

JULY 09 2025

Subject instructions for improved characterization of auditory representations using reverse correlation

Gidey W. Gezae; Nelson V. Barnett; Benjamin Parrell ; Divya A. Chari; Adam C. Lammert



JASA Express Lett. 5, 074402 (2025)

<https://doi.org/10.1121/10.0037102>



Articles You May Be Interested In

A microscopic investigation of the effect of random envelope fluctuations on phoneme-in-noise perception

J. Acoust. Soc. Am. (February 2024)

High-resolution mapping and backscatter studies in support of the Ripples Program

J. Acoust. Soc. Am. (November 2006)

Effect of early dialectal exposure on adult perception of phonemic vowel length

J. Acoust. Soc. Am. (September 2017)




Advance your science and career as a member of the

Acoustical Society of America

LEARN MORE

Subject instructions for improved characterization of auditory representations using reverse correlation

Gidey W. Gezae,¹ Nelson V. Barnett,² Benjamin Parrell,^{3,4}  Divya A. Chari,^{5,6} and Adam C. Lammert^{2,7}

¹Human and Health Development, Pennsylvania State University, State College, Pennsylvania 16802, USA

²MGH Institute of Health Professions, Charlestown, Massachusetts 02129, USA

³Communication Sciences and Disorders, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA

⁴Waisman Center, University of Wisconsin-Madison, Madison, Wisconsin 53705, USA

⁵University of Massachusetts Chan Medical School, Worcester, Massachusetts 01655, USA

⁶Massachusetts Eye and Ear Infirmary, Boston, Massachusetts 02114, USA

⁷Mathematics and Computer Science, College of the Holy Cross, Worcester, Massachusetts 01610, USA

gig5425@psu.edu, nbarnett@mghihp.edu, bparrell@wisc.edu, divya.chari@umassmemorial.org, alammert@holycross.edu

Abstract: Reverse correlation (RC) is an established method for reconstructing auditory representations and has recently emerged as a tool for characterizing the sounds experienced by tinnitus patients. To further optimize RC for auditory research, the present work investigated the influence of subject instructions on characterization quality of tinnitus-like sounds. A validation study was conducted in which 18 normal-hearing subjects were randomly assigned 1 of 3 candidate instruction sets, each inspired by the RC literature. Results show a significant effect of instruction set on characterization quality and reveal that instructing subjects to detect a hidden signal in the RC stimuli resulted in the best reconstruction. © 2025 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC) license (<https://creativecommons.org/licenses/by-nc/4.0/>).

[Editor: Qian-Jie Fu]

<https://doi.org/10.1121/10.0037102>

Received: 21 March 2025 **Accepted:** 18 June 2025 **Published Online:** 9 July 2025

1. Introduction

Internal auditory representations play an important role in mediating our perceptual experience of sound by providing the critical link between the sensory signals available at the cochlea and our perceptual experience of relatively abstract auditory objects. Auditory representations are used as the basis for interpreting sensory signals in accordance with various auditory percepts, ranging from low-level spatial and temporal features (e.g., pitches, timbres, and formants) to higher-level cognitive constructs (e.g., phonemes, speaker identities, and environmental sound sources). Auditory representations that are relatively localized in the auditory pathway (e.g., auditory receptive fields in primary auditory cortex) have seen substantial exposition, whereas auditory representations that are more distributed and/or abstract remain poorly understood by comparison, perhaps because of a lack of approaches for measuring them directly.

One notable exception to the scarcity of approaches for directly measuring higher-level auditory representations is the method of reverse correlation (RC). RC asks subjects to render subjective judgments over random and/or ambiguous stimuli that have no direct relation to the representation of interest, from which unbiased and arbitrarily accurate estimates of those representations can be rendered (Murray, 2011).¹ Although RC has been applied to the study of different modalities, most notably vision, some of the earliest applications of RC were in auditory perception. Ahumada and Lovell (1971) used RC to estimate perceptual representations of simple tones (Ahumada, 2002). This pioneering work was later given a theoretical grounding based on a linear signal detection model by Richards and Zhu (1994). Additionally, auditory applications of RC have surpassed the perception of tones to also explore the quality of more complex sounds, primarily in the context of speech perception. Brimijoin *et al.* (2013) used RC to perform estimation of English vowels in 60 frequency bands spaced over the full audible range of 0.1–22 kHz. Varnet and colleagues have applied RC in several studies aimed at uncovering the fine acoustic cues—e.g., due to coarticulation—that promote robustness in speech recognition (Varnet *et al.*, 2013a,b; Varnet *et al.*, 2015; Varnet *et al.*, 2016). Ponsot *et al.* (2018b) has studied prosodic pitch variation associated with acoustic correlates of social judgments of, for example, dominance, trustworthiness, and smiling. In addition, a closely related methodology, known as *profile analysis*, has resulted in a considerable body of work aimed at probing the internal representations of complex sounds (Green, 1988; Berg and Green, 1990; Lutfi, 1995; Richards and Buss, 1996; Lentz, 2007; Leibold *et al.*, 2009).

