Convergence of temporal and spectral information into acoustic images of complex sonar targets perceived by the echolocating bat, *Eptesicus fuscus*

James A. Simmons, Cynthia F. Moss*, and Michael Ferragamo

Department of Psychology and Section of Neurobiology, Division of Biology and Medicine, Brown University, Providence, R1 02912, USA

Accepted September 14, 1989

**Summary.** 1. FM echolocating bats (*Eptesicus fuscus*) were trained to discriminate between a two-component complex target and a one-component simple target simulated by electronically-returned echoes in a series of experiments that explore the composition of the image of the two-component target. In Experiment I, echoes for each target were presented sequentially, and the bats had to compare a stored image of one target with that of the other. The bats made errors when the range of the simple target corresponded to the range of either glint in the complex target, indicating some trace of the parts of one image interfered with perception of the other image. In Experiment II, echoes were presented simultaneously as well as sequentially, permitting direct masking of echoes from one target to the other. Changes in echo amplitude produced shifts in apparent range whose pattern depended upon the mode of echo presentation.

2. *Eptesicus* perceives images of complex sonar targets that explicitly represent the location and spacing of discrete glints located at different ranges. The bat perceives the target's structure in terms of its range profile along a psychological range axis using a combination of echo delay and echo spectral representations that together resemble a spectrogram of the FM echoes. The image itself is expressed entirely along a range scale that is defined with reference to echo delay. Spectral information contributes to the image by providing estimates of the range separation of glints, but it is transformed into these estimates.

3. Perceived absolute range is encoded by the timing of neural discharges and is vulnerable to shifts caused by neural amplitude-latency trading, which was estimated at 13 to 18 μs per dBA from *N*1 and *N*4 auditory evoked potentials in *Eptesicus*. Spectral cues representing the separation of glints within the target are transformed into estimates of delay separations before being incorporated into the image. However, because they are encoded by neural frequency tuning rather than the time-of-occurrence of neural discharges, the perceived range separation of glints in images is not vulnerable to amplitude-latency shifts.

4. The bat perceives an image that is displayed in the domain of time or range. The image receives no evident spectral contribution beyond what is transformed into delay estimates. Although the initial auditory representation of FM echoes is spectrogram-like, the time, frequency, and amplitude dimensions of the spectrogram appear to be compressed into an image that has only time and amplitude dimensions. The spectral information is not lost but manifests itself as equivalent time-domain information.

**Key words:** Echolocation – Bat sonar – Target ranging
Acoustic images – Auditory processing

**Introduction**

Echolocating bats emit sonar sounds and perceive objects from the echoes of these sounds that return to their ears (Busnel and Fish 1980; Griffin 1958; Nachtigall and Moore 1988). Echolocation is a biological acoustic imaging process in which the capabilities of the auditory system for processing sounds are devoted to the task of creating spatial images. The existence and formation of the images that bats perceive is of considerable interest because it illuminates aspects of auditory function as well as principles of spatial perception in general.

Bats that emit constant-frequency (CF) signals for echolocation primarily use the dimension of echo frequency to create images of targets. These bats have evolved unusually sharply-tuned peripheral auditory filters which improve their resolution of the small frequency changes associated with Doppler-shifted echoes (Neuweiler et al. 1980). They take advantage of frequency sidebands introduced by rapid modulations in the frequency and amplitude of echoes returning from flying insects to identify targets in terms of wing-beat
rates (Henson et al. 1987; Schnitzler et al. 1983; von der Emde and Menne 1989). Sharp peripheral filtering of echoes appears to be a crucial part of the process that renders these fluctuations in CF echoes perceptible to the bat.

Bats that emit frequency-modulated (FM) signals have equally effective means for identifying targets (Griffin et al. 1965; Sum and Menne 1988), but the specific nature of this process is not yet well understood. Although ‘FM’ bats can discriminate differences in a variety of target and echo parameters (Schnitzler and Henson 1980; Simmons and Grinnell 1988; Suthers and Wenstrup 1987), no specific dimension of echoes has yet been conclusively demonstrated to support, or be the basis for, the acoustic images these bats perceive. FM bats do not, for example, use especially sharply-tuned auditory filters to analyze echoes (Jen et al. 1989; Suga and Jen 1977), so it is unlikely that measurement of the frequency composition of echoes by itself can play as dominant a role in FM echolocation as it does in CF echolocation. Several experiments have shown that FM bats can discriminate between targets that reflect echoes having different spectra (Bradbury 1970; Habersetzer and Vogler 1983; Schmidt 1988; Simmons et al. 1974). However, these experiments do not specifically address the nature of the images actually perceived by bats. Although perception of target shape may involve the use of echo spectral information, this does not necessarily mean that FM bats perceive the echo spectrum to be target shape.

Do FM bats perceive targets that differ in shape merely as having ‘spectrally distinct’ echoes, or do they perceive the spatial differences in target structure that are responsible for differences in echo spectra? In broader terms, what are the psychological dimensions of the images FM bats perceive? Engineers pose a similar question in seeking to understand the basis for echolocation in information processing. Their approach is to identify the model that the bat uses to describe targets, and then to show how the bat estimates the parameters of this model (see Altes 1976, for example). What, then, are the parameters of the model target that FM bats must estimate to describe the target’s shape? The fact that bats can distinguish between targets whose echoes have different spectra does not prove that the dimensions of the bat’s model of targets are best expressed as spectral parameters. The bat could equally well build time-domain parameters into its model, and just estimate these parameters from echo spectra by transforming them from the frequency-domain to the time-domain. The bat’s auditory system most likely uses a hybrid time-frequency representation of FM echoes that resembles a spectrogram showing instantaneous frequency (Simmons 1973; Simmons and Kick 1984; see below), so temporal and spectral information are jointly available to represent target features (Altes 1980, 1984). But, while the bat’s model of echoes may be a spectrogram, its model of targets could be transformed further. Target shape and echo spectral discrimination experiments do not even show whether the model’s representations actually serve the bat as an ‘image’. The model may simply consist of a cluster of separate parameters to be estimated. However, if the bat can be shown to remember the full shape of a target, whose representation is formally distributed across several separately-estimated echo-acoustic parameters, then these parameters must somehow have been transposed into terms that unify the representation along a common psychological scale prior to storage. The model, by being remembered as a whole in spite of its disparate parts, is thus perceived as an image.

An experimental jittered-echo technique that examines perception of very small differences in the delay of echoes by the big brown bat, *Eptesicus fuscus* (Simmons 1979) reveals a fine structure to the image of a target along the delay axis that corresponds to the cross-correlation function between emissions and echoes (Simmons 1979, Simmons et al., in press). (Two recent, inexact replications of the original jittered-echo experiment failed to obtain the crosscorrelation result [Moss and Schnitzler 1989; Menne et al. 1989] due to use of a different alignment of the jittering and nonjittering stimuli that seemed insignificant but that converts the experiment into a masking paradigm. Exact replication of the original experiment has reproduced the crucial crosscorrelation result and extended it to show a basis for crosscorrelation in the time-of-occurrence of neural discharges [Simmons et al., in press].) This finding raises the possibility that *Eptesicus*, and perhaps other FM bats, perceive the structure of complex targets along the range axis (Simmons 1979, 1980, 1987; Simmons and Stein 1980). That is, the perceptual dimension of echo delay and target range might support the creation of acoustic images in FM echolocation in a manner similar to the role of echo frequency and target velocity in CF echolocation. This possibility exists because echoes having different spectra would also have equivalently different crosscorrelation functions.

If the bat in fact perceives the echo crosscorrelation function as its image of a target, the target’s shape would be represented along the range dimension as a series of discrete reflecting points, or glints (Altes 1976) corresponding to side-peaks in the crosscorrelation function. The placement of these side-peaks would reflect the spectral composition of echoes, and echoes differing in their spectra would still be discriminable, but the bat would perceive these differences as changes in the range profile of targets rather than merely as changes in echo spectra. To the degree that the spectrum of echoes specifically conveys information about target shape, that spectrum would have to be converted into a set of estimates of the locations of glints along the range axis prior to the final display of the image perceived by the bat. This transformation would be required even if echoes were represented by a spectrogram; the placement of peaks and notches in the ensemble spectral slices making up the spectrogram (Beuter 1980; Altes 1984) would not automatically appear as glints at their appropriate places along the delay axis without systematic, reciprocal-taking communication from the frequency axis of the spec-
Program. In this case, the transformation would convert
the spectrogram into the crosscorrelation (time-frequency
or magnitude-unsquared ambiguity) function of
echoes (Altès 1980; van Trees 1971).

This paper describes a series of experiments that
identify the composition of the acoustic images of com-
plex targets perceived by *Eptesicus fuscus*. The results
indicate how these images must be displayed in the bat's
sonar receiver. Specifically, they demonstrate that *Epi-
tesicus* perceives the shape of a target in terms of the tar-
get's range profile by converting depth-structure informa-
tion from the spectrum or spectrogram of FM echoes
back into a range-axis image of the target. The bat's
model of targets is based on glint structure, with param-
eters being estimated jointly from the arrival-time of
echoes and from their spectra. In this sense, the use of
a spectrogram representation of echoes is confirmed.

**General methods**

Behavioral experiments. We conducted a series of psychophysical
experiments aimed at dissecting the structure of acoustic images
perceived by bats. The animals used in these experiments were
two big brown bats, *Eptesicus fuscus* (family Vespertilionidae),
attained from the attic of houses in Rhode Island and southeastern
Massachusetts. The general procedures pertaining to all of the
behavioral experiments are described first (see also Simmons et al.
1988), and then the specifics of each particular stimulus condition
are considered separately.

Each bat was trained in a two-alternative forced-choice proce-
dure to discriminate between an electronically-simulated complex
target containing two components at slightly different distances
and a simulated simple target containing only one component.
A simple sonar target consists of a single reflecting point, or glint,
located at a discrete distance, or range, from the bat (Altès 1976).
It reflects a single filtered replica of the incident sound that
arrives at a discrete delay after each emission. A complex target
consists of two or more acoustic glints located at slightly different
distances. It therefore reflects compound echoes composed of two
or more filtered replicas of the incident sound separated by corre-
spondingly small differences in delay. Each millimeter of distance
to the glint adds approximately 5.8 μs to the delay of its echo
component.

If an FM sonar signal impinging on a complex, multiple-glint
target is longer in duration than the time separation of the individ-
ual echo components within the compound echo, the echo compo-
nents overlap. The resulting spectrum for the compound echo con-
tains alternating peaks and notches or nulls created by interference
between the overlapping echo components. The bat's auditory sys-
tem imposes the constraint that the time separation of the echo
components must be shorter than the integration-time of the per-
ipheral band-pass filters in the sonar receiver (Beufer 1980)
for interference to occur. The integration-time for echo reception in
a task similar to that used below has been estimated to be about
350 μs in *Eptesicus* (Simmons et al., 1989). If energy from two
separate echoes falls within the integration-time of individual band-
pass filters, the filtered echo waveforms mix to form a complex
spectral interference pattern. Because the bat's sonar sounds are
FM and therefore excite peripheral auditory band-pass filters in
succession (from high to low frequency), a spectrogram-like audi-
tory representation is created (Altès 1980; Simmons and Kick
1984; again, see below). Spectral peaks and notches are then pres-
ent in the ensemble spectrum slices that constitute the spectrogram
(Altès 1984). The frequencies at which these peaks and notches
occur are related to the time separation of the echo components.

