A level of stimulus representation model for auditory detection and attention

Ervin R. Hafter[a] and Kourosh Saberi[b]
Department of Psychology, University of California, 3210 Tolman Hall, Berkeley, California 94720

(Received 21 May 1998; revised 29 May 2001; accepted 1 June 2001)

A model is offered here to address an asymmetry of cueing in signal detection [Hafter *et al.* (1992)] where the effect of frequency uncertainty on the detection of a randomly chosen tone was ameliorated by cueing with a sequence of its harmonics, but detection of a randomly chosen sequence of harmonics was not improved by cueing with their fundamental. The model proposes that signal detection can be based on various levels of neural representation that, for the case at hand, refer to levels organized either by frequency or by complex pitch. Experiments offered to test the model used three-tone complexes for both cues and signals. These stimuli consisted of either three randomly chosen frequencies or three randomly chosen harmonics (from the set $2f_1$ to $7f_1$) of a randomly chosen fundamental. Support for the idea of cueing and detection at different levels of representation was found in higher performance with uncued detection of harmonic complexes relative to that found with complexes of unrelated tones and by successful cueing of each type of information with cues created to remove uncertainty about the relevant information. A final comparison suggests independence of performance (presumably of the limiting noise) at each of the putative levels of representation.
[DOI: 10.1121/1.1394220]

PACS numbers: 43.66.Hg, 43.66.Dc, 43.66.Lj, 43.66.Fe [JWH]

I. INTRODUCTION

A. Uncertainty reduces performance

The peripheral auditory system is generally characterized as a bank of bandpass filters called critical bands, or auditory filters, whose widths are roughly proportional to their center frequencies. Because they derive from neuromechanical processes in the cochlea, these filters are commonly thought to be immutable in shape and bandwidth. However, the possibility of top-down control over their *effective* bandwidths has been discussed in speculations on the role of efferent neural connections to the cochlea (Scharf *et al.*, 1994, 1997), proposed as the basis for interactions between motivational instructions and the effects of frequency uncertainty in signal detection (Hafter and Kaplan, 1976), and observed directly in a study of uncertainty using a probe-signal method (Schlauch and Hafter, 1991).

Traditional psychoacoustic descriptions of the auditory filters have relied upon studies of the detection of pure-tone signals in the presence of noise. An indirect approach infers the bandwidths from so-called ''critical ratios'' defined as the signal level at threshold divided by the spectrum level (level/Hz) of a wideband masker, while more direct methods describe ''critical bandwidths'' in terms of the relation between performance and the bandwidths of either band-limited maskers or spectral notches in a wideband masker. (For review, see Scharf, 1970; Patterson and Moore, 1986; Moore, 1997.) In all such measures, it is tacitly assumed that the subject monitors and responds only to frequencies falling within the appropriate auditory filter which, in turn, means that he or she *knows* the signal's frequency. If it is unknown, performance must decline, even for the ''ideal observer'' of signal detection theory (SDT) (Green and Swets, 1966) due to the increased probability of large peaks in the noise appearing in wrong, i.e., nonsignal filters. Studies where the signal's frequency has been drawn at random on each trial from a list of $M>1$ possibilities have found that the masked threshold relative, to the case of $M=1$, rises to an asymptotic value of 3 to 5 dB for large values of M (e.g., Green, 1961; Schlauch and Hafter, 1991).

B. Reducing uncertainty with cues

The deleterious effects of frequency uncertainty on detection can be reduced or even eliminated by presenting pretrial cues that tell the subject what to listen for. The most effective such cue is a tone matched in frequency to the signal (e.g., Swets and Sewall, 1961; Hafter and Kaplan, 1976; Johnson and Hafter, 1980; Schlauch and Hafter, 1991; Dai *et al.*, 1991). While this implies a crucial role for shared phenomenology whereby a cue works because it ''sounds like'' the signal, successful reductions of uncertainty have also been found with a variety of cues that are not identical to the signal. These include a tone whose frequency relates to the signal by a small integer ratio such as 5/4 (the musical third) (Hafter and Kaplan, 1976) or 3/2 (a musical fifth) (Hafter *et al.*, 1993), a melodic sequence of tones for which the signal is a musically acceptable extension of the melody (Howard *et al.*, 1984, 1986), a chord made up of harmonic frequencies whose missing fundamental is the same frequency as the signal (Hafter *et al.*, 1992), a multi-tonal complex made up of randomly chosen frequencies, one of which matches the signal (Schlauch and Hafter, 1991), and even a

[a]Author to whom correspondence should be addressed. Electronic mail: hafter@socrates.berkeley.edu
[b]Now at: Department of Cognitive Sciences, University of California, Irvine, CA 94720-1650. Electronic mail: kourosh@uci.edu

visual cue that describes the frequency of the signal in musical notation for subjects who have absolute pitch (Plamondon and Hafter, 1990). Thus, the "sounds-like" hypothesis must be extended to include comparisons made in the *mind's ear* between similar percepts that arise through different auditory mechanisms.

C. Is "sounds alike" sufficient for cueing?

The present experiment was begun in part because of a result that did not seem readily explained by a sounds-like hypothesis. In this condition, where the stimuli used for cues and signals shared a common percept, their pitch, reduction of the effects of uncertainty was asymmetric, depending on their order of presentation. One stimulus was a single, randomly chosen tone whose frequency was dubbed f_1; the other was a harmonic sequence of either $2f_1 \rightarrow 6f_1$ or $3f_1 \rightarrow 7f_1$. While a pure tone and a set of its harmonics are quite different from one another in timbre, a subject asked to adjust the pitch of a pure tone until it matches the pitch of a set of harmonics typically picks the fundamental frequency (f_1) of the sequence, even when f_1 itself is absent from the complex. The common pitch of tones and complexes is thought to arise through separate auditory mechanisms. "The former has traditionally been called a *place* pitch in reference to a place of maximum displacement in the cochlear excitation patten, though analysis of periodic activity in auditory neurons shows that information about frequency also exists in the time-domain, even for tones." Conversely, the pitch of the latter relies on further analysis of the complex stimulus based on interactions between its components. Variously called a "residue," "periodicity," or "virtual-pitch" (e.g., Schouten, 1940; Licklider, 1956; Terhardt, 1974; Terhardt *et al.*, 1982), it has been the object of a long-standing debate about its origin. Whether it is derived in the time or frequency domain is irrelevant to the present discussion and so we will use the more neutral term, *complex* pitch. (Excellent discussions of the mechanisms of pitch can be found in De Boer, 1976; Houtsma, 1995).