RC has also emerged recently as a promising method for characterizing tinnitus sounds. Tinnitus, a prevalent auditory condition characterized by the perception of sound without an external source, affects millions of people worldwide (Tunkel *et al.*, 2014). Accurate characterization of tinnitus sounds across individuals is viewed as critical to progress in tinnitus research (Henry *et al.*, 2013; Korth *et al.*, 2021; Landgrebe *et al.*, 2017; Lentz and He, 2020; Mohan *et al.*, 2022; Ukaegbe *et al.*, 2017), diagnosis (Chari and Limb, 2018; Heijneman *et al.*, 2013; Norena *et al.*, 2002; de Ridder *et al.*, 2021; Roberts

et al., 2008; Tyler *et al.*, 1992), and treatment (Davis *et al.*, 2007; Langguth *et al.*, 2007; Nickel *et al.*, 2005; Okamoto *et al.*, 2010; Schaette *et al.*, 2010; Stein *et al.*, 2015; Tass *et al.*, 2012). Recent evidence from applying RC to tinnitus suggests that the method can accurately characterize complex tinnitus sounds (Hoyland *et al.*, 2023; Barnett *et al.*, 2024). Moreover, a fundamental promise of RC is that when presented stimuli that vary within the space of possible representations, estimation of arbitrary representations is possible. RC is, thus, especially appropriate for characterizing tinnitus given its wide heterogeneity in sound character (e.g., roaring, buzzing, etc.), which is in contrast to existing methods—for example, pitch matching—that are best suited only for patients who experience tonal sounds (e.g., ringing; Barnett *et al.*, 2024).

However, it is not clear from these early applications whether RC protocols have been optimized in the auditory domain, much less in the domain of tinnitus. RC protocols described in the research literature vary in several ways, including in the nature of the stimuli presented to subjects (e.g., Bernoulli-distributed; Gosselin and Schyns, 2003) or Gaussian-distributed (Ringach and Shapley, 2004), as well as pre-whitened (Compton *et al.*, 2022) or containing embedded targets (Dotsch and Todorov, 2012; Brinkman *et al.*, 2017), the method of estimation (e.g., linear (Gosselin and Schyns, 2003), sparse-constrained (Roop *et al.*, 2024), or maximum *a posteriori* (Mineault *et al.*, 2009), and instructions given to subjects, the lattermost of which is the focus of the present study. Whereas the subjective judgments that RC participants make are, ideally, well-informed by their internal perceptual experience, there are nevertheless many ways that participants might be instructed to listen, interpret, and, ultimately, respond to stimuli that are vague and ambiguous. The question of optimal subject instructions for RC seems especially pertinent in domains where the objects of interest are themselves often described as noisy—e.g., tinnitus sounds or fricative sounds—which may confuse participants from the outset as to whether their task is one of finding *overt* matches to the percept of interest from among the stimuli or to identify aspects of similarity to the percept that may be *hidden* within the stimuli. Up to 50% of tinnitus patients may experience tinnitus sounds that are considered nontonal and may have noise-like qualities (Henry *et al.*, 2013; Landgrebe *et al.*, 2012; Lentz and He, 2020; Ukaegbe *et al.*, 2017).

Overt matching asks subjects to identify noisy stimuli that simply match, to sufficient quality, the percept of interest. Perhaps because of the fact that most applications of RC are not interested in percepts that are noisy in and of themselves, this approach is much less widely represented in the literature. A notable exception to this generalization, however, is represented by the overlapping literature on *pareidolia* (Rekow *et al.*, 2022; Salge *et al.*, 2021), where the explicit aim is to study situations in which illusory percepts emerge from—as opposed to appearing “inside” or “underneath”—vague stimuli. Still, much more widely represented in the literature is the *hidden* approach to subject instructions, which typically involves asking a subject to identify the percept of interest “within the noise.” Even within the hidden approach, there exist two broad subclasses of participant instructions in the literature: those based on *detection*, in which subjects identify stimuli containing the percept of interest from among those that contain no percept (Ahumada and Lovell, 1971; Gosselin and Schyns, 2003; Brimijoin *et al.*, 2013), and those based on *discrimination*, in which subjects identify stimuli that contain the percept of interest from among those that contain other percepts (Brinkman *et al.*, 2017; Ponsot *et al.*, 2018a; Varnet and Lorenzi, 2022).

There is no reason to believe *a priori* that any of these approaches to RC subject instructions—overt, hidden/detection, or hidden/discrimination—will result in obviously more accurate characterization of auditory percepts, yet, preliminary evidence from our research group, uncovered in the preparation of Hoyland *et al.* (2023) and in [supplementary material A](#), indicate that there may be reason to expect that subject instructions will have a measurable and, perhaps, substantial impact on the quality of the results. In any case concerning RC estimation quality, no direct comparison of subject instructions and their effect on estimation quality has been reported in the literature.

By investigating the impact of various instruction sets, we aim to identify the most effective instructions, which enhance the validity and efficacy of RC for the reconstruction of auditory representations, in comprehending how sound characteristics contribute to tinnitus perception. Our approach involves the careful design of three distinct instruction sets in a validation study similar to that reported by Hoyland *et al.* (2023). The stimuli used were designed specifically to match potential qualities of tinnitus sounds, although the task and instructions are broadly applicable across the auditory domain. The results indicate that participant instructions have a large impact on the accuracy of RC reconstruction accuracy, with instructions consistent with a hidden/detection task providing the most accurate reconstructions.

2. Methods and materials

Following the procedure described by Hoyland *et al.* (2023), we asked participants without tinnitus to complete an augmented RC experiment in which they listened to a target tinnitus-like sound and a random-noise stimulus (described below) on every trial. Target sounds were subsequently used as the basis for validating the estimation quality of the RC-estimated tinnitus sounds by direct comparison.