![Diagram of the two-choice discrimination procedure for studying perception of complex targets.](image)

Fig. 1. Diagram of the two-choice discrimination procedure for studying perception of complex targets. Bats chose between a two-component complex target (a₁ and a₂) with electronically-simulated echoes arriving 100 μs apart (Δt) at fixed delays of 3.275 and 3.375 ms (simulated ranges of 56.47 and 58.19 cm, respectively) and a one-component simple target (b) whose echoes arrived at delays varying from 3.125 to 3.450 ms (ranges from 53.88 to 59.48 cm). The bat's sonar sounds were picked up at microphones (m₁, digitally delayed, and then returned to the bat from loudspeakers (s₁) as echoes

and, hence, the range profile (depth-structure) of the target. The
peaks occur at frequencies that are integer multiples of the reciprocal
of the time separation, while the notches are interposed between the
peaks. The frequency separation of either the peaks or the
notches is equal to the reciprocal of the time separation. The peaks
are broad and flat-topped, while the notches are sharp and well-
defined (see below, especially Fig. 5). Consequently, the notches
are likely to be a better index of the characteristics of the echo
spectrum than the peaks (Schmidt 1988; Simmons et al. 1974),
especially at high echo signal-to-noise ratios (Altès 1984). From
the periods or wavelengths of the frequencies at which these peaks
or notches occur, one can determine the equivalent echo-delay or
target-range differences that produced them.

**Target simulation.** Figure 1 shows schematically the design of
the experiment and the method used to present the bat with elec-
tronically-reproduced echoes that simulate sonar targets at different
distances. The bat was trained to sit on an elevated, Y-shaped
platform and broadcast its sonar sounds to the front - somewhat
to the left and the right - to discriminate a complex target (a₁
and a₂ in Fig. 1) from a simple target (b). The simple, single-glint
target consisted of a single replica of each of the bat's sonar
emissions, simulating an echo from a point target. The complex, two-
glint target consisted of two replicas of each emitted sound deliv-
ered with a time separation of 100 μs. It thus contained two reflect-
ing elements. The apparent distance to the simple and complex
targets was regulated by controlling the delay of the echoes elec-
tronically. The bat was rewarded with a piece of a mealworm of-
f erred in forcepts for each correct response (arrow in Fig. 1),
which consisted of crawling forward onto the left or right arm of
the platform - whichever was in the direction of the complex (positive)
stimulus for that particular trial. The bat's ability to determine
which stimulus was the complex target varied according to the
range of the simple target in a manner that revealed the apparent
shape of the complex target in the data. The appearance of the
complex stimulus on the left or right was determined by a pseudo-
random schedule and set by a switch located beneath the platform. The experiment was set up in a 4.5 × 3.3 × 2.4 m chamber whose walls, ceiling, and floor were lined with convoluted polyurethane foam (Perma Foam Corp., Irvington, NJ) that reduced the amplitude of ultrasonic reverberation by at least 20 to 30 dB compared with what would occur if the chamber had smooth, hard walls.

The electronic system for simulating sonar targets was built around the bat's observing position on the Y-shaped platform. Two Bruel & Kjaer Model 4138 condenser microphones (m in Fig. 1) were mounted at the ends of the arms of the platform to pick up the bat's echolocation sounds. The electrical signal from each microphone was amplified, filtered to a 20 to 100 kHz band with a Rockland Model 442 band-pass filter, delayed by a controlled amount, and then returned to the bat from an RCA electrostatic loudspeaker (Part No. 112343, s in Fig. 1) that was mounted next to the microphone. During representative trials, echolocation sounds recorded from the bats were stored on analog magnetic tape with a Racal Store-4 tape recorder and subsequently reproduced for digital signal analysis on an IBM PC AT computer operating with ILS programs from Signal Technology, Inc. Spectrograms of the sounds were made with a Unigon digital sound spectrograph. Both the microphones and the loudspeakers were located 20 cm from the bat's observing position at the center of the platform, so that together they provided a propagation delay of 1.16 ms for any sound emitted by the bat and returning to the bat's ears after passing through the simulator. The angle separating the two sets of microphones and speakers was approximately 40°.

The bats were tested about 6 days a week, and on the test days each bat was run on a number of trials that was determined by its current body weight and the quantity of mealworms consumed after correct trials. Each day's run constituted a block of trials for one of the experimental conditions (various combinations of echo-presentation regimes for echoes a1 and a2, and various delay and amplitude settings for echo b), and the stimuli were set to a new condition on the following day. The bats typically worked through 35 to 60 trials in each block. If the number of trials achieved on a single day was less than this, the same stimulus conditions were repeated the next day to accumulate more trials.

The data take the form of percentages of errors made over all trials at any particular stimulus condition and delay setting, and the primary mode of presentation of the data is a graph of percentages as a function of the delay of echo b. The data were not arbitrarily divided into above- and below-threshold states because our concern is with the form of the images perceived by the bat as revealed by the shape of the entire curve for each condition (Simmons 1973, 1979), rather than with reduction of the data to a single, artificial estimate of discriminability.

During individual experimental trials, which lasted for several seconds, each bat emitted FM echolocation sounds with durations of about 1.5 to 2.5 ms at rates of 5 to 15 sounds s−1 (see Sect. on sounds, below). Each sound was received at both microphones, with an amplitude at each microphone that depended on the aim of the bat's head during head-scanning movements as the bat searched for the simulated targets (Simmons and Vernon 1971). The amplified and filtered signal representing each of the bat's sonar sounds was delayed electronically by a digital delay system (delay in Fig. 1) designed and built by the Science Services Shops at the University of Oregon (Simmons et al. 1988). The signal was digitized with 12-bit accuracy at a rate of 730 kHz, stored in a circulating buffer memory, and then read out and reconstructed as an analog signal after a preset delay. The total gain of the analog circuitry feeding into the delay lines was about 80 dB to bring the peak-to-peak amplitude of the majority of signals to a level just below the 12-bit limit of the digitizer for maximum signal-to-noise ratio. The magnitude of the electronic delay was chosen so that, when added to the 1.6-ms propagation delay from the bat to the microphone and from the loudspeaker to the bat, it created a total delay corresponding to the desired simulated target range. Each microphone-loudspeaker channel was equipped with a dual-delay system, so the bat could be presented with multiple simulated glints that appeared on the bat's right and on its left from the observing position on the platform.

Acoustic calibration. The target simulator is an acoustic recording and reproducing system whose performance can be summarized by a frequency-response curve or transfer function. The frequency response of the left and right channels of the target simulator shown in Fig. 1. The sonar signals of Eptesicus contain energy from about 24 to 100 kHz, and the simulator returned the entire first harmonic (60 to 15 kHz) as well as most of the second harmonic up to about 85 kHz.

Fig. 2. A graph of the frequency response of the left and right channels of the target simulator shown in Fig. 1. The sonar signals of Eptesicus contain energy from about 24 to 100 kHz, and the simulator returned the entire first harmonic (60 to 25 kHz) as well as most of the second harmonic up to about 85 kHz.

The amplitude of the echoes delivered to the bat were set to fixed levels with respect to an estimate of the bat's threshold for detecting a single-glinit test target (at the location of a1) presented in isolation. To determine the threshold for each bat prior to the beginning of the experiments, the amplitude of test echos was reduced from about 80 dB SPL peak-to-peak in steps of 5, 2, or 1 dB on successive blocks of trials to determine the level at which the bat's performance in the two-choice detection task fell to about 50% correct responses. The threshold was found by plotting percent correct responses against echo sound pressure level in dB, and the echo level corresponding to 75% correct responses was arbitrarily defined as threshold. In all experimental conditions, the amplitude of echoes representing target glints a1 and a2 was set at 15 dB above this detection threshold.
Experiment I

In the first series of experiments, each bat performed a sequential discrimination task in which the simple target (negative stimulus) and the complex target (positive stimulus) were presented one-at-a-time in the following manner: the bat emitted sonar sounds while at its observing position on the platform (Fig. 1), and a voltage comparator determined whether the left or right microphone picked up the stronger signal (i.e., the side towards which the bat's head was aimed). Only the channel on that particular side was activated. Thus, the bat received echoes from either the left or the right loudspeaker – never both at once. The bat learned to scan on the left and the right to sample echoes from both simulator channels. Since only the left or the right channel could be activated at any one time, the bat's scanning determined the sequential presentation of the simple and complex targets through the two channels.

The sequential-presentation paradigm forces the bat to rely upon remembered images of the targets to select the side on which the complex target is presented. The bat's performance depends upon remembering the image of one target for comparison with the image of the other. In comparing images, the bat makes errors when a residue or trace of part of the image of one target interferes with perception of part of the other image. Specifically, when echoes from the simple, single-glint target arrive at the same delay as echoes from either the nearer or the farther glint of the complex target, the bat makes more errors than if the delays are different (see below). In terms of the acoustic stimuli, the sequential-presentation paradigm creates masking for perceived echo delays that lasts longer than the echoes themselves and exerts its effect only locally over a small region of time intervals between emissions and echoes corresponding to the echo delay actually delivered to the bat. The interfering effect between image traces thus is range-specific, and the increased incidence of errors that results can be used to map the structure of the image of the complex target along the bat's psychological range axis.

Condition A

The bat was trained in the sequential-presentation task to select a complex target that reflected two echoes delivered together and separated by a 100 μs delay. The complex target (a1–a2 in Fig. 1) was presented through either the left or the right channel with a propagation delay of 1.16 ms added to electronic delays of 2.115 (glint a1) and 2.215 (glint a2) ms to create overall echo delays of 3.275 and 3.375 ms. (These are nominal delays that varied somewhat in practice with the movement of the bat's head during trials. The 100-μs separation of the echo components remains invariant, however.) These delays simulate the first and second glints of a depth-structured target with glints at ranges of 56.47 and 58.19 cm, respectively. The 100-μs delay separation between the two glints corresponds to a spatial separation of 17.2 mm, which is a realistic distance for insect-sized targets that FM bats are capable of identifying (Simmons and Chen, 1989). Since the sonar sounds of *Eptesicus*, in this experiment are approximately 2 ms in duration (see below), the electronically-reproduced echoes of the first and second glints of the complex target are largely overlapping, creating peaks and notches across frequencies in the composite spectrum at 10 kHz intervals. The delay separation of 100 μs was chosen because it is shorter than the approximately 350-μs integration-time for echo reception by *Eptesicus* (Simmons et al., 1989), and because earlier range discrimination data show that *Eptesicus* can distinguish between two glints separated by this delay, even if the perceived range of the individual sonar targets is smeared by head movements (Simmons and Grinnell 1988: see Simmons 1973). When bats are trained to discriminate between two single-glint targets at different simulated ranges (echo delays), the bat's error curve shows a peak that is less than 100 μs wide. Figure 3 shows the results of an echo-delay discrimination experiment for 2 single-glint targets (Simmons 1973) to serve as a reference for reading the graphs presented below.

In Experiment I, the simple target (b) consisted of electronically reproduced echoes of the bat's sonar emissions arriving at a single delay. The echo delay for the simple target was constant on any given test day but varied from day-to-day over a range from 50 μs before the first glint (a1) of the complex target to 150 μs after a1. The total echo delay of the simple target thus varied between 3.225 and 3.425 ms (target ranges between about 55.6 and 59.1 cm). The bat's performance at choosing the complex target thus was assessed as a function of the delay of echo b. Echoes representing a1, a2, and b were all presented at a level 15 dB above the detection threshold measured for a1 alone.

