konkle, K. Pelphrey, R. Saxe, A. Pascual-Leone, Sensitive period for a multimodal response in human visual motion area MT/MST. *Curr. Biol.* **20**, 1900–1906 (2010).
60. C. Poirier *et al.*, Auditory motion perception activates visual motion areas in early blind subjects. *Neuroimage* **31**, 279–285 (2006).
61. F. Jiang, G. C. Stecker, G. M. Boynton, I. Fine, Early blindness results in developmental plasticity for auditory motion processing within auditory and occipital cortex. *Front. Hum. Neurosci.* **10**, 1–14 (2016).
62. G. S. L. Coullon, F. Jiang, I. Fine, K. E. Watkins, H. Bridge, Subcortical functional reorganization due to early blindness. *J. Neurophysiol.* **113**, 2889–2899 (2015).
63. A. Gurtubay-Antolin *et al.*, Direct structural connections between auditory and visual motion-selective regions in humans. *J. Neurosci.* **41**, 2393–2405 (2021).
64. H. Abe *et al.*, Axonal projections from the middle temporal area in the common marmoset. *Front. Neuroanat.* **12**, 89 (2018).
65. M. Gamberini, C. Galletti, A. Bosco, R. Breveglieri, P. Fattori, Is the medial posterior parietal area V6A a single functional area? *J. Neurosci.* **31**, 5145–5157 (2011).
66. S. Pitzalis, P. Fattori, C. Galletti, The functional role of the medial motion area V6. *Front. Behav. Neurosci.* **6**, 1–13 (2013).
67. K. Nelissen, P. A. Fiave, W. Vanduffel, Decoding grasping movements from the parieto-frontal reaching circuit in the nonhuman primate. *Cereb. Cortex* **28**, 1245–1259 (2018).
68. C. Grefkes, G. R. Fink, The functional organization of the intraparietal sulcus in humans and monkeys. *J. Anat.* **207**, 3–17 (2005).
69. D. H. Brainard, The psychophysics toolbox. *Spat. Vis.* **10**, 433–436 (1997).
70. H. H. Schütt, S. Harmeling, J. H. Macke, F. A. Wichmann, Painfree and accurate Bayesian estimation of psychometric functions for (potentially) overdispersed data. *Vision Res.* **122**, 105–123 (2016).
71. W. J. Park, I. Fine, Auditory motion perception in sighted and early blind. GitHub. https://github.com/VisCog/aud-motion-sighted-earlyblind. Deposited 5 May 2023.