2.1 Experimental protocol

Subjects listened to A-X trials, which contained a target sound (A) followed by a stimulus (X). X was randomly generated for each trial, whereas A remained the same across all trials (Fig. 1). Trials were presented in blocks of 100, with voluntary breaks offered between blocks. Subjects completed five blocks per target sound ($p = 500$ total trials per subject). The total duration of the experiment was approximately 30 min.

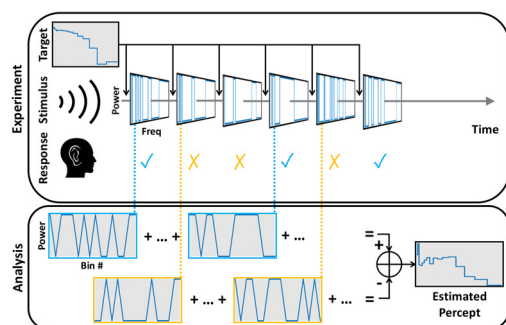


Fig. 1. Illustration of the experimental protocol. Subjects listen to a sequence of randomly generated auditory stimuli, each preceded by a target tinnitus-like sound. Subjects make a subjective judgment over the stimulus relative to the target sound, responding “yes” or “no,” in light of the specific instruction set given to them. The resulting collection of all stimulus-response pairs is used to form an estimate of the target.

Subjects listened over Sennheiser HD 650, open-back, circumaural earphones (Sennheiser, Wedemark, Germany) at a self-determined comfortable level in a quiet auditory environment. Stimuli were presented from a MacBook Pro (Apple, Cupertino, CA) at a sampling rate of 44 100 Hz with 24-bit encoding. Presentation level was not recorded but was estimated to be in the range of 60–80 dB sound pressure level (SPL).

2.2 Participants

Based on a pilot study showing that key instruction sets resulted in an absolute difference in mean reconstruction accuracy of 0.412 and Tukey range statistic of 8.293 (see [supplementary material A](#)), a total of 18 participants were recruited and assigned randomly to 1 of 3 groups, with 6 participants in each group. However, three participants over age 40 years old were assigned (randomly) to different groups to balance the distribution of ages in these three groups. Each group received a distinct instruction set as described below. Group 1 had a mean age of 25.0 years old with a standard deviation of 13.4 years (range 18–55 years). Group 2 had a mean age of 27.2 years old with a standard deviation 18.2 years (range 18–68 years). Group 3 had a mean age of 22.8 years old with a standard deviation of 8.6 years (range 18–42 years). All participants reported no tinnitus or hearing loss.

2.3 Target stimuli

The target sound was drawn from online examples of tinnitus-like sounds, maintained by the [American Tinnitus Association \(2022\)](#). The examples selected for the present study corresponded to the one described as “roaring.” The target sound was selected for its spectra complexity and because of its use in prior RC validation studies ([Hoyland et al., 2023](#); [Barnett et al., 2024](#)). The roaring sound was selected also to further mitigate concerns related to high-frequency hearing loss in participants, as its power decreases with frequency, such that power levels above 10 kHz are at least 30 dB down relative to the 0–2 kHz band, which contains the highest power level. The sound was downloaded and truncated to 500 ms in duration, down from 1000 ms. The purpose of this target sound was primarily to serve as the representative tinnitus-like target for the purposes of estimation quality assessment for all participants, as detailed below.

2.4 Random-noise stimuli

Random-noise stimuli were generated in the frequency domain as random frequency spectra. The frequency domain from 100 to 13 000 Hz was divided into a total of 16 contiguous bins using logarithmically spaced dividing frequencies following the Mel scale. Frequencies within a given bin were all assigned the same power level for each stimulus. To generate a given stimulus, a number of bins between three and nine were selected, with uniform probability, to be assigned a power level of 0 dB, whereas the remaining bins were assigned -100 dB (following [Hoyland et al., 2023](#); [Barnett et al., 2024](#)). Each frequency was then assigned random phase using uniformly distributed random numbers between zero and 2π , and inverse Fourier transforms were subsequently used to compute audio waveforms of 500 ms duration corresponding to each frequency spectrum.

Stimuli composed of binary random variables such as these are well-represented in the RC literature (e.g., [Gosselin and Schyns, 2003](#); [Barnett et al., 2024](#)) and can be contrasted with stimuli composed of less quantized or even continuous values (e.g., [Brimijoin et al., 2013](#); [Abbey and Eckstein, 2006](#)). Note that this approach will result in the estimate β being composed of values that are also quantized, albeit much less than the values in the stimuli. Specifically, as a result of the fact that $y_j \in \{-1, +1\}$, the i th value of β , $\beta_i = 500^{-1}(\pm X_{1,i} \pm X_{2,i} \pm \dots \pm X_{500,i})$, where the value in row j and column i of X , $X_{j,i} \in \{-100, 0\}$. In the present case, β_i will be some number between -100 and 0 , rounded to the nearest increment of $100/500 = 0.2$.