Fig. 3. Echo delay discrimination data from *Eptesicus* for two single-glint simulated targets, replotted from Simmons (1973). The procedure is equivalent to discrimination of a1 from b in Fig. 1. The shape of this curve is not changed as absolute delay or range changes. The zero point on the delay axis corresponds to the delay of echoes for a1.

Results

Figure 4A shows the performance curves of two bats trained to discriminate the complex two-glint target (a1–a2) from the simple one-glint target (b), with varying
Fig. 4A–C. Graphs from Experiment 1 showing percentage of errors produced by two bats (circles and triangles) in 3 different experimental conditions. A with two glints \( a_1 \) and \( a_2 \) presented together as one complex target. B with the spectral effects of adding two glints created instead by filtering echoes of the first glint without presentation of the second glint at all. And C with separate, alternating presentation of the two glints from one emission to the next, using a jitter technique. Percentage errors is plotted as a function of the delay separation between the simple target \( h \) and the first glint \( a_1 \) of the complex target. Zero indicates that the delay of echoes \( h \) and \( a_1 \) coincided.

Delays for the echoes of the simple target. The two curves plot the individual data for each bat, showing percent errors for 9 different echo delays of the simple target. The delay plotted on the abscissa is that of the simple target with respect to the first glint of the complex target. Thus, an echo delay of zero \( \mu s \) represents the condition in which the first glint of the complex target coincides in range with the only glint of the simple target (at an actual delay of 3.275 ms). Around this delay, both bats show an increase in errors compared to neighboring delays. At an echo delay of 100 \( \mu s \) in Fig. 4A, the second glint of the complex target coincides in simulated range with the simple target. At this delay the bats’ errors increase again. The location and spacing of the error peaks corresponds to the range profile of the complex target.

Discussion of Condition A

In the task for Condition A, the targets were presented sequentially, so the peaks in the error curves represent interference from one target’s image to the other; some trace of one image must remain long enough to partially mask perception of the other image. Moreover, this trace-masking effect is range-specific. The particular parts of the images that participate in the masking effect correspond to glints at specific ranges. In the case of the interference between \( a_1 \) and \( b_0 \) that produces errors around zero \( \mu s \) in Fig. 4A, it is understandable that the bat should be able to represent the range of each glint and make errors when their ranges are the same. Both are represented by the arrival-time of echoes and probably encoded by the timing of neural discharges. In effect, \( a_1 \) and \( b_0 \) represent conditions similar to \( a \) and \( h \) in a range discrimination task (Fig. 3), but with sequential presentation of targets. However, the error peak for interference between \( a_2 \) and \( b_0 \) is less straightforward in origin. Because the separation of \( a_1 \) and \( a_2 \) is 100 \( \mu s \), which is substantially smaller than the integration-time for echo reception (about 350 \( \mu s \); Simmons et al., 1989), it is most unlikely that \( a_2 \) can evoke neural discharges of its own (Beuter 1980). For \( a_2 \) to create an error peak comparable to \( a_1 \) in Fig. 4A, the complex spectral interference pattern generated by mixing echoes of \( a_1 \) and \( a_2 \) with a 100-\( \mu s \) time separation somehow must manifest itself in the image of the complex target as a point along the image’s delay or range axis, which is a psychological dimension.

The peaks in Fig. 4A only reach about 25% errors because two glints are present to identify the correct target. Only one of these can be masked by the single-glint target, leaving the other to be recognized by the bat. As Fig. 3 shows, when only one glint is present in the positive stimulus, the bat’s performance reaches 50% errors. Aspects of the factors responsible for the reduced height of the peaks in Fig. 4 compared to Fig. 3 are examined below in Condition C of this experiment.

Condition B

If the bat indeed perceives the 100-\( \mu s \) delay separation of \( a_1 \) and \( a_2 \) in the complex target from the spectrum of the compound echo, an appropriately filtered single echo for \( a_1 \) alone should be perceived by the bat as a complex target with two glints separated by 100 \( \mu s \). We tested this possibility by electronically filtering echo \( a_1 \), with 6 parallel, voltage-tunable band-pass filters (Variable Systems, Model 330A) to mimic the spectral peaks at 30, 40, 50, 60, 70, and 80 kHz and the intervening spectral notches at 25, 35, 45, 55, 65, 75, and 85 kHz that are ordinarily created by mixing \( a_1 \) and \( a_2 \) into a compound echo. Figure 5 shows the frequency response of the components creating this filtered echo in comparison with the frequency response associated with two overlapping glints separated by 100 \( \mu s \). These measurements were obtained by passing a 1.0-ms FM calibration signal (100 to 10 kHz sweep) through one channel of the target simulator and determining the transfer function at the point in the circuit where the filters were introduced. The curves in Fig. 5 show the interference peaks and notches created by overlap of 2 echo components (dashed line) and these same notches
Fig. 5. A graph showing the relative frequency response of the target simulator when two echo components are presented 100 µs apart (glimpse dashed line) compared with the frequency response created when parallel band-pass filters are used to impose peaks and notches in the spectrum of the echo from a single glint (filtered solid line).

Simulated by placing filter resonances at frequencies corresponding to the interference peaks (solid line). The filtered version of the complex echo is not a perfect replica of the interference effect, particularly with respect to the depth of the notches, but it is good enough for two distinct time delays to appear in the compound echo. Figure 6 shows the impulse response for the filters in comparison with the impulse response created by two separate echo components being added together. Both impulse waveforms contain 2 discrete clusters of peaks separated by about 100 µs. These clusters represent the 2 glints created by spectral notches originating either from interference or from filtering. The relatively shallower notches in the filtered echo result in a smaller amplitude for the second set of peaks relative to the first set of peaks, but the 2 glints are clearly present in both cases.

In a formal sense, Fig. 6 shows that it is not strictly necessary to carry out Condition B because the stimuli for Conditions A and B are nearly identical in every sense conveyed by the impulse response. However, there is one difference between the waveforms of echoes in Conditions A and B that makes Condition B a useful experiment. When the echoes for \( a_1 \) and \( a_2 \) are separated by 100 µs as in Condition A, the leading (onset) and trailing (offset) 100-µs segments of the compound echo do not overlap. The filtered echoes in Condition B do not have such nonoverlapping leading and trailing segments. Consequently, visual inspection of the echo waveform alone easily distinguishes between the 2 conditions. The purpose of Condition B is to create the second component of the impulse response - that is, the \( a_2 \) glint - by a physically different mechanism than addition of two echo components with a separation of 100 µs. Although not formally necessary, it seems intuitively desirable to show that the same impulse response arrived at by 2 different means (see Fig. 6) is perceived equivalently by the bat.

In Condition B of the experiment, only echo \( b \) and the filtered echo \( a_1 \) were presented to the bats, whose task still was to select the complex target (filtered echo \( a_1 \)). Echo \( b \) was moved to different locations to determine whether the bat perceived the complex spectral structure of echo \( a_1 \) as if the simulated target contained two glints. The delay of echo \( b \) was always 3.275 ms, the same as that in Condition A. The delay of echo \( b \) varied between 3.225 and 3.450 ms - that is, from 50 µs before to 175 µs after echo \( a_1 \). In Condition A, the delay of \( b \) was changed in steps of 25 µs and this same step size was used in Condition B with the exception of 2 values that had a step size of only 12.5 µs. These immediately-spaced delays of \( b \) were at 12.5 µs and 112.5 µs relative to the delay of \( a_1 \). Smaller-than-usual step sizes were used here because the filters which created the notches in the spectrum of \( a_1 \) also retarded the phase of the signal by an average of about 90° across much of the frequency range used by the bat. The impulse response in Fig. 6 shows the phase lag of the filtered echo relative to the two-glint echo, an effect that is especially pronounced for the part of the impulse that corresponds to the nearer glint in the simulated target. The additional echo delay caused by the phase lag was about half of the 25-µs step size used in condition A. Accordingly, two half-step points were inserted in the immediate vicinity of the objective delays of \( a_1 \) and \( a_2 \) to avoid missing the exact apparent delay while collecting the data. The amplitude of the filtered echo \( a_1 \) was adjusted to match that of echo \( a_1 \) in Condition A, and echo \( b \) again was 15 dB above each bat's detection threshold for \( a_1 \) alone.

Results

Both of the trained bats made the transition from discrimination of the two-glint simulated target in Condition A to discrimination of the filtered echoes in condition B. Their performance did not suffer as the switch was made, suggesting that the bats regarded the tasks as similar. Figure 4B shows the performance curves of 2 bats trained to discriminate the filtered echo \( a_1 \) from echo \( b \). Each curve plots the percentage of errors made by one bat for the different delays of echo \( b \) with reference to the delay of the filtered \( a_1 \). Again, a delay of zero µs on the abscissa marks the time at which echo \( b \) is presented with the same delay as echo \( a_1 \). As in Condition A, the bats both show an increase in errors for echo delays close to zero µs and to 100 µs. The actual peaks of the error curves are at 12.5 and 112.5 µs, the
two points interposed to account for the phase lag of the filtered echoes, particularly that of the first impulse component corresponding to \( a_1 \). One of the 2 bats (triangular data-points) was tested at each of these interposed delays for a limited session of 40 trials in the stimulus regime for Condition A and achieved 12.5% errors at both points. The peaks in Fig 4B thus appear to be shifted slightly to the right in comparison with the peaks in Fig 4A.

Discussion of Conditions A and B

The waveforms of the echoes delivered to the bats in Conditions A and B were formed by different electronic mechanisms, nevertheless, they have nearly the same impulse representations (Fig. 6). The bats perceived both types of waveforms in the same way in the experiments, yielding error peaks at delays corresponding to the main peaks in the impulse response. In Condition B, *Epitesicus* perceives the filtered echo \( a_1 \), with its resonance-created peaks and notches, as though it consisted of 2 discrete components separated by 100 μs, which actually was the stimulus for Condition A. Somehow the bat interprets the echo spectrum or the frequency axis of the echo spectrogram to signify the existence of more than one glint in the target, so that a point corresponding to \( a_2 \) is inserted onto the delay or range dimension of the image. This point does not appear automatically on the delay axis from information represented on the time axis of echo spectrograms owing to the 100-μs glint separation being shorter than the 350-μs integration-time for echo reception (Simmons et al., 1989). The image component for \( a_2 \) must arise through neural computations that move information from a frequency to a time scale.

(To verify whether the echo spectrum or the frequency axis of the spectrogram is indeed the origin of the perceived \( a_2 \) glint while direct neural encoding of echo arrival-time is the origin of the perceived \( a_1 \) glint, the 2 glints in the image must be dissociated experimentally, in effect demonstrating the physiological separateness of their representations. Experiment II, below, was carried out for this purpose.)

