Ensemble Coding in Auditory Processing: Outlier Detection and Discrimination

Francis Indaheng [1], Steve Jiang [1], Mantej Panesar [1], Andy Peng [1]
[1]Department of Psychology, University of California, Berkeley

ABSTRACT

It has been previously established that ensemble coding (also known as ensemble perception), the averaging of perceived data, is present in auditory perception (Piazza et al., 2013). However, the relationship between ensemble perception of auditory stimuli and detection of outlier stimuli remains unclear. Specifically, are individuals who are better at discriminating the mean tone automatically better at discriminating outlier tones? To answer this question, we ran an auditory outlier discrimination task (our primary task). Subsequently, we compared the subject's performance on the outlier task to their performance on ensemble and control tasks. Our findings indicate a positive correlation between the ability to ensemble code and the ability to detect and discriminate outlier tones, suggesting that the mechanisms used in auditory ensemble coding have a direct relationship to detecting and distinguishing auditory outlier tones.

KEYWORDS

Ensemble perception; ensemble coding; auditory outliers; summary statistic; auditory system.

INTRODUCTION

The human auditory system, similar to the human visual system, possesses the ability to ensemble code. For visual systems, features such as emotion, orientation, size, etc. can be averaged from a crowd of objects or people (Haberman & Whitney, 2012). Regarding low-level impressions, researchers report that participants are able to accurately extract the average brightness (Bauer, 2009), color (de Gardelle & Summerfield, 2011), speed (Watamaniuk & Duchon, 1992), motion direction (Williams & Sekuler, 1984), and size (Ariely, 2001; Chong & Treisman, 2003) from a group of objects. Regarding high level social information, prior research has found that participants can successfully identify the average emotion (Haberman & Whitney, 2007), gender (Haberman & Whitney, 2007), and the family resemblance of a crowd of individuals.

Likewise, ensemble coding has been shown to play a role in auditory processing; in other words, tones of different frequencies can be averaged from a set of tones (Piazza et al., 2013). Additionally, we know that subjects can be trained to improve their ability to ensemble code (Chang et al., 2016). However, the correlation between ensemble perception and the ability to detect auditory outliers remains unclear, begging the following question:

If someone is skilled at finding the general summary statistic of an auditory percept, will they also be skilled in detecting an auditory outlier?

This question was answered through three empirical experiments - one control experiment, one standard ensemble coding experiment, and one outlier detection experiment. The experiment presented below focuses specifically on the outlier detection paradigm. We hypothesize that there should be a positive correlation between a subject's ability to ensemble code and their ability to detect and distinguish outlier tones, with their outlier detection capabilities increasing as the outlier tone frequency deviates farther from the mean tone frequency.

METHODOLOGY

GENERAL EXPERIMENT - In each trial, subjects listen to seven subsequent tones, one of which is an outlier tone. After the tones are played, subjects are prompted to identify whether the outlier tone was higher or lower frequency than the mean tone, which is self-determined through the subject's own auditory ensemble coding capabilities. The six other tones are within a set range of a mean frequency, while the outlier tone's frequency deviates by a considerable margin from the mean. Here, evidence is provided to demonstrate that the human auditory system is in fact capable of detecting auditory outliers.

OBSERVERS - Five subjects (undergraduates and community members of UC Berkeley) gave consent to participate. There were two females and three males, with the mean age being 20.4 years, and all had normal (or corrected-to-normal) visual and auditory acuity. They were each tested individually in an isolated, quiet room.

STIMULI - A series of seven tones were played consecutively, with each tone played for a duration of 0.3 seconds. Each of the seven different frequency tone was played in a random order, including the outlier tone. The order of the tones and the frequencies' distance from the mean were all counterbalanced to ensure consistency. These frequencies were generated using MATLAB's built-in Musical Instrument Digital Interface (MIDI). To convert from MIDI values to tone frequencies, we used the following equation:

$$Tone\_Frequency = 440 * 2^{((Midi\_Value-69)/12)}$$

Using this tone frequency, we constructed a playable sinusoidal wave using the following equation:

$$\sin(2 * \pi * Tone\_Frequency * Tone\_Length)$$

Finally, an onset and offset ramp, in the form of cosine waves, were created to play smoother, more distinct tones without static. The following equations were used:

$$Onset = (1 + 2 * \pi * Frequency\_Ramp * Ramp\_Vector ./ fs - \pi/2)/2$$
$$Offset = (1 + 2 * \pi * Frequency\_Ramp * Ramp\_Vector ./ fs + \pi/2)/2$$

The six tones produced in the random sequence were ±6, ±4, and ±2 semitones from the mean. The one outlier tone was randomly chosen from the set of tones ±16, ±14, ±10, and ±6 semitones from the mean. The mean semitone was randomly chosen from the MIDI value range 50 (D3) to 80 (A#5), inclusive.