2.5 Instruction sets

In the first instruction set (Q1, hidden/detection), participants were tasked with listening to each random-noise stimulus and responding based on the perceived presence or absence of the target roaring sound hidden in that stimulus. Subjects were informed that a target sound would be hidden in approximately 50% of the random-noise stimuli. This instruction set was assigned to participant group 1. A second instruction set (Q2, hidden/discrimination) asked subjects to listen to

each random-noise stimulus and provide a response based on whether the sound hidden in that stimulus was a good match to the target sound. Subjects were told that 100% of stimuli would contain a hidden sound, including a random assortment of roaring, buzzing, and ringing sounds. Q2 was assigned to participant group 2. Finally, the third instruction set (Q3, overt) asked participants to listen to each random-noise stimulus and provide a response based on whether the stimulus was a good match for the target sound. Q3 was assigned to participant group 3. Note that the target stimulus was present in every trial and followed in every trial by a random-noise stimulus. Further, note that none of the random-noise stimuli employed in the experiment contained any embedded sounds, nor were they intended to sound like the target or generated with any regard to the target sound whatsoever.

2.6 Tinnitus sound estimation

Estimates of individual participants' percept of the target sound were obtained using the standard RC reconstruction procedure. That is, for each subject, given a 500×16 stimulus matrix X , where each row of X is a bin-power representation of one random-noise stimulus, and a 500×1 response vector y , where entries in y have corresponding values of either -1 ("no" response) or $+1$ ("yes" response), estimates of the subject's percept of the target sound may be calculated as $\beta = 500^{-1}X^T y$. It has been widely depicted in the literature that this estimation formula is a restricted form of the Normal equation, which is the least squares solution to linear regression (Gosselin and Schyns, 2003). In this specific case, the formula minimizes the sum of squared errors between the subject's responses and the responses estimated using $y_{\text{pred}} = X^T \beta$.

2.7 Estimation quality assessment

The power spectral density of the target sound was determined using Welch's method (Welch, 1967) and subsequently transformed into a spectral representation, allowing for direct comparison with the RC estimates. Frequency bins of the same number and spacing as those used for random-noise stimulus creation were defined, and the mean power at frequencies within each bin was calculated, resulting in a bin-power representation, θ , directly comparable to the bin-power representation of the random-noise stimuli. Quality of the estimates was quantified by correlating β with θ using Pearson's r . P -values for each correlation were also calculated and then compared against the significance level of $\alpha = 0.05$, corrected using the Holm-Bonferroni method. Variations in estimation quality between subjects given different instruction sets were tested using a one-way analysis of variance (ANOVA) to test for differences in the mean values of Fisher-transformed (Fisher, 1915) Pearson's r (see Table 1) among the subject groups receiving different instructions sets (i.e., Q1, Q2, or Q3). Further analysis using Tukey's HSD (honestly significant difference) test was performed to compare the mean difference of estimation quality among the three instruction set groups.

3. Results

RC estimates of the participant's percept of the target sound are shown in Fig. 2, overlaid with a bin-power representation of the target sound spectrum. Qualitative differences in the correspondence between the estimates and target sound can be observed across instruction set groups as well as within those groups. The results of correlation analysis between each subject's estimate and the target sound—including Pearson's r and associated p -values—are listed in Table 1. The only three subjects who showed statistically significant correlations were subjects S1, S2, and S3, all of whom received instruction set Q1.

Participant instructions had a significant effect on the reconstruction accuracy, measured by the strength of the correlation between participants' percepts and the true target spectra [$F(2) = 6.26$, $p = 0.011$; Fig. 3]. *Post hoc* testing showed that Q1 (hidden/detection) led to significantly higher reconstruction accuracy than Q3 [overt; mean improvement = 0.147 , $q_s(5) = 4.856$, $p = 0.010$]. There was no statistically significant difference between Q1 and Q2 [$q_s(5) = 3.470$, $p = 0.066$], or Q2 and Q3 [$q_s(5) = 1.386$, $p = 0.600$]. In a follow-up analysis, it was found that the variance in reconstruction quality was significantly higher with Q2 (hidden/discrimination) than Q1 (hidden/detection) using a two-sample F -test for difference in variance between Q1 and Q2 such that $F(5,5) = 0.1281$, $p = 0.042$.

Table 1. Pearson's r and associated p -values for the correlation between the RC estimate and the target sound spectrum for each participant. Bold p -values indicate significant correlations at the Holm-Bonferroni correlated threshold value [$\alpha = 0.05$, degree of freedom (DoF) = 15].

Instruction set		Correlation between target and reconstruction					
Q1 (hidden/detection)	Subject identification	S1	S2	S3	S4	S5	S6
	Pearson's r	0.819	0.786	0.706	0.535	0.495	0.344
	p -value	<0.001	0.001	0.003	0.040	0.061	0.210
Q2 (hidden/discrimination)	Subject identification	S7	S8	S9	S10	S11	S12
	Pearson's r	0.635	0.615	0.327	0.292	−0.403	−0.588
	p -value	0.011	0.015	0.234	0.291	0.136	0.021
Q3 (overt)	Subject identification	S13	S14	S15	S16	S17	S18
	Pearson's r	0.368	0.137	−0.110	−0.133	−0.280	−0.397
	p -value	0.177	0.626	0.696	0.637	0.312	0.143

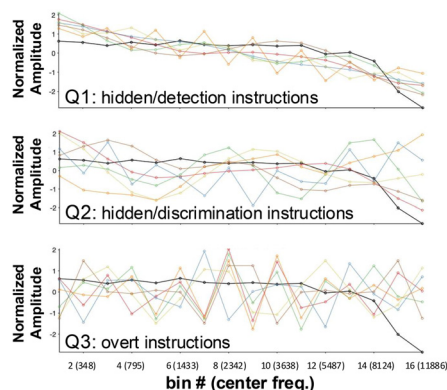


Fig. 2. RC estimates of the target roaring sound from each of three instruction sets, Q1 (hidden/detection), Q2 (hidden/discrimination) and Q3 (overt). Subjects-specific estimates are indicated by the color traces. Black traces indicate the bin-power representation of the target sound. Power levels for each estimate were standardized prior to being plotted.