As described earlier, Hafter *et al.* (1992) found that a set of harmonics used as a cue improved detection of a randomly chosen tone set to its missing fundamental.[1] However, when those roles were reversed, that is, when randomly chosen f_1's were used to cue signals that were a sequence of their harmonics, there was no improvement due to cueing. We will refer to this dependency on order of presentation as an asymmetry in cueing.

II. A MODEL OF CUEING BASED ON LEVELS OF STIMULUS REPRESENTATION

A. The level of stimulus representation used for detection or LSRD

In seeking to explain the asymmetry of cueing described earlier, this model concentrates less on the shared phenomenology of a "sounds-like" hypothesis than on shared elements at the level of stimulus representation or LSR whose neural activity serves as the basis for signal detection. We call this the level of stimulus representation used for detection or LSRD. It is well known that various features of an acoustic stimulus are represented at multiple LSRs throughout the auditory nervous system (Pickles, 1988); for example, tonotopic maps of acoustic frequency have been identified in regions ranging from cochlea to cortex. While band limitations found in detection experiments are often said to result form "cochlear filters," what is generally meant by this is that the filtering began in the cochlea and not that the subject's judgments were based directly on neural activity at the interface of cochlea and auditory nerve. Because it is unlikely that a neural site simply relays information, one assumes that additional processing at each LSR provides a unique version of the original stimulus. By further assuming that the decision process in a signal-detection task has access to many, if not all, of the LSRs, it follows that best performance requires that the LSRD be the level whose data present the highest signal-to-noise ratio (S/N). From this perspective, the optimal subject must use knowledge of the signal's parameters to select both the optimal LSRD and the appropriate elements within it. This is especially important for multidimensional signals whose potential LSRDs represent different stimulus dimensions.

B. Formal assumptions of the LSRD model

(1) Neurons in LSRs are organized topologically in accord with their sensitivity to values along a stimulus dimension. The dimension may be based on a primitive feature of the stimulus such as acoustic frequency or on a more complex feature that derives from interactions between primitives. Examples of the latter include complex pitch and locations in auditory space.

(2) Each neuron in an LSRD is best tuned to a specific value of the represented dimension but responds, to some extent, to nearby values of the dimension falling into its "receptive field." Pooled responses from adjacent receptive fields then determine the effective bandwidth of a masker, providing a kind of filter in the represented dimension.

(3) For signals represented at multiple LSRs, optimal performance requires that the LSRD be the level providing the highest S/N. When signals are represented both by primitive dimensions and by complex interactions between those primitives, the larger S/N will generally be found at the higher-order representation (for a discussion, see Sec. II C).

(4) For maximum effectiveness in reducing signal uncertainty, a cue should specify, unambiguously, both the level being used as the LSRD and the correct filter within that representation.

C. Application of the model to tones and multi-tonal complexes without uncertainty

In the framework of signal detection theory (SDT) (Green and Swets, 1966), each observation is judged in accord with the likelihood (λ) that it arose from signal-plus-noise rather than from noise alone. For detection of a single tone of *known* frequency in a background wideband noise, the optimal decision rule is to compute λ on the basis of energy within the single auditory filter centered on the frequency of the signal. Thus, in terms of the model, the LSRD used to detect those signals is organized according to frequency. If the frequency of the signal is *unknown*, the optimal subject must calculate a likelihood for every filter in the LSRD that might contain the signal, and performance must decline due to the increased probability of high likelihoods produced by noise alone in nonsignal filters.

For a multi-tonal signal made up of J *known but unrelated* frequencies, the optimal rule is to decide on the basis of the product of the J individual likelihood ratios,

$$\lambda = \prod_{i=1}^{J} \lambda_i , \tag{1}$$

based on energies in the appropriate J auditory filters. While this rule could also be used for a signal made up of J harmonically related frequencies with a known fundamental, an alternative decision rule might be to listen for the signal's complex pitch. In that case, detection would be limited by a pitch-masker which both reduced the effectiveness of signals by adding noise to their individual components and produced false pitches through the accidental occurrence of peaks in the acoustic noise at harmonics of the signal. In terms of the LSRD model, detections would be based on neural activity at an LSRD that is topologically arrayed according to complex pitch. Similarly, the extent to which adjacent elements are pooled at this level would define the effective bandwidth of the pitch masker. Unfortunately, without knowing the statistics of the putative pitch masker, one cannot say whether a J-tone harmonic complex with *known* frequencies would be better detected on the basis of the J individual frequencies or on its complex pitch.

D. Application of the model to multi-tonal complexes with uncertainty

The situation with J-tone signals is quite different with frequency uncertainty. If the individual tones are unrelated, the best LSRD is still one organized by frequency, only here, the optimal statistic, $\Pi\lambda_i$ [Eq. (1)], is based on the J highest individual likelihoods found across all auditory filters in the range of possible frequencies. However, when the unknown frequencies in the signal are known to be harmonically related, albeit with no knowledge of their f_1, the superior strategy is to listen for and respond to the emergence of a complex pitch. That is because noise alone can produce a false positive in the pitch domain only if the highest peaks in the acoustic noise happen to fall into filters related by a common fundamental frequency. Thus, signal uncertainty should have less of an effect on performance using a LSRD organized by complex pitch than one organized by frequency.