The waveform of the echo created by adding \( a_1 \) and \( a_2 \) differs from the waveform of the notch-filtered \( a_1 \) alone. The leading and trailing segments of the compound \( a_1 + a_2 \) echo, each segment being 100 μs long, are not involved in the overlap of the 2 echo components and can easily be seen in the raw waveform to be distinct from the overlapping portion. The waveform of the filtered \( a_1 \) echo has no such ‘nonoverlapping’ leading and trailing segments, only spectral peaks and notches at the same locations as the compound echo (Fig. 5). Nevertheless, both stimulus conditions give rise to a two-glint image. Thus, the bat does not perceive the separation of \( a_1 \) and \( a_2 \) from some simple observation of the echo waveform itself. Instead, the bat converts the shape of the echo spectrum into an estimate of the equivalent underlying time separation of summing-and-interfering echo components and then uses this estimated time separation to perceive the range profile of the target. From Fig 6 it is evident that the bat perceives an impulse-like or time-domain representation of the echoes from the complex target even though it must combine auditory time and frequency representations to do so. The emergence of a time-domain image of targets from a joint time-frequency representation of echoes is a central thesis of the spectrogram-correlation model of echolocation (Altes 1980, 1984). It follows, too, from the observation that bats perceive an image that corresponds approximately to the crosscorrelation function of echoes even though echoes themselves are coded in spectrogram or instantaneous-frequency terms (Simmons 1975, 1979; Simmons et al., in press).

Condition C

The combined results of Conditions A and B suggest that the bat uses the frequency axis of the spectrum or spectrogram of echoes to perceive the structure of the target. However, this does not necessarily imply that the bat perceives the spectrum entirely as the equivalent of target depth-structure. Although the echo spectrum evidently is used to determine the separation of \( a_1 \) and \( a_2 \), there may be aspects of the spectrum that the bat perceives without specifically using them to perceive the second glint. That is, the echo spectrum could contribute to discrimination of the complex target without this contribution being manifested in the appearance of the second glint in the image. This point is related to the observation that the error peaks in Fig 4A and B only rise to about 25% errors, rather than the 50% that occurs when only a single glint is present in both of the simulated targets (see Fig. 3). Some feature of the two-glint target makes it more discriminable than a one-glint target in what is essentially a target-range discrimination task. Condition C was intended to determine whether the echo spectrum contributes any information beyond what is used to estimate the separation of the two glints. In this condition, the bats were tested with a complex target that returned two echo components separated by 100 μs without the spectral information created by presenting two overlapping echo components together.

The complex target again contained two glints, \( a_1 \) and \( a_2 \), at delays of 3.275 and 3.375 ms. However, in Condition C, the echoes of \( a_1 \) and of \( a_2 \) were returned to the bat in alternation. That is, the glints of the complex target were presented as though they really were a single glint that jittered back and forth in range between 56.47 and 58.19 cm from one echo to the next. Following each of the bat’s sonar emissions, either \( a_1 \) or \( a_2 \) returned, but not both together — the two glints of the complex target were never presented in tandem for the same sonar emission. A digital circuit was added to each channel of the target simulator that counted the bat’s sonar sounds and alternated the state of a flip-flop to gate \( a_1 \) and \( a_2 \) on and off in alternation (Simmons 1979; Simmons et al., in press). Thus, the bat was presented with two glints separated by 100 μs (or 17.2 mm).
without the spectral consequences of the direct mixing of their echoes. To the degree that each glint somehow activates a neural representation of its own particular range, Condition C is similar to Conditions A and B, but the spectral notches that ordinarily signify the presence of the second glint are removed. The simple target, containing only glint \( h \), was presented at different delays from 50 \( \mu s \) before \( a_1 \) to 175 \( \mu s \) after \( a_1 \). As in Conditions A and B, all of the echoes were presented at 15 dB above the detection threshold for \( a_1 \) alone.

**Results**

Both bats smoothly transferred from Condition B to Condition C without showing any initial decrement in performance, again suggesting that they considered these tasks to be very much alike. Figure 4C shows the individual performance curves of 2 bats trained to discriminate the complex target that jitters between glints \( a_2 \) and \( a_1 \) from the simple target \( h \). As before, the percentage of errors is plotted against the delay of the single glint \( h \) versus the delay of \( a_1 \). The results show an increase in errors when target \( h \) is presented at the same delay as either the first glint (zero \( \mu s \) delay) or the second glint (100 \( \mu s \) delay) of the complex target, even though the spectral cues present in Conditions A and B were removed by precluding overlap of echo components. By itself, the presence of glints alternating between two ranges is enough to create a 'complex target' in terms of the bat's ability to discriminate one target from another. Most significantly, the peaks in the error curves in the region of zero and 100 \( \mu s \) delays closely resemble each other in height and width for all 3 experimental conditions. To the bats, the two-glint complex target is equally discriminable from the simple target whether the 2 glints are presented in compound echoes, presented in a filtered echo, or presented in alternation.

**Echolocation sounds.** During selected trials from conditions A, B, and C, the echolocation sounds used by the bats for discriminating the complex target from the simple target were tape-recorded and then reproduced for analysis to determine their properties. Both of the bats used essentially the same signals in all 3 conditions, with variations primarily in the overall duration of the signals from about 1.5 to 2.5 ms. Figure 7 shows spectrograms of representative sonar sounds emitted by the two bats during trials of Condition C. In Experiment II (below), the bats also used sounds that were the same. The signals are frequency modulated, with 3 harmonics that together cover the frequency range from 23 to about 100 kHz. These sounds are similar to signals used by *Eptesicus* during other laboratory discrimination experiments (Simmons et al. 1988, for example) and during the approach or tracking stage of interception of prey (Griffin 1958; Simmons 1987). The frequency response of the target simulator (Fig. 2) shows that the echoes returned to the bat contained predominantly the first harmonic and part of the second harmonic of these sounds, and that the small segment of the third harmonic around 90 kHz probably played little part in either bat's performance.

**Discussion of Experiment I**

In all 3 experimental conditions of Experiment I, the bats explicitly perceive the location along the range axis of 2 glints in the complex target. Whether the second glint in the complex target is actually represented by its echo component (Conditions A and C) or only represented by an artificially created spectrum (Condition B) seems not to matter: *Eptesicus* treats the 3 stimulus regimes as perceptually alike. The bats evidently derive their ability to perceive the complex target as being complex because it contains 2 glints located at different ranges. The peaks in Fig. 4 only rise to about 25% errors instead of 50% errors as in Fig. 3 because the complex target's range profile makes it appear different from the simple target. The bats perceive 2 glints in one target and only one glint in the other, so that, even when the simple target is at the same range as either of the glints in the complex target, the bats can still perceive the other glint to identify the correct stimulus. However, each bat's performance declines when the single glint coincides in range with one of the double glints, but the decline is not total (that is, to 50% errors) because the other glint is still perceived and contributes to the discrimination. The errors occur because the glints in the image of one target leave behind a *trace or memory* that can interfere with perception of glints in the other target a second or two later when the bat scans from one simulator channel to the other. This persistence of the image of the complex target as a whole, incorporating both glints, even though the underlying acoustic representation may incorporate separate time and frequency elements, is the justification for using the term ‘image’ in this context.
The principal conclusion from Experiment I concerns the role of the frequency axis of the spectrum or spectrogram of echoes in relation to the time axis for perception of complex targets. Under ordinary acoustic conditions, when the presence and range of the second glint, \( a_2 \), in a complex target is represented by interference peaks and notches in the spectrum of the compound echo, *Eptesicus* perceives this spectral information as though it signifies the existence of an 'event' along the range axis rather than simply events along the frequency axis. This is also true when the spectral notches are created artificially with filter resonances. The location and spacing of the notches and or peaks along the frequency axis is sufficient to evoke an event at the correct, reciprocally-related echo-delay or target-range separation. That is, the bat perceives the shape of a complex target from the spectrum of echoes, but it perceives this shape directly in terms of the underlying delay or range separations required to produce spectral notches at specific frequencies.

The shape of the echo spectrum is converted by the bat into an image that expresses spectral information in the domain of time rather than frequency, as would be expected if the bat used a crosscorrelation-like sonar receiver (Simmons 1979, 1980; Simmons and Stein 1980; Simmons et al., in press). Figure 6 suggests that *Eptesicus* perceives an image that corresponds to the target impulse response, and the echo crosscorrelation function is merely the expression of the impulse response for the particular incident sound wave being used. Systematically-arranged spectral peaks and notches would appear in the crosscorrelation function of echoes as one or multiple discrete side-peaks in delay or range, and this corresponds to what the bat actually perceives. In contrast to \( a_2 \), the distance to the nearer glint, \( a_1 \), must be perceived from the arrival-time of the echo as a whole, which presumably is represented directly as a delay by the auditory system. The presence of components for both glints \( a_1 \) and \( a_2 \) in the image perceived by the bat shows that the bat integrates time and frequency dimensions of its auditory representation of echoes, as would be expected of a process utilizing spectrograms. The results of Experiment I thus are consistent with the use of an auditory spectrogram representation of echoes that ultimately is displayed in terms of the crosscorrelation function by a mechanism that transforms information present along the frequency axis and the time axis of spectrograms into an image wholly along the time axis of the crosscorrelation function.

Not only does *Eptesicus* use the echo spectrum to perceive complex sonar targets in terms of their range profile, but for these purposes the echo spectrum is entirely transformed into a time-domain representation. In Experiment I, the bat derives no capacity to discriminate complex targets from their spectra beyond what is used to estimate range profile. The range of the second glint may be specified from the individually-presented delay of separate, alternating echo components (Fig. 4C) or more naturally from the spectral consequences of mixing the two echo components together (Fig. 4A) without affecting the bat's performance. *Eptesicus* perceives the spectrally estimated range profile of the complex target with the same level of performance as the two glints presented in alternation, with no available spectral cues. The spectrum thus does not contribute to the discriminability of the complex target beyond what is used to estimate the underlying delay separation implied by the presence of peaks and notches at particular frequencies.

The auditory system of *Eptesicus* may establish a physiological representation of echoes that separately incorporates information about the arrival-time and the spectrum of echoes, presumably using physiological axes of echo delay and echo frequency to display echo parameters in neural terms (Jen et al. 1989; Sullivan 1982; Suga 1988; Wong and Shannon 1988). A hybrid representation of this sort previously has been proposed from considerations of peripheral auditory coding of echoes and the performance of bats in target-range discrimination experiments (Simmons 1973). The neural displays of time and frequency must coexist, since the acoustic parameters they represent are encoded quite differently at early stages of the auditory pathways (Bodenheimer and Pollak 1981; Simmons and Kick 1984; Suga 1988). The process of 'reading' the two displays entails conversion of the terms of one (the spectral display) into the terms of the other (the delay display). The quite distinct peripheral physiological manifestations of delay and spectral information (time-of-occurrence of nerve discharges versus tuned frequency of neural 'channels') provide a means for confirming experimentally that these two representations indeed are independent of each other prior to the convergence of their outputs onto the image-display perceived by the bat.