4. Discussion

We tested the effect of participant instructions on the accuracy of reconstructing auditory percepts in a signal detection task using RC. Our results show that instructions make a significant difference in the accuracy of these reconstructions, where the highest accuracies stem from instructing subjects to detect the presence of the target sound hidden in a random-noise stimulus, which participants were told may or may not be present (Q1, hidden/detection). Results indicate that these instructions provided a significant improvement, in particular, over instructions that asked subjects to match the random-noise stimulus and target sound (Q3, overt). Qualitatively, the hidden/detection instructions (Q1) also resulted in better accuracy than instructions asking participants to expect different embedded sounds on every trial and subsequently respond when this embedded sound matched the target (Q2, hidden/discrimination), although this difference was not statistically significant.

The reason for this effect is not immediately clear and may merit future study. Notwithstanding the empirical evidence presented here, the model of subject responses in RC does not readily admit any theoretical reason for a difference in subject responses patterns, much less reconstruction quality driven by such differences. Still, anecdotal evidence from our subjects suggests that hidden-type instructions narrow listeners' focus to the presence of any component features of a given stimulus that rise "above the noise," rather than keeping attention on the overall impression of the stimulus. In assessing the overall impression of stimuli, subjects may find that relatively few constitute "good matches," resulting in less reliable responses and a more difficult task overall.

Notably, variance in the accuracy of the reconstructions of the target sound was significantly higher with Q2 (hidden/discrimination) than Q1 (hidden/detection). The reason for this is unclear but may be related to anecdotal reports from subjects that the Q2 instructions were confusing or difficult to understand. This effect may also corroborate the finding of [Abbey and Eckstein \(2006\)](#), which showed wide differences in efficiency in discrimination vs detection tasks. Furthermore, it is important to acknowledge that Q2, in addition to having distinct instructions, represents a different psychophysical task. Q2 represents a discrimination task, in which subjects were asked to compare the target sound to alternatives. In such a task, the RC-derived estimate would, in general, be expected to reflect not an absolute representation of the target sound but a relative one, which would complicate comparisons between the estimate and the target sound. However, given enough alternatives in a discrimination task, the estimate should converge to the absolute representation (see [supplementary material C](#)), which was the intention in the present study. No examples of alternative sounds were ever played, and the instructions were kept broad with respect to the nature of the alternatives to encourage subjects to discriminate the target sound from a vague, general idea of some other tinnitus sound. That is, the use of a discrimination task in Q2 is a means to an end, where the goal of the task is still to reconstruct the absolute representation, consistent with a detection task and the RC model of subject responses. Still, the possibility should be acknowledged of some subjects misinterpreting the task as comparing the target sound to a small number of specific alternatives, which is an outcome that could potentially explain some of the increased variability observed in

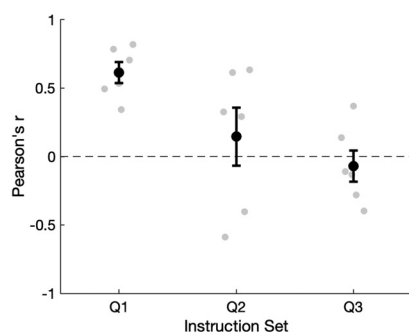


Fig. 3. Scatterplot shows the distribution of Pearson's r —a measure of RC estimation quality—for each instruction set as gray circles. Black circles indicate the mean value, and black lines depict the upper and lower edges of the 95% confidence interval.

Q2 condition and, may, by extension, be partially responsible for the absence of a significant difference between this instruction set and the others.

Evaluating RC estimation quality by means of correlating the spectra of the RC-derived estimates and the target sounds is appropriate despite the effects of spectral transformations of the stimuli and target sounds along the auditory pathway. Correlation values will not be changed for any reasonable transformation of the sound spectra in the context of these experiments, because the stimuli and target sounds are subject to those same transformations. A formal statement of this fact is supplied in [supplementary material B](#). Note that quality assessment via correlation would be similarly unchanged as a result of spectra transformations stemming from hearing-related factors that might influence the RC estimates, including those caused by hearing loss (especially age-related), presentation level of the target and stimulus sounds, and the frequency response of equipment used to present sound.

It should be commented, however, that the formal statement in [supplementary material B](#) assumes that all frequencies remain audible to some degree and, therefore, does not account for situations in which some frequencies become entirely inaudible. Although this does not affect the relative distribution of correlation values in a way that would influence our primary result, the expected effect of inaudible frequencies given the model of subject responses in RC would be to lower the absolute magnitude of correlation between the target sound and its reconstruction because any variation across stimuli at the inaudible frequencies would be unaccounted for when calculating the reconstruction. This effect might partially explain why the observed correlations between the target and reconstruction are less than ideal and why relatively few of those correlations themselves were determined to be statistically significant. One might expect this effect to be especially pronounced in the context of targets containing abundant high-frequency power as a result of the prevalence of high-frequency hearing loss. However, it is notable that [Hoyland et al. \(2023\)](#) showed comparable results using a “buzzing” target, containing primarily high-frequency power, as compared to the roaring target used here, which does not.