The fourth assumption of the model was proposed to address the asymmetry of cueing described in Sec. I C. According to the model, a missing-fundamental harmonic was a useful cue for detection of f_1 because it specified the single, appropriate filter for detection at a LSRD organized by frequency. Conversely, cueing was ineffective in the reverse condition because a single tone does not specify, unambiguously, a single complex pitch. Consider, for example, a tonal cue of 600 Hz. If treated as part of a complex pitch, it could thought of as the second harmonic of 300 Hz (an octave), the third harmonic of 200 Hz (an octave plus a just-fifth), or the fifth harmonic of 120 Hz (two octaves plus a just-major third).

E. A test of the LSRD model

While the model offered above would seem to explain the asymmetry of cueing with sets of harmonics and their fundamentals, it arose as a *post hoc* analysis of that result. The present study was planned as a more rigorous test of the model, restricting comparisons to stimuli more similar to one another on features not directly addressed in predictions of the model. Here, all stimuli would be three-tone complexes whose frequencies would be either unrelated to one another, thus offering no complex pitch for detection, or related by a common fundamental. As discussed in Sec. II D, under conditions of signal uncertainty, the model predicts better performance with harmonic complexes than with the randomly chosen tones. In line with assumption 4, the cues would also all be three-tone complexes whose frequencies would be chosen to ameliorate uncertainty about either the individual frequencies in the signal, its complex pitch, or both.

III. PROCEDURE

A. Stimulus generation

The experiment compared performance across five conditions, each of which measured the detectability of three-tone complexes. Frequencies in the signals were different on every trial, but the levels of the three tones were set to be equally detectable through reference to an empirically derived, equal-detectability function (EDF). This function was found by measuring thresholds for the three subjects at five frequencies covering the range from 400 to 4725 Hz. A straight line was fitted to these data in dB/Hz to provide an approximation to an EDF (Green *et al.*, 1959; Schlauch and Hafter, 1991). When the individual differences between the subjects proved to be insignificant, a single, averaged EDF was constructed and used throughout the experiment for all subjects. Signals were generated digitally, with a sampling rate of 50 kHz, and played through a locally constructed 16-bit D/A converter and a low-pass filter with a cutoff frequency of 20 kHz and slope of 48 dB/oct (Frequency Devices Model 901).

The continuous wideband masker was produced by an analog white-noise generator and filtered only by the frequency response of the Stax (SR5) electrostatic headphones. Its spectrum was essentially flat across the frequency range of interest. The spectrum level of the noise was 20 dB SPL as determined with a Hewlett-Packard (3582A) spectrum analyzer.

B. Psychophysical measurement

Performance was measured in a two-interval, forced-choice psychophysical (2IFC) task which presented signals with equal probability in one of two 300-ms intervals. Times between those two intervals as well as between the cue

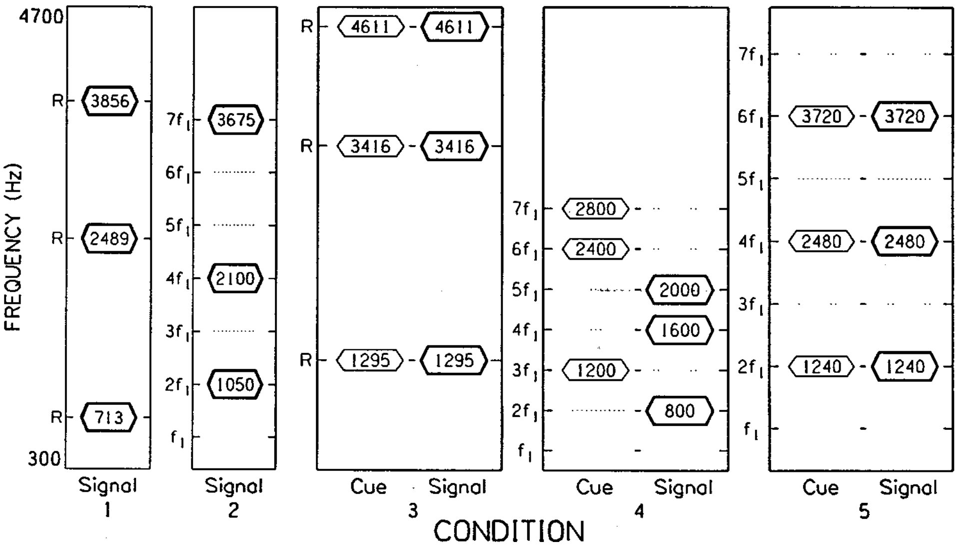


FIG. 1. Time-lines descriptive of five kinds of trials from the present experiment. In condition 1, uncued signals were three randomly chosen tones (see text). In condition 2, uncued signals were three harmonics chosen at random from the set $2f_1 \rightarrow 7f_1$ of a randomly chosen f_1. In condition 3, a random signal as in condition 1 was preceded by a cue matched to it in frequencies. In condition 4, a harmonic signal as in condition 2 was preceded by a cue matched to it in complex pitch. In condition 5, a harmonic signal as in condition 2 was preceded by a cue matched to it in both frequencies and pitch.

(when present) and the first interval were 250 ms. Durations of the cues and signals were 300 ms, including 10-ms linear onset and offset ramps. Each trial was terminated by the subject's response and followed by visual feedback that identified the correct response interval. Proportions of correct responses, $P(C)$'s, were obtained over blocks of 50 trials. Each subject ran at least seven blocks of each condition. Experimental sessions generally consisted of ten blocks, with only a single stimulus condition presented within a block. Before each block, a subject was allowed as many practice trials as he or she wished, although such practice trials never exceed ten. All conditions were practiced until performance seemed stable before the actual experiment began; at that point, the order of conditions was randomized and each subject was tested in a different order. Subjects were students at the University of California, including one of the authors, KS. All reported normal hearing. Testing was done in a double-walled, audiometric listening booth.

C. Signals and cues

A single EDF was used to set levels of cues and signals throughout all five conditions of the experiment. Thus, conditions are compared in units of performance. The five stimulus conditions are described below. In addition, a spectrogramlike depiction of a representative trial from each condition is shown in Fig. 1. In order to save space, the figure is not drawn to scale and interstimulus intervals are omitted. Cues, when used, were 6.3 dB higher than the signals, leaving them weak but clearly audible. To be more specific, each component in a cue was 6.3 dB higher than the level of that same frequency if drawn from the EDF used to generate signals.