**Experiment II**

The results of Experiment I provide evidence that the auditory representation of a depth-structured target in *Eptesicus* involves a hybrid of frequency-domain and time-domain processing of echoes. Experiment II seeks to dissociate the components of the image of a complex target that have their origin in a time-domain representation from components that have their origin in a frequency-domain representation. The perceived distance to the first reflecting element of a complex target (or to a single-glint, simple target) is represented by the delay of echoes after emissions, which most likely is initially encoded by the time-of-occurrence of neural discharges marking the different frequencies in the FM sweep of emissions followed by similar discharges for echoes (Bodenheimer and Pollak 1981; Pollak et al. 1977; Simmons and Kick 1984). The resulting array of neural discharges for the emission or the echo resembles a spectrogram, with each point in the representation having both time-domain and frequency-domain information (Altes 1980, 1981, 1984; Simmons and Kick 1984). The time that elapses between neural discharges evoked by echoes and discharges evoked by previous
emissions ultimately is displayed in the bat's auditory cortex (Sullivan 1982; Suga 1988; Wong and Shannon 1988). The neural representation of this time separation remains experimentally vulnerable to any factor capable of perturbing the faithful neural registration of the time at which echoes occur prior to the stage where the range display is actually computed. The most obvious candidate for perturbing the timing of neural discharges to echoes is the relation between stimulus amplitude and neural response latency. As echoes are increased in strength, the latency of the discharges they evoke becomes shorter (Bodenhamer and Pollak 1981; Bodenhamer et al. 1979; Pollak et al. 1977; Suga 1970). It should be possible to shift the apparent range of the first glint in a complex target just by changing the amplitude of its echoes.

The spectral peaks and notches in echoes specify only the range separation of the two glints, with the absolute range of the second glint necessarily being referred to the absolute range of the first glint. In Fig. 4. *Epipicus* perceives both glints at their correct ranges, indicating that the range separation somehow is expressed in terms of absolute range. The locations of the notches in the spectrum initially are represented physiologically by the amount of neural activity evoked in neural channels tuned to different frequencies across the frequency range of the FM sweep of echoes. The individual frequencies in the sweeps are registered as 'excitatory frequencies' (Bodenhamer and Pollak 1981) in neural response profiles that are not much influenced by the amplitude of echoes. Only the latency of the neural discharges in any particular frequency channel changes with stimulus amplitude, not the frequency that evokes the discharges in that channel. In other words, the time-domain part of the spectrogram representation of echoes is very sensitive to echo amplitude, but the frequency-domain part is not. If two echoes overlap and interfere with each other so as to create spectral notches, these notches remain at the same frequencies even though one echo changes in amplitude relative to the other. Only the depth of the notches changes with the relative amplitude of the echoes. The differing effects of changes in echo amplitude on neural response latencies and the neural representation of spectral notches is exploited in Experiment II to distinguish between the images of the first and second glints in a complex target.

Auditory evoked potentials

We measured latency changes for neural responses to FM sounds at biologically appropriate stimulus levels (Kick and Simmons 1984) in five *Epipicus*. Bats were anesthetized by intramuscular injection of a ketamine-rompun mixture (50 mg kg\(^{-1}\)) supplemented with metofane inhalation. The skull was exposed over the midbrain, and a small hole was made in the skull to insert a tungsten wire electrode onto the dorsal surface of the inferior colliculus to record auditory evoked potentials from the auditory nerve (\(N_1\)) and nucleus of the lateral lemniscus (\(N_4\)). FM sounds, 1.0 ms in duration and sweeping from 110 down to 20 kHz, were broadcast through a specially-built electrostatic loudspeaker (Simmons et al. 1989) located 10 cm from the bat's ear. These acoustic stimuli were presented over a range of amplitudes from 20 to 90 dB SPL peak-to-peak. The neural responses to these sounds were amplified with a WPI Model DAM-5A physiological amplifier and digitized at 100 kHz (12 bit accuracy). Sets of 256 responses for each amplitude were stored and averaged (JRC Electronics Model ISC-16 data acquisition system). We defined neural response thresholds to be the lowest sound pressure that still produced recognizable averaged auditory evoked potentials. Our methods are otherwise similar to those used earlier with *Epipicus* (Feng et al. 1978), and our observed thresholds are within 10 dB of previously-reported behavioral or physiological thresholds when the frequency response of our acoustic system is taken into account (Dalland 1966; Gardner and Garvey 1974; Kick 1980).

Results

As sound pressure level increases, the latency of auditory evoked responses in *Epipicus* to brief FM sounds becomes shorter (Fig. 8). The latency of the \(N_1\) evoked potential is plotted as a function of sound pressure level relative to threshold. For both animals, the evoked-potential latency shortens by approximately 0.55 ms as stimulus amplitude increases from near threshold to 40 dB above threshold. The inset of Fig. 8 shows the entire waveform of averaged evoked potentials from one bat for a series of amplitudes above threshold to illustrate how the latency of the whole ensemble of the response systematically shortens as the FM sounds become stronger. At a level 15 dB above threshold, the slope of the \(N_1\) latency function is approximately \(-13 \mu\text{sec dB}^{-1}\) and the slope of the \(N_4\) function is nearer \(-18 \mu\text{sec dB}^{-1}\). If the bat uses the timing of neural discharges making up such evoked responses to FM sonar echoes to determine target range, an increase in echo amplitude in this region should shorten the target's apparent range by about 2 to 3 mm dB.

![Fig. 8 A, B](image-url)

A graph showing the dependence of the time-of-occurrence of neural responses on the amplitude of FM sounds in the auditory system of *Epipicus*. A The \(N_1\) (auditory nerve) evoked-potential latency for 2 bats (circles and triangles) shortens by 0.55 ms as stimulus amplitude increases from near threshold to 40 dB above threshold. B Averaged evoked potentials (inset) from one bat illustrate how the latency of the \(N_1\) through \(N_4\) (lateral lemniscal) responses systematically shortens as the FM sounds become stronger. At a level 15 dB above threshold, the slope of the \(N_1\) latency function is about \(-13 \mu\text{sec dB}^{-1}\) and the slope of the \(N_4\) function is nearer \(-18 \mu\text{sec dB}^{-1}\).
Behavioral Experiments

To determine which parts of the image of a complex sonar target are susceptible to changes in apparent distance due to amplitude-induced latency shifts, we again studied the bat’s performance at discriminating between an electronically-simulated complex target containing two components separated by 100 μs and a simulated simple target containing only one component (Fig. 1). In Experiment I the echoes for \(a_1\) and \(a_2\) were presented sequentially with the echoes for \(h\) by taking advantage of the directionality of the bat’s sonar sounds and its head movements while scanning the targets. In Experiment II these echoes were presented sequentially in one case and simultaneously in another, to examine the effects of directly mixing the echoes for \(h\) with those for \(a_1\) and \(a_2\). In the case of sequential presentation, the bat’s performance is affected only by interference or masking between traces of the remembered images of the targets. In the case of simultaneous presentation, masking can occur directly from one sound to the other without bringing the images into play. When the bat receives the echo for \(a_1\) or for \(a_2\), it also receives the echo for \(h\), and, depending upon the relative timing of these echoes, simultaneous, forward, or backward masking could occur. Our intention was to use changes in the latency of neural discharges to the echoes of \(h\) to identify whether it is the timing of these discharges or the spectral consequences of mixing \(h\) with \(a_1\) and \(a_2\) that determines the delay at which the bat perceives \(h\) to be aligned with either \(a_1\) or \(a_2\). Echoes of each component of the complex target (\(a_1\) and \(a_2\)) were presented at a fixed strength 15 dB above behavioral detection threshold, while echoes of the single-component simple target (\(h\)) were presented at several different strengths ranging from 9 to 24 dB above threshold in 2 separate tasks.

Two different stimulus-presentation regimes were used in Experiment II to control the mixing of \(h\) with \(a_1\) and \(a_2\). The first was a two-choice simultaneous discrimination task in which both the left and the right channels of the target simulator were activated together by each sonar emission. (The electronic comparator that determines which way the bat aimed its sonar sounds was simply switched off.) In this condition, the complex and simple targets always were presented together. For each sonar emission, all three echoes — \(a_1\), \(a_2\), and \(h\) — were delivered at their appropriate delays to the bat’s ears, creating the opportunity for masking to occur from one echo to another. Since the bat receives overlapping echoes from both the simple and complex targets each time it emits a sonar sound, echoes impinging on the bat’s ears from both targets can interfere with each other and create additional spectral interference patterns. In particular, even as \(a_2\) interferes with \(a_1\) to create spectral notches signifying the separation of \(a_1\) and \(a_2\), so too, can \(h\) interfere with \(a_1\) and create competing spectral notches according to its own delay separation from \(a_1\). If the range separation of \(a_2\) relative to \(a_1\) is represented by spectral notches in the compound echo of \(a_1\) and \(a_2\), then alignment of \(h\) with \(a_2\) could obscure these notches through masking, and the bat should make discrimination errors because it perceives \(h\) and \(a_2\) to be at the same distance. Changing the amplitude of the echoes for \(h\) should have no effect on the bat’s perception of the alignment of \(h\) with \(a_2\) because the spectral notches associated with \(a_1 + h\) will not change in their frequencies. In other words, any delay-specific masking effect of \(h\) on \(a_2\) should not shift in delay-specificity with small changes in the amplitude of \(h\). For the simultaneous-presentation regime, the \(a_2\) peak in the error performance curve of the type shown in Fig. 4 should thus not move left or right on the graph as the amplitude of \(h\) changes.

When \(h\) is aligned with \(a_1\) in the simultaneous-presentation regime, the bat also should make errors due to masking. If the absolute ranges of \(a_1\) relative to \(h\) are represented by the timing of neural discharges evoked by each echo, a change in the amplitude of the echoes for \(h\) should shift its apparent range relative to \(a_1\) as a consequence of the amplitude-latency trade-off. That is, any delay-specific masking effect of \(h\) on \(a_1\) should shift in delay-specificity as the amplitude of \(h\) changes. For example, in the case of a decrease in the amplitude of the echoes for \(h\), which should retard its corresponding neural discharges, these echoes would have to be moved to an earlier electronic delay to coincide with \(a_1\) in the bat’s perception. For the simultaneous-presentation regime, the \(a_1\) peak in an error performance curve of the type shown in Fig. 4 should thus move to the left on the graph as \(h\) decreases.

The second stimulus-presentation regime in Experiment I was a two-choice sequential discrimination task, as described above in the methods for Experiment I. In the sequential discrimination task, a voltage comparator determined whether the left or right microphone received the greater signal according to the aim of the bat’s head, and echoes were returned to the bat from only that side. Since just one target is presented to the bat for each sonar emission, echoes from the complex and simple targets no longer can impinge on the bat’s ears together, and any masking that does occur must depend upon traces of the images lingering after the echoes have terminated. In particular, echoes for \(h\) no longer mix with echoes for \(a_1\) and \(a_2\), so no competing spectral notches would be created. In this sequential condition (which is identical to Condition A of Experiment I except that the echoes for \(h\) are changed in amplitude), the distance from \(a_1\) to \(a_2\) can only be known from spectral cues created when \(a_1\) and \(a_2\) mix together because their 100-μs separation is well within the integration-time for echo reception. The bat’s perception of the absolute range of \(a_2\) thus depends upon perception of the delay-determined absolute range of \(a_1\), because the spectral cues for \(a_2\) are referred to \(a_1\) when they are converted into images. Changing the amplitude of echoes for \(h\) thus should shift the error peaks for both \(a_1\) and \(a_2\) together (as in a graph such as is shown in Fig. 4) in the sequential condition, rather than shift only the
peak for $a_1$ as in the simultaneous condition. This should occur because the entire, remembered image of the complex target has a perceived absolute range that should depend on the latency of neural discharges to $a_1$ relative to neural discharges to $h$.

In both the simultaneous and sequential tasks, the complex target contained two glints, $a_1$ and $a_2$, whose echoes arrived after delays of 3.275 and 3.375 ms, respectively. The delay of echoes from the simple target, $h$, varied in steps of 25 μs from 150 μs before the first glint of the complex target to 175 μs afterwards. The performance of the bats was measured for each delay of echoes from the simple target, and at amplitudes of echo $b$ from $-6$ to $-9$ dB relative to echoes from the individual components of the complex target, which again were $15$ dB above detection threshold.