Although not the primary purpose of the study, it is also notable that these results represent the highest-resolution reconstructions of tinnitus-like sounds using RC reported in the literature, where the frequency resolution (i.e., the number of frequency bins in reconstructions) is double that reported by [Hoyland et al. \(2023\)](#), a study that also made use of Q1-type instructions, which are shown in this work to be the preferred set. Moreover, these results are among the highest-reported reconstruction accuracies of tinnitus-like sounds, comparable in estimation quality to lower-resolution reconstructions reported by [Hoyland et al. \(2023\)](#). Note that maintaining accuracy along with increased resolution is not necessarily expected because of the widely observed fact that for a given number of trials in RC experiments, an increase in the resolution (i.e., dimensionality) of the stimuli and reconstructions will result in a reduction in accuracy (e.g., [Roop et al., 2024](#)). Therefore, the comparable estimation quality observed in the present study is likely attributable to the 150% increase in the number of trials collected relative to those of [Hoyland et al. \(2023\)](#). Increasing the number of trials is a standard approach to accommodate the accuracy-resolution trade-off.

The present study provides evidence for a preferred approach to instructing subjects engaged in auditory applications of RC, with regard to producing high quality reconstructions of auditory percepts. It was found that asking subjects to detect the percept of interest embedded within the noisy stimuli—as distinguished from those stimuli where no such percept was embedded—provided estimates of the highest accuracy. This was demonstrated in a specific experimental context and may be somewhat limited in generality because of the use of a single tinnitus-like target sound, the focus on a single-stimulus paradigm (and not, for example, a forced-choice paradigm), and a small, if statistically justified, sample size. However, it seems probable that subject instructions should be carefully considered in applications of RC to other aspects of auditory perception—as in, e.g., the growing literature on RC in speech ([Brimijoin et al., 2013](#); [Varnet et al., 2016](#); [Mandel et al., 2016](#))—to follow the language of the hidden approach used here. By refining the use of RC, researchers may be able to gain valuable scientific insight into how auditory phenomena are linked to personalized, cognitive representations of those phenomena. Improvements to RC may also enable improved diagnostic and treatment approaches for a variety of auditory conditions by allowing more precise characterization of pathological representations (e.g., as in speech processing disorders) or more effectively uncovering representations that are entirely internal to the individual (e.g., as in tinnitus).

Supplementary Material

See [supplementary material](#) for (A) the details of our pilot data and power analysis, (B) a formal statement regarding the appropriateness of target-estimate correlation for quality assessment, and (C) a formal statement regarding convergence of discrimination tasks to detection tasks.

Acknowledgments

This work was supported by the University of Massachusetts Center for Clinical and Translational Science and the People's Research Grant from the American Tinnitus Association.

Author Declarations

Conflict of Interest

The authors have no conflicts to disclose.

Ethics Approval

Procedures were approved by the University of Massachusetts (UMass) Internal Review Board (IRB; Study identification No. H00024058). Informed consent was obtained from all participants.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

¹Although the pursuit of internal, auditory representations is a primary motivation for using RC, it must be acknowledged that whether internal representations are what RC actually estimates is a matter of debate in the literature. Specific methodological concerns exist as a result of, e.g., occasional negations of the estimate relative to the true representation (Neri, 2018), in addition to general concerns stemming from the relative simplicity of the model underlying RC, which does not take into account the transformations of perceptual pathways, and assumes strictly linear comparisons with the stimuli. From a statistical point of view, what RC actually estimates—as a form of regression—is the representations which, given the model, best explains the observed subject responses.