1. Condition 1: Random-complex signals: No cues

The idea here was to present useful information only at an LSRD organized by frequency. To this end, each signal consisted of three unrelated tones selected at random from a uniform distribution of frequencies that ranged from 400 to 4725 Hz, with the sole restriction being that the ratio between adjacent frequencies must exceed 1.10. In the example shown in Fig. 1, the three unrelated frequencies are 713, 2489, and 3856 Hz. We assumed that this condition would provide the poorest performance, thus allowing room for improvement in conditions with less uncertainty. For this reason, pretesting was used to pick a signal level that would produce especially weak scores of $P(C) \sim 0.60$, a value well below the $P(C) = 0.75$ conventionally used to define threshold. The EDF so-chosen was anchored at 500 Hz to 33 dB SPL, corresponding to a signal-energy-to-noise-power ratio (E/N_o) of 7.7 dB. To reiterate, this EDF was then used for all five conditions.

A subject in condition 1 could, of course, adopt a nonoptimal strategy that ignored the fact that there were three tones in the signal and respond to magnitudes of the two or even single largest values of λ across the range of frequencies. In order to see if subjects were doing that, informal tests were run during the preexperimental period with signals consisting of either one or two randomly chosen tone(s). Performance was lower with only two tones and still lower with one and, leading us to conclude that subjects in the experiment proper would listen for (at least) three tones.

2. Condition 2: Harmonic-complex signals: No cues

Unlike the case in condition 1, the three tones in these signals bore a simple harmonic relation to one another. At the beginning of each trial, a fundamental frequency (f_1) was selected at random from the range 200 to 675 Hz. The next six harmonics of that fundamental ($2f_1 \rightarrow 7f_1$) were computed and, from these, three were selected at random to be the signal. Thus, no signal contained f_1, and the harmonic number of components in the signal differed from trial to trial. This procedure ensured a minimum ratio between adjacent components in the signal of 1.17 (7/6). In the example in Fig. 1, the randomly chosen f_1 is 525 Hz and the randomly chosen tones in the signal are $2f_1$ (1050 Hz), $4f1$ (2100 Hz), and $7f_1$ (3675 Hz). While the individual components were represented by increased energy in the three auditory filters, just as in condition 1, this signal also presented information potentially useful for detection on the basis of its complex pitch (525 Hz). The model predicts higher performance based on the complex pitches of these signals than on the three frequencies drawn purely at random in condition 1.

3. Condition 3: Random-complex signals: Cues matched to the signal's frequencies

Here, the three frequencies were selected purely at random, as in condition 1, but now each trial began with a three-tone cue made up of the same three frequencies. The example in Fig. 1 shows both cues and signals at 1295, 3416, and 4611 Hz. We call these frequency or *F* cues because, in terms of the model, their effectiveness should indicate specification of the appropriate auditory filters at an LSRD organized by frequency. The prediction here is of performance better than that found in condition 1.

4. Condition 4: Harmonic-complex signals: Cues matched to the signal's pitch

Harmonically related signals here were chosen in the same way as in condition 2, but each trial now began with a cue intended to remove uncertainty about the signal's complex pitch without sharing its elements at a LSR organized by frequency. For this, the cue was a three-tone harmonically related complex with the same fundamental as the signal but composed of different harmonic numbers. Thus, after three of the components $2f_1 \rightarrow 7f_1$ had been designated as the signal, the remaining three made up the cue. In the example in Fig. 1, the randomly chosen f_1 was 400 Hz and the randomly selected components for the signal were $2f_1$ (800 Hz), $4f_1$ (1600 Hz), and $5f_1$ (2000 Hz). The cue was made up of the remaining components, $3f_1$ (1200 Hz), $6f_1$ (2400 Hz), and $7f_1$ (2800 Hz). Again, cues were 6.3 dB higher than signals. We call these pitch or *P* cues because, in terms of the model, their effectiveness would indicate specification of the appropriate pitch filters at a LSRD organized by complex pitch. Thus, the prediction is of higher performance than in condition 2.

5. Condition 5: Harmonic complexes: Cues matched to the signal's frequencies and pitch

Signals here were again harmonic complexes chosen as in conditions 2 and 4. However, the cues were made up of the same three harmonics as the signals. Thus, in the example in Fig. 1, both signals and cues were $2f_1$ (1240 Hz), $4f_1$ (2480 Hz), and $6f_1$ (3720 Hz) of the randomly chosen f_1 (620 Hz). We call these "bi-dimensional" cues FP because they presented information about both the individual frequencies in the signal and its complex pitch. The idea was to see if these cues could enhance performance by cueing both the individual frequencies as in condition 3 *and* the complex pitch as in condition 4. If signals at the putative LSRDs were limited by independent noise, informational enhancement produced by the FP cues should be additive. Using $(d')^2$ as the SDT measure of transmitted information (Green and Swets, 1966), the prediction would be that the d' for condition 5 should equal the root-mean-square (rms) value of the d'-values found in conditions 3 and 4.

TABLE I. Summaries of the individual scores as well as averages across the three subjects for the five experimental conditions depicted in Fig. 1. Individual performance in proportion correct, $P(C)$, as well as means across subjects.

Condition	Cues	Signals	Subject 1	Subject 2	Subject 3	$\bar{x}$
C1		Random	0.630	0.597	0.571	0.599
C2		Harmonic	0.702	0.732	0.678	0.704
C3	F	Random	0.764	0.788	0.751	0.763
C4	P	Harmonic	0.827	0.780	0.765	0.791
C5	FP	Harmonic	0.893	0.917	0.913	0.908

IV. RESULTS

Results from the individual subjects as well as the averaged means are shown in Table I. The averaged data are also plotted in Fig. 2 for visualization. The significance of differences between conditions predicted by the model were tested through use of individual *z*-score tests as described in the Appendix.