Results

Simultaneous discrimination. As in Experiment I, the data take the form of percentage errors recorded for each bat while discriminating the simulated two-glint target from the simulated one-glint target as the one-glint target appeared at a succession of different distances around the distance of the two-glint target. Figure 9 shows the results of the simultaneous discrimination condition for two amplitudes of the echoes for $h$ relative to echoes for $a_1$ and $a_2$. When echoes for the single-glint ($h$) target are equal in strength to echoes of each of the glints ($a_1$, $a_2$) in the complex target ($0$ dB condition, upper panel of Fig. 9), both glints in the complex target appear as peaks in the error curves at their correct or objective, echo delays. This result is similar to that shown in Fig. 4A, even though the simple and complex targets were presented simultaneously in this condition. However, when the echoes for the single-glint target are $6$ dB weaker than echoes for each of the glints in the complex target ($-6$ dB condition, lower panel), the error peak corresponding to the first glint in the complex target shifts to the left by about 75 to 100 μs. That is, the bats perceive the single glint ($h$) to be at the apparent range of the first glint ($a_1$) of the complex target when its echoes are presented about 75 to 100 μs earlier than the objective delay for echoes of that first glint. In contrast, the error peak corresponding to alignment of the single glint ($h$) with the second glint ($a_2$) still appears at the objective range for the second glint.

The leftward shift of the error peak for the first glint by 75 to 100 μs with a $6$-dB amplitude change in echoes for $h$ is consistent with a $-13$ to $-17$ μs dB trading relationship between the amplitude of echoes and the latency of neural discharges to those echoes. In the masking conditions created by simultaneous presentation of both targets, the error peaks for $a_1$ and $a_2$ thus are dissociated by their differing response to changing the amplitude of $h$.

Figure 10 shows in more detail the shift in the apparent range of the single target ($b$) with respect to the first glint ($a_1$) of the complex target for one of the bats.

---

Fig. 9. Results for two bats (circles and triangles) on simultaneous discrimination between the simple and complex targets (Experiment II). Percentage of errors is plotted as a function of the delay separation between the complex and simple targets. At zero on the abscissa echoes from the simple target ($h$) coincide in delay with echoes from the first glint ($a_1$) of the complex target. Data are shown for conditions in which the simple target's echoes are equal in strength to echoes from each of the components of the complex target ($0$ dB, upper plot) and in which the simple target's echoes are half as strong ($-6$ dB, lower plot). Scale C on right refers to Fig. 10.

Fig. 10. More detailed data for one bat from Fig. 9 (circles) plotted as percentage-error contours over a $15$-dB span of amplitude differences between echoes. Heavy sloping line shows the $-13$ μs dB latency shift of $X_1$ evoked responses (Fig. 5). Contour intervals based on a scale of 1-6 (shown as Scale C to the right of Fig. 9).

The bat's performance is shown on a contour graph of percentage errors for a wider range of amplitudes of the single-glint echoes than is shown in Fig. 9. In Fig. 10, the first glint appears as a ridge sloping upward to the right, and the second glint appears as a ridge...
aligned vertically. As the amplitude of the echoes simulating the single target is increased from $-6$ to $+9$ dB relative to the individual echo components of the complex target, the bat perceives the single target as aligning with the first glint at progressively greater ranges. In contrast, the delay of echoes from the single target aligns with the objective delay of echoes from the second glint at all amplitudes tested. The $15$-dB variation in echo amplitude forces the apparent delay of echoes from the single target to shift over a span of about $200 \mu s$, which corresponds to $3.4$ cm. The sloping contour ridge even appears to invert the complex target when the single glint must, paradoxically, appear further away than the second glint to align with the first glint. (This illusion is a consequence of manipulating the amplitude of echoes for $b$ rather than for $a_1$ or $a_2$, which was only a methodological convenience here.)

**Sequential discrimination.** When the bat has to rely on memory of acoustic images to discriminate the complex and simple targets, and direct echo-to-echo masking cannot occur, the effects of changing echo amplitude are different. Figure 11 shows the results of the two-glint versus one-glint sequential discrimination experiment with *Eptesicus*. When the echoes of the single-glint target have the same amplitude as echoes from either glint in the complex target (0 dB condition, upper panel in Fig. 11), the error peaks corresponding to the two glints again fall at their objective ranges. (These curves are the same as in Fig. 4A because this is Condition A from Experiment I.) However, when the echoes for the single-glint target are reduced in amplitude by 6 dB, the apparent range of the single glint shifts with respect to the range of both of the glints in the complex target (−6 dB condition, lower panel in Fig. 11). The size of the shift is in the range of 50 to 75 or 100 $\mu s$. The two glints of the complex target are locked together relative to the single glint in the sequential-presentation task, where the image of the complex target must be retained to perform the discrimination. Figure 12 shows in more detail the performance of one bat for a wider range of amplitudes of the single-glint echoes than is shown in Fig. 11. Here, the contour ridges associated with both glints in the complex target slant upward to the right. The apparent range of the target as a whole follows the shift that, in the simultaneous-presentation task, only affects the relation between the single-glint target and the first glint of the complex target. The size of the shift for the two peaks together is consistent with an amplitude-latency trading relationship of about $-13 \mu s$ dB.

**Discussion of Experiment II**

In Experiment II, the dissociation of the first and second glints in the complex target from each other relative to the single glint in the simple target confirms the result of Experiment I— that different mechanisms encode their positions along the range axis. Starting at a level 15 dB above the bat's threshold, the latency of $N_1$ responses in *Eptesicus* is lengthened by about 78 $\mu s$ when FM stimuli decrease in amplitude by 6 dB (Fig. 8). The latency of $N_4$ responses is lengthened by about 108 $\mu s$. To compensate for a greater neural response latency, the weaker echoes from the single-glint target in the $-6$ dB condition of Fig. 9 would have to be presented 78 to 108 $\mu s$ earlier (to the left along the delay axis) if their neural responses are to coincide with those representing echoes of the first glint of the complex target. The bats suffer this shift in the delay of the single-glint echoes, indicating
that the delay of echoes of both the single glint and the first of the double glints probably is represented by the timing of neural discharges in response to the FM sweep (Bodenhamer and Pollak 1981; Bodenhamer et al. 1979; Pollak et al. 1977; Suga 1970; Kick and Simmons 1984) In contrast, the single-glint echoes still appear to the bat to be aligned with the objective delay of the echoes from the second glint in the complex target even when they are made 6 dB weaker. Thus, it is not likely that the bat represents the range of the second glint relative to the first by the timing of neural discharges.

It is more likely that the bat uses notches and or peaks in the overall spectrum or the spectrogram of the composite echo from the complex target to represent the separation of the 2 echo components. The arrival of the single-glint echoes on top of the echoes from the second glint of the complex target would obscure these spectral notches in a manner independent of amplitude changes as small as those which produce dramatic shifts in the apparent range of the single glint relative to the first glint in the complex target.

Experiment II confirms another result obtained in Experiment I. The tandem movement of the error peaks corresponding to the double glints in Figs 11 and 12 demonstrates that, notwithstanding their differing physiological representations early in the auditory processing of echoes, the eventual neural representation of the complex target is positioned in absolute range at a point determined by the delay of echoes from the nearer of the two glints. The spectral information initially used to represent the range of the second glint relative to the first eventually is expressed in terms of absolute target range as this is registered by the timing of the discharges evoked by the first echo component. Evidently, the bat converts the spectral information about the second glint to an estimate of its spatial separation from the first glint, using the time-based estimate of the range of the first glint to support construction of the image of the target as a whole. The output of the sonar receiver of Eptesicus thus is displayed in the natural spatial domain of the distance along which the target is distributed.

General discussion

There already is a great deal of evidence that echo spectral information may be used by FM bats for perception of target shape (Beuter 1980; Bradbury 1970; Griffin 1967; Habersetzer and Vogler 1983; Schmidt 1988; Simmons et al. 1974; Simmons and Lavender 1976). But the crucial step taken in the experiments described here is to identify specifically how echo spectral information is incorporated into acoustic images of targets. It is widely assumed that spectral cues are used by FM bats, and FM bats certainly can distinguish between targets reflecting echoes that differ in their spectra, but none of this addresses what bats actually perceive when presented with complex targets. From our experiments, Eptesicus perceives images organized along a psychological target-range axis that is formed by combining direct estimates of echo delay with delay estimates arrived at indirectly through representation of the frequency axis of echo spectra or spectrograms.

The results obtained here may reconcile the bat’s extraordinarily fine acuity of less than 1 μs for echo delay (Menne et al. 1989; Moss and Schnitzler 1989; Simmons 1979) with the likelihood that the bat uses echo spectra to perceive fine features of targets (Simmons 1987; Simmons and Grinnell 1988). If the bat uses the spectrum of echoes to determine target shape, this would be a perceptual phenomenon requiring relatively accurate (fractions of a millimeter) perception of differences in range separation from one target’s shape to another, but a phenomenon actually achieved by accurate registration of small changes in the frequencies of spectral peaks and notches (Beuter 1980; Schmidt 1988; Simmons et al. 1974). If the bat’s only use for the range axis in perception is to coordinate reaching out with its wing or tail to seize an airborne target, there would be no need for a target-ranging accuracy of a fraction of a millimeter (Simmons 1987, 1989). To perceive the small range separations between glints in biologically realistic targets (Simmons and Chen 1989), the bat would have to register the absolute range of the target with considerable accuracy—certainly greater accuracy than the 1 to 2 cm needed to intercept insect prey (Trappe 1982; Webster and Griffin 1962). That is, if the bat converts spectrally-represented target shape into a true range-axis estimate of the separation of glints in the target, the range axis would need to have an intrinsic accuracy great enough to accept this transformed spectral information. If the range axis did not have such inherent accuracy, the effort put into preserving details of shape down to a fraction of a millimeter in the spectrum would be lost when the transformation to a coarser time scale takes place. For example, if the bat perceives a change of less than a millimeter in the range separation of two glints in a target, as Eptesicus has been shown to do (Simmons et al. 1974), the axis of range must be perceived in increments at least as small as a fraction of a millimeter to support such perception (Simmons and Stein 1980).

In contrast to the prevailing view of echolocation by FM bats, in which target ranging with delay cues and shape perception with spectral cues are treated as separate processes (Haberzetzer and Vogler 1983; Neuweiler et al. 1980; Schmidt 1988; Suga 1988), we find that perception of range and shape are integrated together through creation of a synthetic perceptual dimension of range that expresses both kinds of information. In principle, this mutual arrangement might be expected in echolocation (Simmons et al. 1974) because the time and frequency domains are no more than different representations of the same information, and FM echoes are represented psychologically by a spectrogram-like code that has both time and frequency dimensions (Altes 1980, 1984). Any property of echoes that might be conveyed by the spectrum and initially encoded on the auditory system’s tonotopic maps will naturally appear in
the time domain as an 'event' along the delay or range axis if the output leading to perception indeed does take a time-domain form after the proper transformation is performed. In this respect, although the representation of echoes is spectrogram-like, the representation of targets is more like the ambiguity function along the time axis. Previous experiments also implicate the time domain as the form taken by the output of the sonar receiver in *Eptesicus* (Simmons 1979; Simmons and Stein 1980). The possibility that FM bats actually might perceive time-domain images (Simmons 1980; Simmons and Stein 1980) typically has been rejected because, although the anatomical basis for spectral coding (tonotopy) is well-known, the physiological basis for fine temporal acuity in the microsecond range is not so obvious (Schmidt 1988). Such fine acuity is, however, demonstrated experimentally by behavioral methods (Menne et al. 1989; Moss and Schnitzler 1989; Simmons 1979; Simmons et al., in press).

