

Music and Brain Symposium 2013:  
Hearing Voices

*Acoustics of Imaginary Sound*  
Chris Chafe

Center for Computer Research in Music and Acoustics,  
Stanford University



<http://www.youtube.com/watch?v=cGZtC4m52zM>

The Mommyheads in a 2010 Time Warner Cable Ad

*IODA Marketing*





<https://www.youtube.com/watch?v=X073Ou7EMbo>

The real deal: rehearsing over the Internet



# A decade of Internet music performance

- open jamming
  - rehearsals, concerts
  - auditions
  - teaching
  - recording, remote overdub
- ...and physically avoiding players... (for whatever reason)

A new medium translates into new experiences

- no distance cue
- sound arrives before visual
- latencies anywhere from network rooms to stadiums

Most importantly, how does delay affect ensemble synchronization?

## Effect of temporal separation on synchronization in rhythmic performance

---

Chris Chafe, Juan-Pablo Cáceres, Michael Gurevich§

Center for Computer Research in Music and Acoustics (CCRMA), Stanford University, Stanford, CA 94305, USA; e-mail: cc@ccrma.stanford.edu; §Sonic Arts Research Centre (SARC), Queen's University Belfast, Belfast BT7 1NN, Northern Ireland, UK

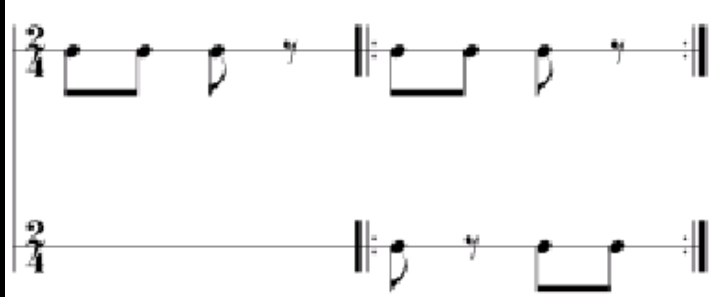
Received 14 May 2009, in revised form 18 April 2010

---

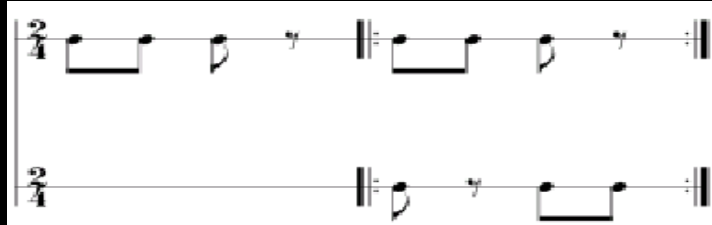
**Abstract.** A variety of short time delays inserted between pairs of subjects were found to affect their ability to synchronize a musical task. The subjects performed a clapping rhythm together from separate sound-isolated rooms via headphones and without visual contact. One-way time delays between pairs were manipulated electronically in the range of 3 to 78 ms. We are interested in quantifying the envelope of time delay within which two individuals produce synchronous performances. The results indicate that there are distinct regimes of mutually coupled behavior, and that ‘natural time delay’—delay within the narrow range associated with travel times across spatial arrangements of groups and ensembles—supports the most stable performance. Conditions outside of this envelope, with time delays both below and above it, create characteristic interaction dynamics in the mutually coupled actions of the duo. Trials at extremely short delays (corresponding to unnaturally close proximity) had a tendency to accelerate from anticipation. Synchronization lagged at longer delays (larger than usual physical distances) and produced an increasingly severe deceleration and then deterioration of performed rhythms. The study has implications for music collaboration over the Internet and suggests that stable rhythmic performance can be achieved by ‘wired ensembles’ across distances of thousands of kilometers.

Subjects = students and staff at  
Stanford (paired randomly)

Task = play rhythm accurately,  
keep an even tempo  
(no strategies given)