A. Detection based on complex pitch

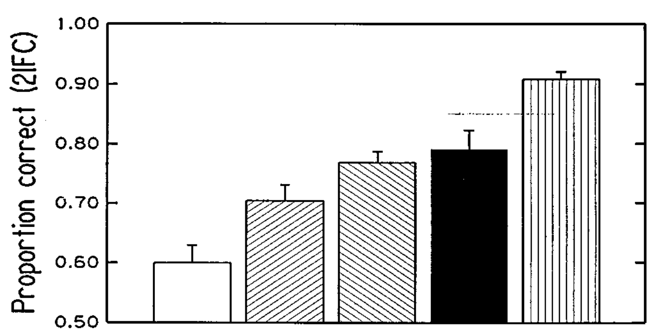
Based on the LSRD, we predicted in Sec. II D that subjects should be better at detecting the harmonically related complexes in condition 2 than the unrelated complexes of condition 1. This prediction, C2>C1, proved to be true (see the Appendix).

B. Cueing at the level of frequency

Because cues matched in frequency are highly effective for one-tone signals, one would expect a similar improvement using matched-frequency (*F*) cues with the unrelated three-tone signals. This predicted amelioration of frequency uncertainty, C3>C1, was confirmed (see the Appendix).

C. Cueing at the level of complex pitch

Results discussed in Sec. IV A suggest that signals in condition 2 were detected on the basis of an emergent property of the relation between their components, a complex



Condition	1	2	3	4	5
Signals	Rand.	Harm.	Rand.	Harm.	Harm.
Cues			Same	Other 3	Same

FIG. 2. Mean performance in the five conditions described by Table I and Fig. 1. Data are shown as the proportion of correct responses, averaged across three subjects. Error bars depict the standard error of the means.

TABLE II. Data obtained with bi-level cues as well as predicted results for condition 5 based on summation of the information conveyed in conditions 2 and 4 [see Eq. (1)].

	Subject 1	Subject 2	Subject 3
Obtained $P(C)$	0.89	0.92	0.91
Predicted $P(C)$	0.82	0.86	0.84

pitch. Given this, the model suggests that amelioration of pitch-uncertainty requires use of cues matched in pitch (P). This prediction, C4>C2, was also confirmed (see the Appendix).

D. The effect of bi-dimensional cueing at both LSRDs

Although successful cueing in condition 4 would seem to indicate that P cues worked by alleviating uncertainty about the complex pitch of the signals, the $P(C)$'s were not much different from those found with frequency-matched cues in condition 3. Thus, one could argue that subjects in condition 4 might have used a knowledge of harmonicity to determine the frequency of f_1 from the cue and then calculated the tones in the signal for detection at an LSRD organized by frequency. However, this idea was dispelled by results with the FP cues in condition 5. If one assumes independence of the limiting noise at LSRDs organized by frequency and by complex pitch, performance in condition 5 measured in d' should reflect the rms sum of the d's found with the one-dimensional cues in conditions 3 and 4 (Green and Swets, 1966). In order to test this hypothesis, $P(C)$ values from C3 and C4 were converted to d's using Elliot's tables (in Swets, 1964) and used to predict performance with bi-dimensional cues (FP):

$$d'_{\mathrm{FP}} = \sqrt{[d'_F]^2 + [d'_P]^2}. \tag{2}$$

Predicted values of d'_{FP}, converted back into $P(C)$ for comparisons to the obtained data, are presented in Table II. It shows that the obtained values were actually slightly higher than those predicted from the combination of information in Eq. (2). While hyper-additivity of this type is not predicted by classic SDT, one might speculate that the two sources of information in the FP cues somehow enhanced each other's effectiveness. For example, knowing the pitch might have helped the listener to focus more precisely on the appropriate elements in an LSRD organized by frequency and *vice versa*. Regardless, by showing that performance was as least as good as that predicted by additivity of cued performance at the two putative LSRDs, this lends further support to the hypothesis of independent accessibility to information at separate LOPs organized by frequency and complex pitch. Seeking statistical support for the independence-of-cueing hypothesis, we compared the case where cues carried both the frequencies and the pitch of the signals to the one in which the subjects were cued with pitch alone. In support of independence, C5>C4 also proved to be highly significant (see the Appendix).

V. DISCUSSION

A. Detection based on a complex feature of the signal

We have argued that the special performance found in conditions 2, 4, and 5 support the notion that the subjective decision maker had access to neural data at a neural level that specifically represents complex pitch. While it is difficult to point to direct evidence of LSR topologically organized by complex pitch, the musical perception of a sequential relation between successive notes, e.g., C, C♯, D, D♯, etc., even with missing-fundamental, harmonic complexes, would seem to suggest one. Furthermore, while direct physiological results of such an organization have been scarce, studies of the neural code for amplitude modulation and complex pitch (e.g., Schreiner and Langner, 1988; Langner *et al.*, 1997) lend credence to its existence, and one expects that the advent of new brain-imaging techniques will clarify this important issue in the near future. Interestingly, our prediction of improved detectability of a complex based on the relation between its tones reflects a more general principle, namely that detection of any complex signal under stimulus uncertainty should be better if the judgments are based on the relationship between its primitive components rather than on an independent analysis of its primitives alone.

B. Assumptions of the model

Formal assumptions of the LSRD model were made purposefully strong to simplify predictions for the experiments. While assumptions 1–3 seem well justified by the apparent usefulness of complex pitch as a dimension for detection, the assertion in assumption 4 that the shared representation between cue and signal at the LSRD must be unambiguous was probably overstated. A softer proposal might say that while such representation is necessary for maximally effective cueing, a partial relation between the cue and signal could partially reduce the effects of uncertainty. In support of the softer view, we point to the condition described in footnote 1, where a five-component, missing-fundamental harmonic sequence improved the detectability of a signal set to its fundamental frequency, albeit by not as much as a tone of the same frequency as the signal. While performance with the missing-fundamental cue was significantly higher than with no cue at all [$P(C)=0.835$ *vs* 0.68], it was less than with a single-tone cue set to the same frequency as the signal [$P(C)=0.92$] (Hafter and Schlauch, 1989). Whether this was because the multi-tonal cue pointed to additional filters such at subtharmonics of the fundamental or because it produced a correct but poorly defined representation at the LSRD organized by frequency, we simply do not know.