PROCEDURES - Psychological subjects were prompted to input their demographic information including: their initials , gender, and age. . Next, they were presented with following instructions:

You will listen to 7 audio tones. 1 tone is an outlier.If the outlier is a higher tone than the average tone, press the "H" key. If the outlier is a lower tone than the average tone, press the "L" key. Press any button to continue.

Subjects were given unlimited time to read and understand the task before them. Additionally, the experimenter showed them 2 practice trials and verbally confirmed that they fully comprehended the task. During the actual experiment, subjects listen to seven tones presented in a random temporal order, one of which is the outlier tone. After the tones are played, the subject is prompted to press the “H” key if the outlier tone was of higher frequency than the mean tone, or the “L” key if the outlier tone was of lower frequency than the mean tone. For clarification, it is important to understand that the subjects were not directly played the mean semitone. Rather, the six tones were deviated from the aforementioned mean tone. This setup requires the subject to utilize their own auditory ensemble coding to determine the mean, before they can determine if the outlier was a higher or lower frequency.

RESULTS

Utilizing a logistic regression equation, we were able to construct psychometric curves [Fig. 1], plotting the percentage of higher responses (y-axis) as a function of distance from the outlier to the mean (x-axis). Our data concludes that the subjects were capable of detecting and discriminating the outlier tones, with an average point of subjective equality (PSE) of 0.002826592 ($p < 0.005$ for all subjects).

DISCUSSION

In correlation to data from the general auditory ensemble coding tasks, the same subjects found a positive correlation of +0.8897 ($p = 0.0433$) between their performances on the general auditory ensemble coding data and the auditory outlier data [Fig. 2]. In comparison, the data for both the general auditory ensemble coding and the auditory outlier tasks held no significant correlation ($p = 0.2464$ and $p = 0.6767$, respectively) in relation to the control task.

Here, our data points to solid evidence that the human auditory system is in fact capable of detecting outliers, as well as discriminating them against a range of tones centered around a mean tone. Furthermore, subjects who excelled at the ensemble coding task did comparably well on the outlier detection task, indicating an answer to our previously stated question (*If someone is skilled at finding the general summary statistic of an auditory percept, will they also be skilled in detecting an auditory outlier?*). Our initial results suggest that someone who is skilled at finding the general summary statistic of an auditory percept will also be skilled in detecting an auditory outlier. This outcome is theoretically justified because in order to detect a perceptual average, the outlier is necessarily calculated and vise versa. While these are only pilot results, these findings have important real-life implications. We know that ensemble coding is a skill that can be trained and improved upon (Yamanashi, Chang, Whitney, In Review). Therefore, our initial findings strongly suggest that training general ensemble perception will automatically enhance outlier detection. This is a vitally important implication that should be directly tested in future studies.

Such findings would be applicable to real-world situations in which outlier detection, auditory or visual, is a fundamental component (i.e. disaster response). Future research that more specifically characterizes the relationship ensemble perception, outlier detection, and training could provide an additional instructional methods for disaster response personnel seeking to improve their outlier detection performance in real-life situations.

ACKNOWLEDGEMENTS

This study was made possible by the Daniel D. Sonderegger Perception Science Internship, and the faculty of the Whitney Lab at UC Berkeley, who provided invaluable assistance with all steps of the research process. A special thanks to Dr. Allison Yamanashi Leib who believed in us every step of the way, and Jin Jeon for his consistent willingness to help.

REFERENCES

Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science, 12*, 157–162.

Bauer, B. (2009). Does Steven's power law for brightness extend to perceptual brightness averaging? *Psychological Record, 59*, 171–186.

Chang, Kelly, et al. "Training Ensemble Perception." Journal of Vision, vol. 16, no. 12, Jan. 2016, p. 52., doi:10.1167/16.12.52.

de Gardelle, V., & Summerfield, C. (2011). Robust averaging during perceptual judgment. *Proceedings of the National Academy of Sciences, USA, 108*, 13341–13346.

Haberman, J., & Whitney, D. (2007). Rapid extraction of mean emotion and gender from sets of faces. *Current Biology, 17*, 751–753.

Haberman, J., & Whitney, D. (2012). “Ensemble Perception: Summarizing the scene and broadening the limits of visual processing.” In J. Wolfe and L. Robertson (Eds.), *From Perception to Consciousness: Searching with Anne Treisman.* Oxford University Press.