- Abbey, C. K., and Eckstein, M. P. (2006). "Classification images for detection, contrast discrimination, and identification tasks with a common ideal observer," *J. Vis.* **6**(4), 335–355.
- Ahumada, A., Jr., and Lovell, J. (1971). "Stimulus features in signal detection," *J. Acoust. Soc. Am.* **49**(6B), 1751–1756.
- Ahumada, A. J. (2002). "Classification image weights and internal noise level estimation," *J. Vis.* **2**(1), 121–131.
- American Tinnitus Association (2022). "Listen to sample tinnitus sounds," available at <https://www.ata.org/listen-sample-tinnitus-sounds> (Last viewed June 15, 2022).
- Barnett, N. V., Hoyland, A., Chari, D. A., Parrell, B., and Lammert, A. C. (2024). "Reverse correlation characterizes more complete tinnitus spectra in patients," *IEEE Open J. Eng. Med. Biol.* **5**, 589–592.
- Berg, B. G., and Green, D. M. (1990). "Spectral weights in profile listening," *J. Acoust. Soc. Am.* **88**(2), 758–766.
- Brimijoin, O. W., Akeroyd, M. A., Tilbury, E., and Porr, B. (2013). "The internal representation of vowel spectra investigated using behavioral response-triggered averaging," *J. Acoust. Soc. Am.* **133**(2), EL118–EL122.
- Brinkman, L., Todorov, A., and Dotsch, R. (2017). "Visualising mental representations: A primer on noise-based reverse correlation in social psychology," *Eur. Rev. Soc. Psychol.* **28**(1), 333–361.
- Chari, D. A., and Limb, C. J. (2018). "Tinnitus," *Med. Clinics* **102**(6), 1081–1093.
- Compton, A., Roop, B. W., Parrell, B., and Lammert, A. C. (2022). "Stimulus whitening improves the efficiency of reverse correlation," *Behav. Res. Methods* **55**(6), 3120–3128.
- Davis, P. B., Paki, B., and Hanley, P. J. (2007). "Neuromonics tinnitus treatment: Third clinical trial," *Ear Hear.* **28**(2), 242–259.
- de Ridder, D., Schlee, W., Vanneste, S. *et al.* (2021). "Tinnitus and tinnitus disorder: Theoretical and operational definitions (an international multidisciplinary proposal)," *Prog. Brain Res.* **260**, 1–25.
- Dotsch, R., and Todorov, A. (2012). "Reverse correlating social face perception," *Soc. Psychol. Pers. Sci.* **3**(5), 562–571.
- Fisher, R. A. (1915). "Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population," *Biometrika* **10**(4), 507–521.
- Gosselin, F., and Schyns, P. G. (2003). "Superstitious perceptions reveal properties of internal representations," *Psychol. Sci.* **14**(5), 505–509.
- Green, D. M. (1988). *Profile Analysis: Auditory Intensity Discrimination* (Oxford University Press, Oxford, UK), No. 13.
- Heijneman, K. M., De Kleine, E., Wiersinga-Post, E., and Van Dijk, P. (2013). "Can the tinnitus spectrum identify tinnitus subgroups?," *Noise Health* **15**(63), 101–106.
- Henry, J. A., Roberts, L. E., Ellingson, R. M., and Thielman, E. J. (2013). "Computer-automated tinnitus assessment: Noise-band matching, maskability, and residual inhibition," *J. Am. Acad. Audiol.* **24**(06), 486–504.
- Hoyland, A., Barnett, N. V., Roop, B. W., Alexandrou, D., Caplan, M., Mills, J., Parrell, B., Chari, D. A., and Lammert, A. C. (2023). "Reverse correlation uncovers more complete tinnitus spectra," *IEEE Open J. Eng. Med. Biol.* **4**, 116–118.
- Korth, D., Wollbrink, A., Lukas, C., Ivansic, D., Guntinas-Lichius, O., Salvati, V., Paraskevopoulos, E., Pantev, C., and Döbel, C. (2021). "Comparing pure tone and narrow band noise to measure tonal tinnitus pitch-match frequency," *Prog. Brain Res.* **262**, 115–137.
- Landgrebe, M., Azevedo, A., Baguley, D. *et al.* (2012). "Methodological aspects of clinical trials in tinnitus: A proposal for an international standard," *J. Psychosom. Res.* **73**(2), 112–121.
- Landgrebe, M., Hajak, G., Wolf, S. *et al.* (2017). "1-Hz rTMS in the treatment of tinnitus: A sham-controlled, randomized multicenter trial," *Brain Stimul.* **10**(6), 1112–1120.
- Langguth, B., Goodey, R., Azevedo, A. *et al.* (2007). "Consensus for tinnitus patient assessment and treatment outcome measurement: Tinnitus Research Initiative meeting, Regensburg, July 2006," *Prog. Brain Res.* **166**, 525–536.
- Leibold, L. J., Tan, H., and Jesteadt, W. (2009). "Spectral weights for sample discrimination as a function of overall level," *J. Acoust. Soc. Am.* **125**(1), 339–346.
- Lentz, J. J. (2007). "Variation in spectral-shape discrimination weighting functions at different stimulus levels and signal strengths," *J. Acoust. Soc. Am.* **122**(3), 1702–1712.
- Lentz, J. J., and He, Y. (2020). "Perceptual dimensions underlying tinnitus-like sounds," *J. Speech. Lang. Hear. Res.* **63**(10), 3560–3566.