The neural response-latency shifts that form the basis for Experiment II amount to about 13 to 18 µs dB, which is equivalent to a shift of about 2 to 3 mm dB in target range. The behavioral experiments show that *Eptesicus* indeed does experience such shifts in the apparent range of targets, so their occurrence must be taken into consideration when thinking about echolocation under conditions where echo amplitudes can be expected to change. The size of the shift is insignificant with respect to the accuracy of 1 to 2 cm needed to capture prey if the change in echo amplitude is only a decibel or two. However, when a bat approaches a flying insect, echoes from that insect can be expected to grow in amplitude by 40 to 60 dB or more, depending upon how large a segment of the pursuit maneuver one looks at (Griffin 1958; Kick and Simmons 1984). For such large changes in echo amplitude, the target-ranging process would surely be greatly disrupted if latency shifts of up to nearly a millisecond were to occur in neural responses to a succession of echoes. Behavioral experiments have also shown that *Eptesicus* regulates its hearing sensitivity as targets become closer by raising its hearing thresholds for echoes at a rate of about 11 to 12 dB for each halving of range (Kick and Simmons 1984). A substantial portion of this regulation may be accomplished by the degree to which the middle-ear muscles are partially contracted at delays up to 5 to 8 ms following each emission. Regulation of echo sensitivity by the middle ear would keep the amplitude of echoes stimulating the inner ear at a constant level throughout a large part of the approach to a target. Experiment II establishes the practical importance of some such mechanism for preventing latency shifts from distorting the perception of target range.

**Neural mechanisms of image formation** - Target range is determined from echo delay (Simmons 1973) and is represented in the bat's auditory cortex by neurons that selectively respond to pulse-echo pairs with particular echo delays (O'Neill and Suga 1982; Suga 1988; Suga and Horikawa 1986; Suga and O'Neill 1979; Sullivan 1982; Wong and Shannon 1988). The frequencies in the FM sweep of echoes are represented topographically by tonotopic maps that are repeated at every stage of the auditory system from the inner ear to the auditory cortex (Suga 1988, for example). Our working hypothesis is that the formation of acoustic images of targets takes place as a result of the convergence of information from frequency maps, which represent the spectrum of echoes using coordinates of frequency and amplitude, with information from range maps, which represent the delay of echoes using coordinates of time and amplitude (Suga 1988). We presume that the conversion of echo spectral information back into range information occurs because frequency maps can 'talk to' range maps, most likely at the level of the auditory cortex.

The mustached bat, *Pteronotus parnellii*, is the species in which the neural representation of echoes has been most extensively studied. Only in this species is the topographical representation of echo delay on computed target range maps well described. In *Pteronotus*, the neural display of target range occurs in at least 3 distinct cortical zones that are thoroughly segregated from other cortical regions representing the frequency of echoes (Suga 1988). *Pteronotus*, however, emits CF FM echolocation sounds and invests heavily in the neural machinery for identification of prey by detecting rapid modulations of CF echoes caused by the wing-beats of insects (Henson et al. 1987; see also Neuweiler et al. 1980; Schnitzler et al. 1983). In contrast, *Eptesicus* and other FM bats are committed to using only FM signals to identify prey. This does not preclude the perception of wing-beats to determine whether targets are in fact insects (Sum and Menne 1988), but the bat must process FM rather than CF echoes to do so. In the FM bat, *Myotis lucifugus*, the tonotopic and range maps are not completely segregated as they are in *Pteronotus*; instead, they are found to overlap one another in one region of the auditory cortex (Wong and Shannon 1988). In addition, the range map appears to be less precisely defined in *Myotis* than in *Pteronotus*. Our very preliminary observations suggest that this is also the case in *Eptesicus*. The nature of the transformation of spectral cues into images has been specified well enough in the behavioral data presented above to raise the possibility that overlap of frequency and time maps may be an integral part of the process whereby spectral and temporal features of echoes are fused.

Our experimental results lead to several predictions about the neural representation of echo frequency in *Eptesicus*. The range separation of glints in a complex target evidently is represented by the location and spacing of peaks and or notches along the frequency axis of the spectrum or spectrogram of echoes. At high echo signal-to-noise ratios, the notches alone provide a better estimate of target shape than the peaks, while at low signal-to-noise ratios only the peaks are likely to be detected (Altes 1984). As Fig. 5 shows for the case of a 100-µs separation of echo components, the width of an individual spectral interference peak and notch in the set created by any given time-separation of overlapping
echo components is the same as any other member of the set. (Other examples can be seen in Schmidt 1988, and Simons et al. 1974). The peaks and notches at 25, 35, 45, 55, 65, and 75 kHz in Fig. 5 are all the same width, for instance. If the auditory system of Eptesicus were matched to the task of determining the location and spacing of multiple spectral peaks and notches as a means of estimating the time-separation they represent, the tuning curves of neurons at different best frequencies from 25 to 100 kHz ought to be of a constant width in absolute frequency units. (This assumes roughly equal weighting of all echolocation frequencies as potential locations of peaks or notches.) While the location and spacing of these spectral notches or peaks is detected along a frequency scale, it is perceived as a spacing of delays between echo components along a time scale - that is, along a scale that is the reciprocal of frequency. It would be difficult for the spectral representation of the range profile of a target to smoothly and seamlessly be transformed into the range profile itself without taking into account this reciprocal relationship. Thus, while we would expect neural tuning curves to be of equal widths in kHz, we would expect the topographic scale of frequency (frequency versus distance or volume of tissue in the brain) to be hyperbolic rather than linear or logarithmic. Such a reciprocal-of-frequency scale would facilitate the organization of regular neural interconnections that can compute the frequency-to-time transform (see below).

What are the tuning curves and tonotopic axis of frequency like in Eptesicus? At present, sufficiently precise published data exist only for the nucleus of the lateral lemniscus (Covey and Casseday 1986) and the primary auditory cortex (Jen et al. 1989). First, Fig. 13 shows the mean neural tuning curves for 3 sets of 5 cortical neurons tuned to average best frequencies of 27, 37, or 67 kHz (replotted from Jen et al. 1989, Fig. 4). Each set of neurons was recorded from a separate perpendicular electrode penetration of the auditory cortex, and the low variability of the original data within each set (indicating a possible anatomical columnar organization relative to frequency) makes these tuning curves particularly valuable for comparison across frequencies. The mean tuning curves are strikingly alike when displayed on a linear frequency scale, confirming our prediction of constant sharpness of tuning in absolute-frequency units. Second, Fig. 14 shows the tonotopic axis of frequency as determined in the nucleus of the lateral lemniscus (A: Covey and Casseday 1986) and the auditory cortex (B: Jen et al. 1989). The lemniscal data are taken directly from the reference, while the cortical data represent the mean best frequencies at each position as shown in the reference. Both anatomical dimensions are presented as percentage distance along the tonotopic map from low to high frequency to make the two parts of Fig. 14 consistent. These data are plotted on a vertical scale that is linear with period rather than linear with frequency. The corresponding frequency axis would be the reciprocal of period, or hyperbolic with frequency. The data-points for the tonotopic maps fall along straight lines, with a correlation coefficient of \(-0.99 (P<0.001)\) in the lateral lemniscus and \(-0.84 (P<0.005)\) in the cortex. It appears as though Eptesicus indeed represents the frequency of echoes along a scale having equal intervals of time rather than frequency. It is particularly important that both equal tuning widths in kHz and equal scale intervals

![Fig. 13](image1.png)

**Fig. 13.** A graph showing the mean tuning curve for each of 3 sets of 5 cortical neurons in *Eptesicus*. Each set is located on the same electrode penetration for a best frequency of 27, 37, or 67 kHz (Jen et al. 1989). The frequency scale is linear rather than logarithmic to display the similarity of the curves in units of absolute frequency.

![Fig. 14 A, B](image2.png)

**Fig. 14 A, B.** Graphs showing the frequency map of *Eptesicus* derived from experiments in (A) the nucleus of the lateral lemniscus (Covey and Casseday 1986) and (B) the primary auditory cortex (Jen et al. 1989). The anatomical position of frequency-tuned cells across neural tissue (see text) is plotted against a best-period scale instead of best-frequency, making the equivalent frequency axis reciprocal or hyperbolic. The regression lines cover frequencies in the echolocation band for *Eptesicus*. Data-points for the lateral lemniscus taken directly from reference, data-points for the auditory cortex represent mean values of frequency for each position.
in microseconds are found in the auditory cortex because our working hypothesis is founded on this level of the auditory system.

Our specific working hypothesis is that neural connections from the frequency-representing tonotopic axis to the time-representing echo-delay axis of auditory cortical maps are responsible for the conversion of spectral peaks and notches into estimates of echo-delay separation between overlapping echo components. The existence of both frequency and delay axes is documented in Pteronotus (Suga 1988), and selective neural responses tuned to frequencies or to delays have been found in Myotis (Sullivan 1982; Wong and Shannon 1988), although the character of the delay map is presently not as well known as in Pteronotus. Our preliminary results from the auditory cortex of Eptesicus are similar in these respects. These two axes probably are the physiological correlates of the time and frequency dimensions of a spectrogram representation of FM echoes. In principle, a spectrogram representation is capable of preserving information about echo delay and target shape across its two-dimensional surface (Altis 1980, 1984), and it would be simple enough to end our interpretation with this observation. However, we do not believe that the mere simultaneity of representations of echo delay along a time axis and spectral notches along a frequency axis can account for the emergence of the second (a2) component of the perceived image in Fig. 4. For this to occur, the locations of notches in frequency must somehow converge onto an appropriate value of delay separation. This convergence requires that appropriate neural wiring explicitly be in place, or at least that activity in neurons tuned to particular frequencies can be read out in units of delay that depend upon the reciprocal of frequency. The transformation of notch-frequency to delay could be achieved by a pattern of neural connections between frequency-tuned cells and delay-tuned cells, and the regularity (and, hence, plausibility) of this pattern would be greatly enhanced if the frequency scale were linear with time rather than frequency.

A hypothetical process of image formation by Eptesicus is shown schematically in Fig. 15. This diagram is to be considered a cartoon of the process rather than a quantitative model in every detail. The target in Fig. 15, which is shown as a moth at range r, is simplified acoustically into two glints. A and B, separated by a range difference of \(\Delta r\). To perceive the target's location and shape, the bat emits FM sonic sounds with a spectrogram structure similar to what is shown in Fig. 7. The echoes arrive after some delay, \(t\), and have two components separated by the time-interval, \(\Delta t\). The size of this time interval is determined by the range separation of the two glints in the target. The bat's auditory system initially encodes the emission and then the echo as spectrograms consisting of patterns of excitation in the Organ of Corti (shaded areas) and neural discharges (dots) marking the time-of-occurrence of each frequency in the FM sweep. The frequency scale for these spectrograms is presumed to be hyperbolic (Fig. 14). The delay of echoes is encoded by the spectrogram delay of the echo representation, which consists of the time, \(t\), between responses to the emission and to the echo at each frequency, \(f\) (Simmons and Kick 1984; Suga 1988). The range separation, \(\Delta r\), of the glints, \(A\) and \(B\), or the time separation of their echo components, \(\Delta t\), appears in the echo as a series of notches in the spectrum (*), separated by a frequency interval, \(\Delta f\), which is the reciprocal of the time separation of the echo components. These spectral notches have a peculiar manifestation in the neural spectrogram as a consequence of the reduced echo amplitude at frequencies near the center of each notch, which creates longer response latencies at these frequencies compared to other frequencies in the echo. The neural spectrogram thus has a scalloped appearance, with the spectral notches being marked by longer latencies as well as weaker response strength.