C. An alternative explanation for cueing at the level of pitch

A reviewer of an earlier submission of this article (Darwin, 1995) pointed out a potential confound in our methodology that might also have produced an improvement from condition 1 to 2. He noted that while purely random frequencies in condition 1 were drawn from a distribution whose upper boundary was 4725, the harmonically related frequencies in condition 2 were drawn from a distribution whose trial-by-trial upper limit was f_7 of a randomly selected f_1, ranging from 1400 to 4725 Hz. Thus, better performance in condition 2 might simply have reflected less frequency uncertainty. Because of this, the second author recruited a new crew of subjects at the University of Florida to retest the comparison between conditions 1 and 2 when tones in each condition were drawn from the same distributions. To this end, each trial in the revised condition 1 began with the random choice of a separate, range-setting frequency (f_R) from the same 200 to 675-Hz range used to select f_1 in condition 2. Thus, the three unrelated tones for that trial were drawn at random from the range $2f_R$ to $7f_R$, the same as for the related harmonically related tones in condition 2. Concerns that the reduced uncertainty was fully responsible for the results in Fig. 3 were dispelled when thresholds (from a tracking procedure) for the new condition 2 were significantly lower than those for new condition 1 ($p<0.05$)[2].

D. Failure to find successful cueing of a fundamental frequency by a matched complex pitch in "informational masking"

McFadden (1988) did not find what he called "periodicity cueing" for the amelioration of "uncertainty" in a case where the signal was a pure tone and the cue a four-tone sequence of its harmonics. However, we do not find this incompatible with the cueing reported here. In McFadden's study, the masker was a set of six other tones whose frequencies did not relate in a harmonic fashion to the signal. They were played in a temporal sequence, three before the signal and three after it. Because the masker tones were chosen so as to not affect the auditory filter centered on the signal, his paradigm falls into a class often referred to as informational masking, with uncertainty referring to the order of presentation of the individual tones in the masker. Because Watson and Kelly (1981) had shown that masker uncertainty of this kind depresses performance, McFadden (1988) thought that presenting the harmonically related cue during the signal interval would reduce the effects of uncertainty and thus increase detection. When this did not turn out to be the case, he reasoned that "...one might conclude that periodicity cueing does exist for sensory masking but not for informational masking." We agree. From our perspective, a cue chosen to reduce uncertainty about a signal works by reducing the number of potential signals to be listened for and hence the number of filters that must be monitored. Thus, when the masker changes but the signal is always the same, one should not expect that cueing the signal would have an effect.

E. The effects of differences in timbre between cues and signals

The purpose of condition 4 was to see if signals made up of harmonic complexes would be successfully cued by their complex pitch if there were no actual frequencies in common. On first listening to these stimuli when setting up the experiment, we worried that cueing at the level of complex pitch might be overshadowed by the large differences between the timbres of cues and signals. Less problematic, but still of concern, might be differences between them in pitch or pitch strength, the worst case being the rare occasion when a cue was made up of only even harmonics and the signal only odd. However, these fears proved to groundless when the three-tone complexes were presented in noise because at low S/N, the cues and signals sounded remarkably alike in pitch and timbre, despite the differences in harmonic numbers. This is reminiscent of the one-tone residue reported by Houtgast (1976). He presented three-tone harmonic complexes (not including f_1) as standards and asked subjects to detect small changes in the f_1 of a three-, two- or one-tone test signal whose harmonic numbers differed from those of the standard. He found that while the task could, to a small extent, be done when the tones were presented in quiet, it was much easier if they were heard in a background noise. Indeed, with noise, subjects reported hearing the pitch of the fundamental in the test signal, even when it had only a single harmonic. While pattern matching models of pitch (see Moore, 1997) easily address the commonality of pitch between cues and signals in our condition 4, they do not speak to the similarity of timbres at low S/N. However, this seems less puzzling if one considers that while only three of the harmonically related filters from $2f_1$ to $7f_1$ received weak tones, all of the important harmonically related filters, including the one at f_1, were filled with noise. As such, the weak tones may be thought of as having acting as seeds, highlighting a specific complex pitch which then recruited noise-based energy in all of its first 7 harmonics to produce essentially the same noisy pitch and timbre, regardless of the seeded frequencies.

F. Phenomenology and cueing?

An important factor not directly addressed by the model is the role of phenomenology shared by the cue and signal. One could postulate that trial-by-trial feedback led subjects to attend to the appropriate elements in the LSRD without insisting that the cue and signal sounded alike, but it seems more plausible that perceived qualities of the cue should guide the listener both in selection of the appropriate LSRD and of the best filter in it. Thus, while the simple sounds-like hypothesis of cueing with its emphasis on conscious awareness of stimulus features seems insufficient for the asymmetry of cueing found with a fundamental and its harmonics, it seems equally clear that phenomenology must play an important part in focusing the attention on the stimulus dimension to be monitored during search and detection.

VI. SUMMARY

An important factor in signal detection is the extent to which the subject knows what the signal will be. When there is uncertainty about some property of the signal, it is necessary for the subject to monitor more potential filters in the stimulus domain, raising the probability of more false positives due to peaks in the masker. Typically, uncertainty is reduced in the laboratory by offering practice trials, postresponse feedback, and cueing. The present study concentrates on the latter, that is, on the effects of presenting sensory cues that inform the subject what to listen for. In order to stress the importance of cueing, there was a high degree of signal uncertainty, with signals differing from one another on every trial. While it seems obvious to say that successful cueing must, somehow, elicit an internal representation of the signal to be detected, this study began in response to an observation that a complex cue improved the detection of one of its primitives, but that use of the primitive as a cue had little or no effect on detection of the complex. The model proposed to explain this asymmetry notes that complex signals are apt to be represented at multiple neural sites, some organized according to such primitive features as acoustic frequency and others on emergent features, such as complex pitch, derived from interactions between primitives. Based on the argument that ideal performance requires that the level of stimulus representation used for detection, or LSRD, should be the one with the highest S/N, the model predicts that best performance with a complex signal will occur when it is detected on the basis of its complex feature. Finally, the model assumes that successful cueing requires that the cue and signal share unique neural elements at the LSRD.