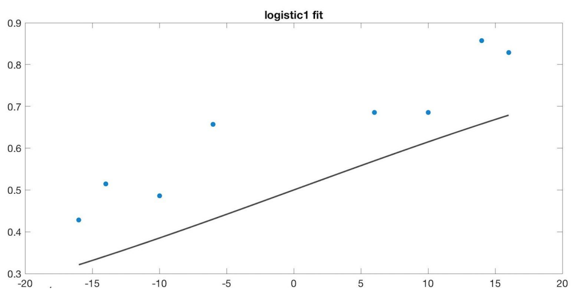
Piazza, Elise A., et al. “Humans Use Summary Statistics to Perceive Auditory Sequences.” Psychological Science, vol. 24, no. 8, Dec. 2013, pp. 1389–1397., doi:10.1177/0956797612473759.

Watamaniuk, S. N. J., & Duchon, A. (1992). The human visual system averages speed information. *Vision Research, 32*, 931–941.

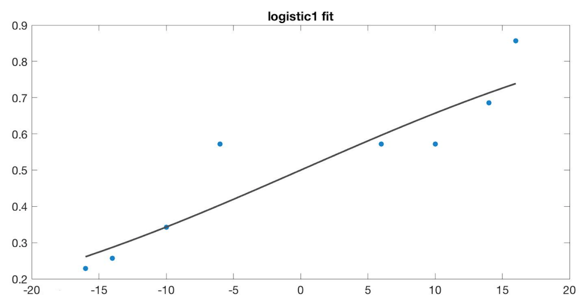
Williams, D. W., & Sekuler, R. (1984). Coherent global motion percepts from stochastic local motions. *Vision Research, 24*, 55–62.

FIGURES

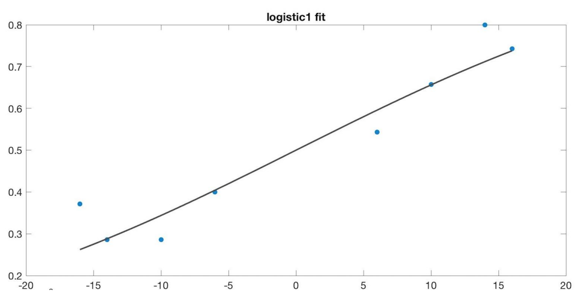
KPA



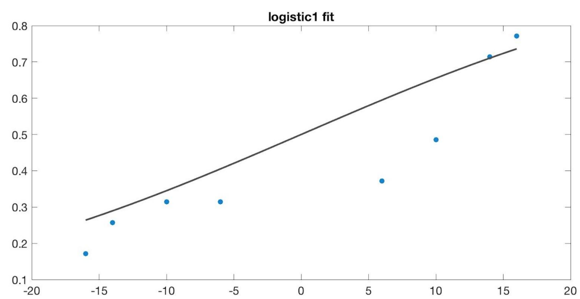
AK



JM



EK



AL

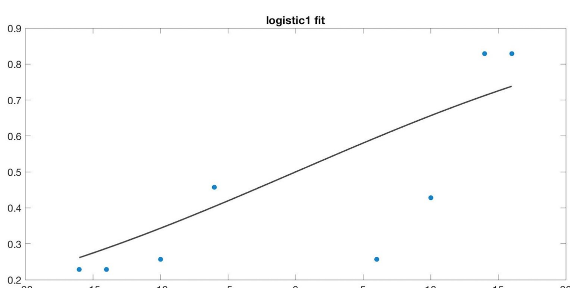


Figure 1

A psychometric curve analysis plotting the percentage of "h" responses ("outlier is higher than mean tone") on the y-axis versus the true distance between the outlier and the mean (in semitones) on the x-axis.

Initials	PSE (auditory outlier)	PSE (auditory ensemble)	Accuracy (auditory control)	Accuracy (auditory outlier)	Accuracy (auditory ensemble)
KPA	0.00064325	4.0222	0.8653846154	0.6214285714	0.57
JM	0.0041507	1.1581	0.8461538462	0.675	0.695
AL	0.004326	1.9675	0.8076923077	0.6464285714	0.68
AK	0.00059451	1.8365	0.8413461538	0.6607142857	0.6826923077
EK	0.0044185	1.7893	0.8413461538	0.6607142857	0.6586538462

Figure 2

A table containing summary data for all five subjects. The first two columns contain the points of subjective equality (PSE) of the psychometric curves for the auditory outlier and auditory ensemble tasks, respectively. The last three columns contain the subjects' average accuracies for the control, outlier, and ensemble tasks, respectively. When correlated, there is a Pearson r correlation of 0.8897 ($p = 0.0433$) between accuracy on the auditory outlier and ensemble tasks. There was no significant correlation between the auditory control task and either the auditory outlier or auditory ensemble tasks ($p = 0.6767$ and $p = 0.2464$, respectively).