The spectrograms delays at each frequency are eventually represented topographically on an echo-delay or target-range map in the auditory cortex (O'Neill and Suga 1982; Suga 1988; Suga and Horikawa 1986; Suga and O'Neill 1979; Sullivan 1982; Wong and Shannon 1988). The schematic range map in Fig. 15 is the 'auditory display' for target range. Neurons in the range map respond selectively to a particular region of delay for a particular frequency or span of frequencies. They also selectively respond to a specific range of echo amplitudes. In effect, the cortical location of delay-tuned neurons represents particular values of \(t\), the spectrogram delay. The separation of tuned delays from one neuron to the next, \(d\), along with the sharpness of delay tuning, must constitute the resolution or 'grain' of the range display at the single-cell level. Estimates of the sharpness of delay tuning fall in the region of hundreds of microseconds to milliseconds, which is very large compared to the behavioral acuity in Eptesicus of at least 1 ms. New experiments indicate that this behavioral acuity is less than 0.4 ms (Menne et al. 1989; Moss and Schnitzler 1989) and may be as small as 10 ns (Simmons et al., in press). The spacing of adjacent delay-tuned neurons is unknown in FM bats, but it has been estimated to be on the order of 50 to 100 \(\mu\)s in the CF FM bat, Pteronotus parnellii (O'Neill and Suga 1982; Suga and O'Neill 1979). Presumably, creation of the perceptual range-axis image of a single-glint target, \(A\) (shown schematically at the bottom of Fig. 15 as a shaded peak that is narrower than the summed neural activity on the map), involves integration of estimates of the spectrogram delay across frequencies in the FM sweep. The demonstrated sharpness of the image of \(A\) requires interpolation, and possibly other processes to 'read' the range map, too.

The presence of notches in the spectrum of echoes should have two effects on the neural activity evoked in the range map. First, in the frequency region of the notches less activity will occur, or activity will be confined to neurons that are tuned to lower amplitudes than at other frequencies. Second, estimates of the spectrogram delay for frequencies close to the notch will be displaced to greater delays as a consequence of the greater response latency at notch frequencies. That is,
the scalloped appearance of the neural spectrogram of echoes will be carried over into the activity on the range map when the spectrogram delays are transposed onto a topographic representation. The spectral notches are also represented along tonotopic maps at other sites in the auditory system, including one that overlaps the range map in the auditory cortex. Our behavioral experiments show that the location and spacing of these spectral notches is used to create new events along the delay axis at time separations related to the periods of the frequencies where the notches occur. The hyperbolic or period scale of the frequency axis may facilitate this computation. In Fig. 15 this transformation is shown as a shaded arrow extending from the notch frequencies to the perceptual image of the second glint, B, in the target’s image. We are presently engaged in further behavioral and physiological experiments to identify the nature of this transformation in more detail.

**Fig. 15.** A schematic diagram showing the process whereby acoustic images of complex targets are formed in the sonar receiver of *Eptesicus*. Echo delay is represented neurally by the timing of discharges and eventually displayed topographically on a target-range map. The time-separation of overlapping echo components is represented spectrally by neurons arranged on a hyperbolic tonotopic map. The image perceived by the bat is created by fusing these two neural displays, in the process converting the spectral representation into a delay estimate (see text).

**Conclusions**

The results of the experiments reported above introduce a complicated new view of the process of image-formation in FM echolocation. Through this complexity can be seen the underlying simplicity of the process from the bat’s vantage point, however. In the context of wideband sonar, a complex sonar target can be described as a series of discrete reflecting points, or glints, distributed along the dimension of target range (Altes 1976). The target’s overall range is conveyed by echo delay, and the range separation of the glints that compose the target is conveyed in the temporal and spectral fine structure of the sonar echoes returned to the sonar receiver (Altes 1984). This description corresponds closely with the content of the acoustic images perceived by the FM echolocating bat, *Eptesicus fuscus*.

In Experiment I, bats were trained in a two-choice
sequential-presentation task to discriminate between an electronically-simulated complex target consisting of 2 glints (separated by 17.2 mm or 100 µs) and a simple target consisting of 1 glint. (Sequential presentation was regulated by the aim of the bat's head and sonar sounds towards one channel of the target simulator at a time. The sequential-presentation regime forced the bat to compare images or remembered traces of images to perform the task.) The range of the complex target was about 56 cm (3.275-ms echo delay), while the range of the simple target was varied around this value. The performance of the bats shows an increase in errors when the simple target coincides in apparent range with either glint in the complex target, thus providing an index of the perceived distance of the simple target with respect to the individual elements of the complex target. In the first experimental condition, echoes for both glints in the complex target were presented together, and the bats made errors when the simple target coincided in simulated range with either glint of the complex target. That is, the bats actually perceive the complex target as having a structure along the range axis. In the second condition, only echoes for the first glint of the complex target were delivered, but they were filtered to recreate spectral notches ordinarily produced by overlap of echoes from the two glints. The bats perceived the filtered echoes as though the target contained two glints, indicating that the spectral composition of echoes conveys the glint structure of complex targets. In the third condition, echoes for the two glints of the complex target were presented alternately, from one sonar emission to the next (echo jitter procedure). The bats perceived the range of each glint separately, as before, but without spectral cues caused by overlap of echoes delivered together. The error peaks in the bats' performance curves are similar in height and width in all 3 conditions, indicating that the range profile of the target rather than its intermediate representation by the spectrum of echoes ultimately is responsible for discrimination of the complex target.

The distance to the target's leading edge, or nearest glint, must be represented by the time-of-occurrence of neural discharges encoding the delay of the first component of echoes. The perceived distance to a target is therefore susceptible to errors introduced by changes in echo amplitude, since stimulus amplitude influences neural response latency. Auditory evoked potentials (N1 and N2) recorded from anesthetized bats show a latency shift of 13-18 µs for each dB change in the amplitude of a frequency-modulated (FM) sweep. In contrast, interference notches in echo spectra do not move to new frequencies as echo amplitude changes. The amplitude-dependence of response latencies and the amplitude-independence of spectral notches were exploited in Experiment II to dissociate the image of the complex target into a delay-encoded first glint and a spectrally-encoded second glint. In Experiment II, the bats again discriminated the simulated complex target from the simple target, but with simultaneous presentation as well as sequential presentation of the target echoes. (In the simultaneous-presentation regime, when echoes from both targets are delivered for each sonar emission, masking between echoes can occur directly, without the bat having to rely upon traces of images to perform the task.) The amplitude of echoes for the simple target was changed over a 15-dB range to elicit neural response-latency shifts that would change the apparent range of the target. For each dB change in the amplitude of echoes in the simultaneous-presentation condition, the perceived range of the simple target relative to the first glint of the complex target changes by 2 to 3 mm, equal to a change in echo delay of approximately 13 to 17 µs. However, the perceived distance of the simple target relative to the second glint of the complex target does not change with echo amplitude in the simultaneous-presentation condition. The simultaneous presence of echoes from the simple target masks the spectral notches that represent the separation of the two glints in the complex target, and changes in the amplitude of any of the echoes should not affect the timing of echoes required for the spectral notches to be masked. In contrast, in the sequential-presentation task, the perceived distance of the simple target relative to both glints in the complex target changes by about 2 mm for each decibel change in echo amplitude, indicating that the apparent range of the second glint in the image of the complex target ordinarily is perceived with reference to the absolute range of the first glint.

*Episicus fuscus* converts echo spectral information about the second glint of the complex target into an estimate of the equivalent delay or range separation of the two glints. In this process, the bat refers the spectrally-derived range of the second glint to an absolute range for the first glint that is derived from the timing of neural discharges. Evidently the bat initially develops a mixed time-frequency auditory representation of echoes from a complex target but ultimately expresses the acoustic image of the target along a single perceptual axis which corresponds directly to the natural spatial dimension of the distance or range over which the target is distributed. The output of the bat's FM sonar receiver thus is displayed in the time domain, even though intermediate auditory representations partake of both the time and frequency domains.

The dissociation of the two glints in the complex target in the results of the simultaneous-presentation condition of Experiment II demonstrates the physiological reality of the separateness of the time- and frequency-domain elements of the bat's auditory representation of echoes. In general terms, the initial neural representations of FM echoes resemble spectrograms whose time and frequency dimensions are operated upon with a certain degree of independence, indicating that the bat may achieve the equivalent of spectrogram computations by fusing the outputs of separate time-processing and frequency-processing subsystems within the auditory pathways. The neural basis for the transformation of the output of the frequency-processing subsystem into terms compatible with a time-domain display probably requires convergence of information between tonotopic displays of echo spectra and topographic displays of
echo delay. We presently assume that these are located in the bat’s auditory cortex. The detection of spectral features in FM echoes (particularly interference notches) would be facilitated if the frequency selectivity (tuning curves) of the neurons that represent echo spectra were constant in bandwidth rather than proportional in bandwidth (constant Q) across frequencies. Similarly, convergence of the spectral estimates upon a time-domain display would be facilitated if the tonotopic axis were to be scaled hyperbolically across frequencies even though the tuning curves themselves would have constant absolute-frequency widths along this hyperbolic scale. Both of these physiological conditions appear to be the case in *Eptesicus*

Acknowledgment: This research has been supported by Office of Naval Research Contract No. N00014-86-K-0401, by Grant No. NNS83-0244 and prior grants from the National Science Foundation, by SIMH Research Scientist Development Award No. MH00621, by a grant from the System Development Foundation, and by a University Research Instrumentation Grant from the Department of Defense. We thank R.A. Altes, R.R. Capranica, J.H. Casseday, E. Covey, D.R. Griffin, A.D. Grinnell, J. Melwani, D. Menne, H.-U. Schnitzler, N. Sugiyama, and W.E. Sullivan for their constructive suggestions. These experiments were conducted in response to challenges posed by J.J. O’Hare, H. Hawkins, and others at the Office of Naval Research in a series of meetings held in 1985 and 1986. We are grateful for having the opportunity to address them.

References


Bodamerhuber RD, Pollak GD, Marsh DS (1979) Coding of the fine frequency information by echolocating neurons in the inferior colliculus of the Mexican free-tailed bat. Brain Res 173:530-535


Covey E, Casseday JH (1986) Connectional basis for frequency representation in the nuclei of the lateral lemniscus of the bat *Eptesicus fuscus*. J Neurosci 6:2926-2940


Simmons JA (1979) Perception of echo phase information in bat sonar. Science 207:1336-1338


Webster FA, Griffin DR (1962) The role of the flight membrane in insect capture by bats. Anim Behav 10: 332–340