Tests of the LSRD model offered here used signals that were three-tone complexes chosen at random for each trial. In support of the assumption that complex features would provide a higher S/N than primitive ones, performance was weakest in the condition where the three tones bore no relation to one another, making it necessary to detect them on the basis of the individual frequencies. Performance improved when the three tones were related harmonically, providing a complex pitch as the basis for detection. Cues chosen to test the final assumption of the model were also three-tone complexes. In support of the argument that they must share representation with the signal at the appropriate LSRD, uncertainty about a signal made up of unrelated tones was ameliorated by a cue that shared its individual frequencies, while uncertainty about a signal made up of harmonically related tones was ameliorated by a cue that shared its complex pitch. Finally, in support of the idea of separate and independent levels of representation for the two types of detection, randomly chosen harmonic complexes were detected still better if preceded by cues that matched them both in frequency and in complex pitch.

TABLE AI. Subject-by-subject planned comparisons between conditions (see text of the Appendix).

Conditions compared	Subject 1	Subject 2	Subject 3
C2>C1	2.0254	3.8216	2.9410
C3>C1	3.8990	5.5966	5.1237
C4>C2	3.9401	1.4808	2.5797
C5>C4	2.5277	5.1497	5.4384

ACKNOWLEDGMENTS

This work was supported by a grant from the USPHS, Institute on Deafness and Other Communication Disorders, No. NIDCD 00087. We wish to thank the many friends and colleagues who offered invaluable comments on earlier versions of the manuscript. These include Poppy Crum, Erick Gallun, Eric Jensen, Gerald Kidd, Christine Mason, Dennis McFadden, and Miriam Valenzuela.

APPENDIX

The LSRD model suggests three planned comparisons based on the argument that complex signals would be more detectable primitives, C2>C1, and the efficacy of the two types of cues, C3>C1 and C4>C2. Separate paired comparisons were made for the three subjects based on the minimum number of trials (350) collected for each condition. That is, for the mth subject, a z-score of the difference between conditions Y and X was calculated as

$$z(m,y,x)=\frac{P(C)_{(m,y)}-P(C)_{(m,x)}}{\sqrt{[P(C)_{(m,y)}][1-P(C)_{(m,y)}]+[P(C)_{(m,x)}][1-P(C)_{(m,x)}]}}. \tag{A1}$$

Results of these analyses are shown in Table AI. Use of a one-tailed test for the planned comparisons based on prior theory (Keppel and Zedeck, 1989) showed that eight of the nine $P(C)$-differences were significant ($p<0.05$), with the worst case being for subject 2 in C4>C2, where the difference was marginally insignificant ($p<0.07$). Indeed, seven of the nine cases were significant at $p<0.01$, now exempting subject 1 in C2>C1 ($p<0.02$).

For condition 5, where both types of cues were present, the argument that bi-dimensional cues aided independent processing of stimuli at the two putative LSRDs is strongly supported by results in Table II, which show that the amount of transmitted information, $(d')^2$ (Green and Swets, 1966), in condition 5 actually exceeded the sum of the $(d')^2$ values from the two singly cued conditions. Signals were the same in conditions 2, 4, and 5, each offering the possibility of detection on the basis of either frequency or complex pitch. The ability to use pitch cues with these signals is clearly shown by the comparison C4>C2. If frequency cueing provided additional improvement by ameliorating uncertainty about the individual frequencies, one would expect still higher performance in condition 5. The comparisons of C5>C4 shown in the fourth row of Table AI were significant for all subjects ($p<0.01$).

[1]Hafter *et al.* (1992) reported a case in which detections of randomly chosen, pure tone signals of frequency f_1 were tested alone and when preceded either by a single-tone cue of the same frequency or by sets cues made up

of five harmonics $2f_1 \rightarrow 6f_1$ or $3f_1 \rightarrow 7f_1$ of the signal. The proportion of correctly identified signals in a two-alternative, forced choice task was 0.68 with no cueing, 0.92 with a same-frequency cue, and an in-between value of 0.835 with the complex cue.

[2]As noted in Sec. V C, conditions 1–4 were repeated using the same limited range of frequencies for randomly drawn frequencies harmonically related signals. Data from three subjects were collected using a two-down, one-up tracking procedure (16 reversals) to produce four to eight tracks per subject. Thresholds obtained in that way are not directly comparable to $P(C)$ values in the main experiment, but the same basic effect accrued. Relative to the case with purely random tones (condition 1), the average threshold for harmonically related tones (condition 2) was -0.66 dB; for random tones cued by the same frequencies (condition 3) it was -2.28 dB; for harmonically related signals cued by the other three harmonics (condition 4) it was 4.05 dB.

Dai, H. P., Scharf, B., and Buus, S. (**1991**). ''Effective attenuation of signals in noise under focused attention,'' J. Acoust. Soc. Am. **89**, 2837–2842.

Darwin, C. (**1995**). Review of a manuscript submitted to J. Acoust. Soc. Am.

de Boer, E. (**1976**). ''On the 'residue' and auditory pitch perception,'' in *Handbook of Sensory Physiology*, edited by W. D. Keidel and W. D. Neff, Vol. V, Part 3, pp. 479–583.

Green, D. M. (**1961**). ''Detection of auditory sinusoids of uncertain frequency,'' J. Acoust. Soc. Am. **33**, 904–911.

Green, D. M., and Swets, J. A. (**1966**). *Signal Detection Theory and Psychophysics* (Krieger, Huntington, New York).

Green, D. M., McKey, M. J., and Licklider, J. C. R. (**1959**). ''Detection of a pulsed sinusoid in noise as a function of frequency,'' J. Acoust. Soc. Am. **31**, 1446–1452.