- Lutfi, R. A. (1995). "Correlation coefficients and correlation ratios as estimates of observer weights in multiple-observation tasks," *J. Acoust. Soc. Am.* **97**(2), 1333–1334.
- Mandel, M. I., Yoho, S. E., and Healy, E. W. (2016). "Measuring time-frequency importance functions of speech with bubble noise," *J. Acoust. Soc. Am.* **140**(4), 2542–2553.
- Mineault, P. J., Barthelme, S., and Pack, C. C. (2009). "Improved classification images with sparse priors in a smooth basis," *J. Vis.* **9**(10), 17.
- Mohan, A., Leong, S. L., De Ridder, D., and Vanneste, S. (2022). "Symptom dimensions to address heterogeneity in tinnitus," *Neurosci. Biobehav. Rev.* **134**, 104542.
- Murray, R. F. (2011). "Classification images: A review," *J. Vis.* **11**(5), 2.
- Neri, P. (2018). "Classification images as descriptive statistics," *J. Math. Psychol.* **82**, 26–37.
- Nickel, A. K., Hillecke, T., Argstatter, H., and Bolay, H. V. (2005). "Outcome research in music therapy: A step on the long road to an evidence-based treatment," *Ann. N.Y. Acad. Sci.* **1060**(1), 283–293.
- Norena, A., Micheyl, C., Chéry-Croze, S., and Collet, L. (2002). "Psychoacoustic characterization of the tinnitus spectrum: Implications for the underlying mechanisms of tinnitus," *Audiol. Neurotol.* **7**(6), 358–369.
- Okamoto, H., Stracke, H., Stoll, W., and Pantev, C. (2010). "Listening to tailor-made notched music reduces tinnitus loudness and tinnitus-related auditory cortex activity," *Proc. Natl. Acad. Sci. U.S.A.* **107**(3), 1207–1210.
- Ponsot, E., Arias, P., and Aucouturier, J. J. (2018b). "Uncovering mental representations of smiled speech using reverse correlation," *J. Acoust. Soc. Am.* **143**(1), EL19–EL24.
- Ponsot, E., Burred, J. J., Belin, P., and Aucouturier, J. J. (2018a). "Cracking the social code of speech prosody using reverse correlation," *Proc. Natl. Acad. Sci. U.S.A.* **115**(15), 3972–3977.
- Rekow, D., Baudouin, J. Y., Brochard, R., Rossion, B., and Leleu, A. (2022). "Rapid neural categorization of facelike objects predicts the perceptual awareness of a face (*face pareidolia*)," *Cognition* **222**, 105016.
- Richards, V. M., and Buss, E. (1996). "Frequency correlation functions for the detection of a tone added to modulated noise maskers," *J. Acoust. Soc. Am.* **99**(3), 1645–1652.
- Richards, V. M., and Zhu, S. (1994). "Relative estimates of combination weights, decision criteria, and internal noise based on correlation coefficients," *J. Acoust. Soc. Am.* **95**(1), 423–434.
- Ringach, D., and Shapley, R. (2004). "Reverse correlation in neurophysiology," *Cognit. Sci.* **28**(2), 147–166.
- Roberts, L. E., Moffat, G., Baumann, M., Ward, L. M., and Bosnyak, D. J. (2008). "Residual inhibition functions overlap tinnitus spectra and the region of auditory threshold shift," *J. Assoc. Res. Otolaryngol.* **9**(4), 417–435.
- Roop, B. W., Parrell, B., and Lammert, A. C. (2024). "A compressive sensing approach for inferring cognitive representations with reverse correlation," *Behav. Res.* **56**(4), 3606–3618.
- Salge, J. H., Pollmann, S., and Reeder, R. R. (2021). "Anomalous visual experience is linked to perceptual uncertainty and visual imagery vividness," *Psychol. Res.* **85**(5), 1848–1865.
- Schaette, R., König, O., Hornig, D., Gross, M., and Kempter, R. (2010). "Acoustic stimulation treatments against tinnitus could be most effective when tinnitus pitch is within the stimulated frequency range," *Hear. Res.* **269**(1–2), 95–101.
- Stein, A., Engell, A., Junghoefer, M., Wunderlich, R., Lau, P., Wollbrink, A., Rudack, C., and Pantev, C. (2015). "Inhibition-induced plasticity in tinnitus patients after repetitive exposure to tailor-made notched music," *Clin. Neurophysiol.* **126**(5), 1007–1015.
- Tass, P. A., Adamchic, I., Freund, H. J., von Stackelberg, T., and Hauptmann, C. (2012). "Counteracting tinnitus by acoustic coordinated reset neuromodulation," *Restor. Neurol. Neurosci.* **30**(2), 137–159.
- Tunkel, D. E., Bauer, C. A., Sun, G. H. *et al.* (2014). "Clinical practice guideline: Tinnitus," *Otolaryngol. Head Neck Surg.* **151**(2_suppl), S1–S40.
- Tyler, R. S., Aran, J. M., and Dauman, R. (1992). "Recent advances in tinnitus," *Am. J. Audiol.* **1**(4), 36–44.
- Ukaegbe, O. C., Orji, F. T., Ezeanolue, B. C., Akpeh, J. O., and Okorafor, I. A. (2017). "Tinnitus and its effect on the quality of life of sufferers: A Nigerian cohort study," *Otolaryngol. Head Neck Surg.* **157**(4), 690–695.
- Varnet, L., Knoblauch, K., Meunier, F., and Hoen, M. (2013a). "Using auditory classification images for the identification of fine acoustic cues used in speech perception," *Front. Hum. Neurosci.* **7**, 865.
- Varnet, L., Knoblauch, K., Meunier, F., and Hoen, M. (2013b). "Show me what you listen to! Auditory classification images can reveal the processing of fine acoustic cues during speech categorization," in *Interspeech*, Lyon, France (August 25–29, 2013) (International Speech Communication Association), pp. 3167–3171.
- Varnet, L., Knoblauch, K., Serniclaes, W., Meunier, F., and Hoen, M. (2015). "A psychophysical imaging method evidencing auditory cue extraction during speech perception: A group analysis of auditory classification images," *PLoS One* **10**(3), e0118009.
- Varnet, L., and Lorenzi, C. (2022). "Probing temporal modulation detection in white noise using intrinsic envelope fluctuations: A reverse-correlation study," *J. Acoust. Soc. Am.* **151**(2), 1353–1366.
- Varnet, L., Meunier, F., Trollé, G., and Hoen, M. (2016). "Direct viewing of dyslexics' compensatory strategies in speech in noise using auditory classification images," *PLoS One* **11**(4), e0153781.
- Welch, P. (1967). "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," *IEEE Trans. Audio Electroacoust.* **15**(2), 70–73.