Hafter, E. R., and Kaplan, R. (**1976**). NASA Ames Report, ''A role for attention in auditory detection under conditions of frequency uncertainty.''

Hafter, E. R., and Schlauch, R. S. (**1989**). ''Factors in detection under uncertainty,'' J. Acoust. Soc. Am. **86**, S112.

Hafter, E. R., Schlauch, R. S., and Tang, J. (**1993**). ''Attending to auditory filters that were not stimulated directly,'' J. Acoust. Soc. Am. **94**, 743–747.

Hafter, E. R., Saberi, K., Schlauch, R. S., Plamondon, L., and Tang, J. (**1992**). ''Use of Stimulus Pre-cues to Study Processing at Various Levels of Auditory and Musical Processing,'' Second International Conference on Music Perception and Cognition, Univ. of California, Los Angeles, CA, 22–26 Februrary.

Houtgast, T. (**1976**). ''Subharmonic pitches of a pure tone at low S/N ratio,'' J. Acoust. Soc. Am. **60**, 405–409.

Houtsma, A. J. M. (**1984**). ''Pitch Salience of Complex Sounds,'' Music Percept. **1**, 296–307.

Houtsma, A. J. M. (**1995**). ''Pitch Perception,'' in *Handbook of Perception and Cognition*, edited by B. C. J. Moore (Academic, San Diego).

Houtsma, A. J. M., Rossing, T. D., and Wagenaars, W. M. (**1987**). ''Auditory Demonstrations, Demonstration 23,'' Institute for Perception Research, Eindhoven, The Netherlands.

Howard, Jr., J. H., O'Toole, A. J., Parasuraman, R., and Bennett, K. B. (**1984**). ''Pattern-directed attention in uncertain-frequency detection,'' Percept. Psychophys. **35**, 256–264.

Howard, J. H., O'Toole, A. J., and Rice, S. E. (**1986**). ''The role of frequency versus informational cues in uncertain frequency detection,'' J. Acoust. Soc. Am. **79**, 788–791.

Johnson, D. M., and Hafter, E. R. (**1980**). ''Uncertain-frequency detection: Cuing and condition of observation,'' Percept. Psychophys. **28**, 143–149.

Keppel, G., and Zedeck, S. (**1989**). *Data Analysis for Research Designs* (Freeman, New York).

Langner, G., Sams, M., Heil, P., and Schulze, H. (**1997**). ''Frequency and periodicity are represented in orthogonal maps in the human auditory cortex: evidence from magnetoencephalography,'' J. Comp. Physiol., A **181**, 665–676.

Licklider, J. C. R. (**1956**). ''Auditory frequency analysis,'' in *Information Theory*, edited by C. Cherry (Academic, New York).

McFadden, D. (**1988**). ''Failure of a missing-fundamental complex to interact with masked and unmasked pure tones at its fundamental frequency,'' Hear. Res. **32**, 23–40.

Moore, B. C. J. (**1997**). *Psychology of Hearing* (Academic, London).

Moore, B. C. J., Glasberg, B. R., and Peters, R. W. (**1986**). ''Thresholds for hearing mistuned partials as separate tones in harmonic complexes,'' J. Acoust. Soc. Am. **81**, 1873–1880.

Patterson, R., and Moore, B. C. J. (**1986**). ''Auditory filters and excitation patterns as representations of frequency resolution,'' in *Frequency Selectivity in Hearing*, edited by B. C. J. Moore (Academic, New York).

Pickles, J. O. (**1988**). *An Introduction to the Physiology of Hearing* (Academic, London).

Plamondon, L., and Hafter, E. (**1990**). ''Selective attention in absolute pitch listeners,'' J. Acoust. Soc. Am. Suppl. 1 **88**, S49.

Plomp, R. (**1964**). ''The ear as a frequency analyzer,'' J. Acoust. Soc. Am. **36**, 1628–1636.

Plomp, R., and Mimpen, A. M. (**1968**). ''The ear as a frequency analyzer II,'' J. Acoust. Soc. Am. **43**, 764–767.

Scharf, B. (**1970**). ''Critical Bands,'' in *Foundations of Modern Auditory Theory*, edited by J. V. Tobias (Academic, New York), Vol 1, pp. 157–202.

Scharf, B., Chays, A., and Magnan, J. (**1997**). ''The role of the olivocochlear bundle in hearing: Sixteen case studies,'' Hear. Res. **103**, 101–122.

Scharf, B., Magnan, J., Collet, L., Ulmer, E., and Chays, A. (**1994**). ''On the role of the olivocochlear bundle in hearing: A case study,'' Hear. Res. **75**, 11–26.

Schlauch, R. S., and Hafter, E. R. (**1991**). ''Listening bandwidths and frequency uncertainty in puretone signal detection,'' J. Acoust. Soc. Am. **90**, 1332–1339.

Schouten, J. F. (**1940**). ''The residue and the mechanism of hearing,'' Proc. K. Ned. Akad. Wet. **43**, 991–999.

Schreiner, C. E., and Langner, G. (**1988**). ''Periodicity coding in the inferior colliculus of the cat. II Topographical organization,'' J. Neurophysiol. **60**, 1823–1840.

Swets, J. A. (**1964**). *Signal Detection and Recognition by Human Observers* (Wiley, New York).

Swets, J. A., and Sewall, S. T. (**1961**). ''Stimulus vs. Response uncertainty in recognition,'' J. Acoust. Soc. Am. **33**, 1586–1592.

Terhardt, E. (**1974**). ''Pitch, consonance and harmony,'' J. Acoust. Soc. Am. **55**, 1061–1069.

Terhardt, E., Stoll, G., and Seewann, M. (**1982**). ''Pitch of complex according to virtual-pitch theory: Tests, examples, and predictions,'' J. Acoust. Soc. Am. **71**, 671–678.

Watson, C. S., and Kelly, W. J. (**1981**). ''The role of stimulus uncertainty in the discrimination of auditory patterns,'' in *Auditory and Visual Pattern Recognition*, edited by D. H. Getty and J. H. Howard (Erlbaum, Hillsdale, NJ).