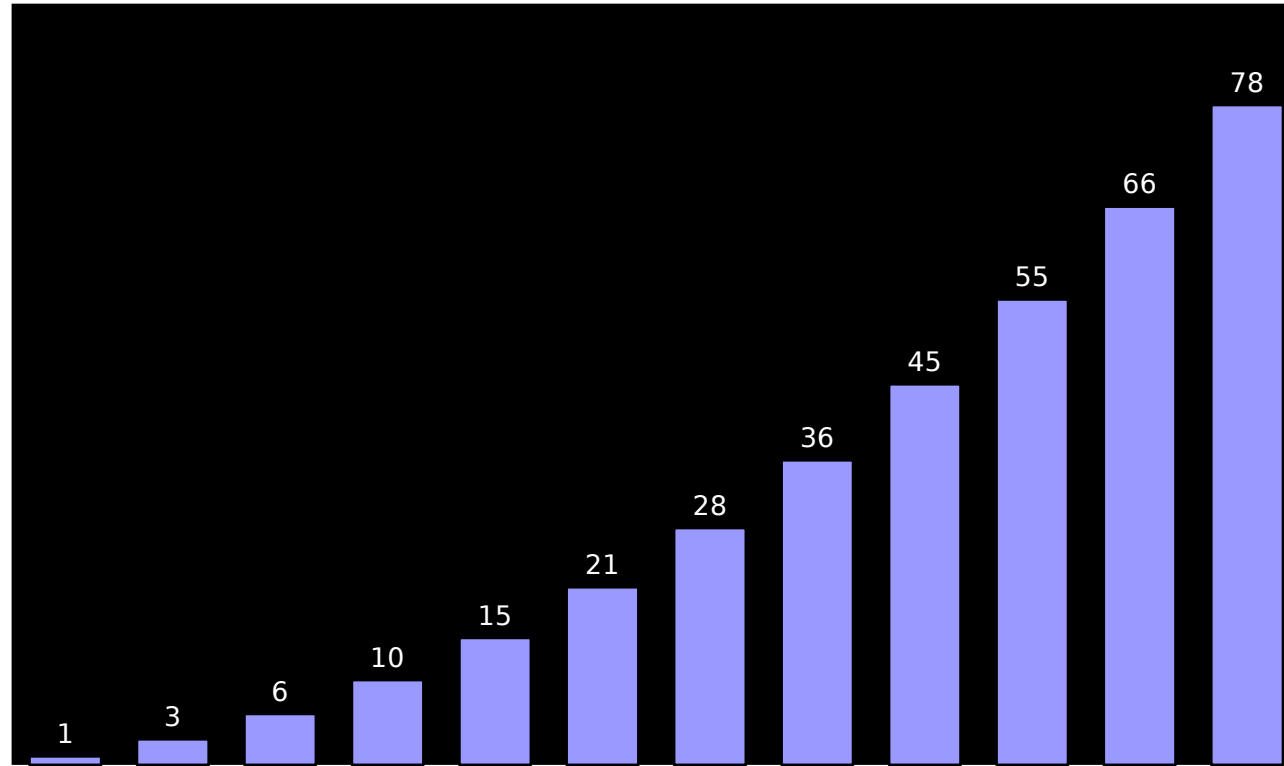
Interlocking rhythm



Sound  
(3ms delay each,  
metronome cue = mm94)

<https://ccrma.stanford.edu/~cc/vox/teleTalkMedia/wav/clap.wav>

# Delays: 1 - 78ms in 12 steps

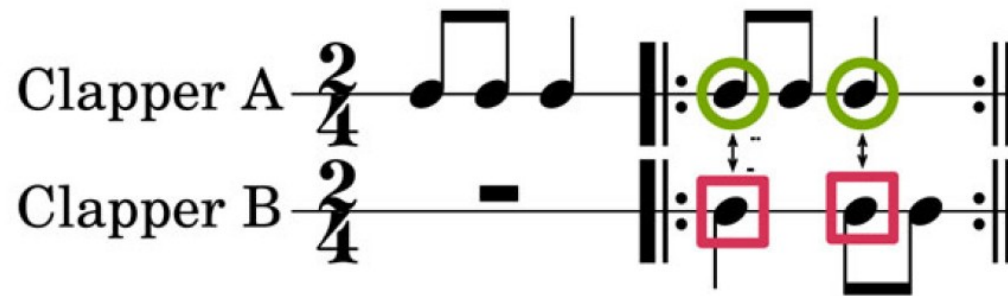


Delay Times (ms) Tested

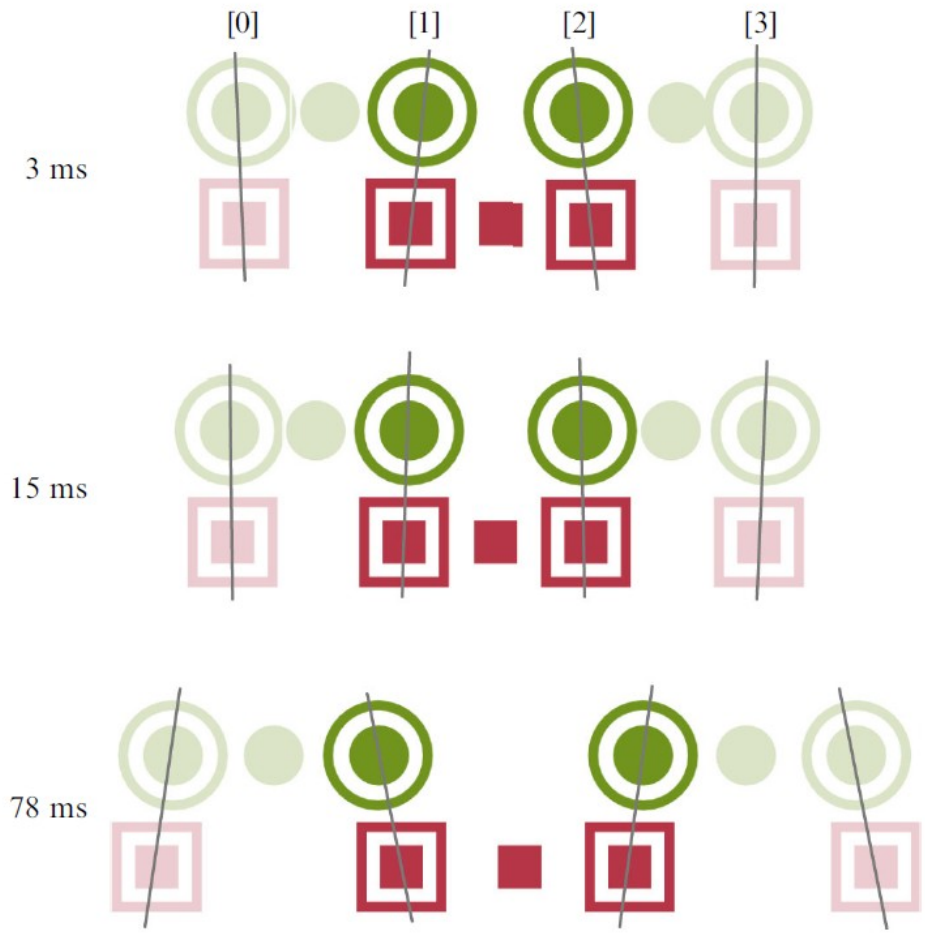
Deceleration from longer delay  
but where does it start to cause  
trouble?

Sound  
(78ms delay each,  
metronome cue = mm90)

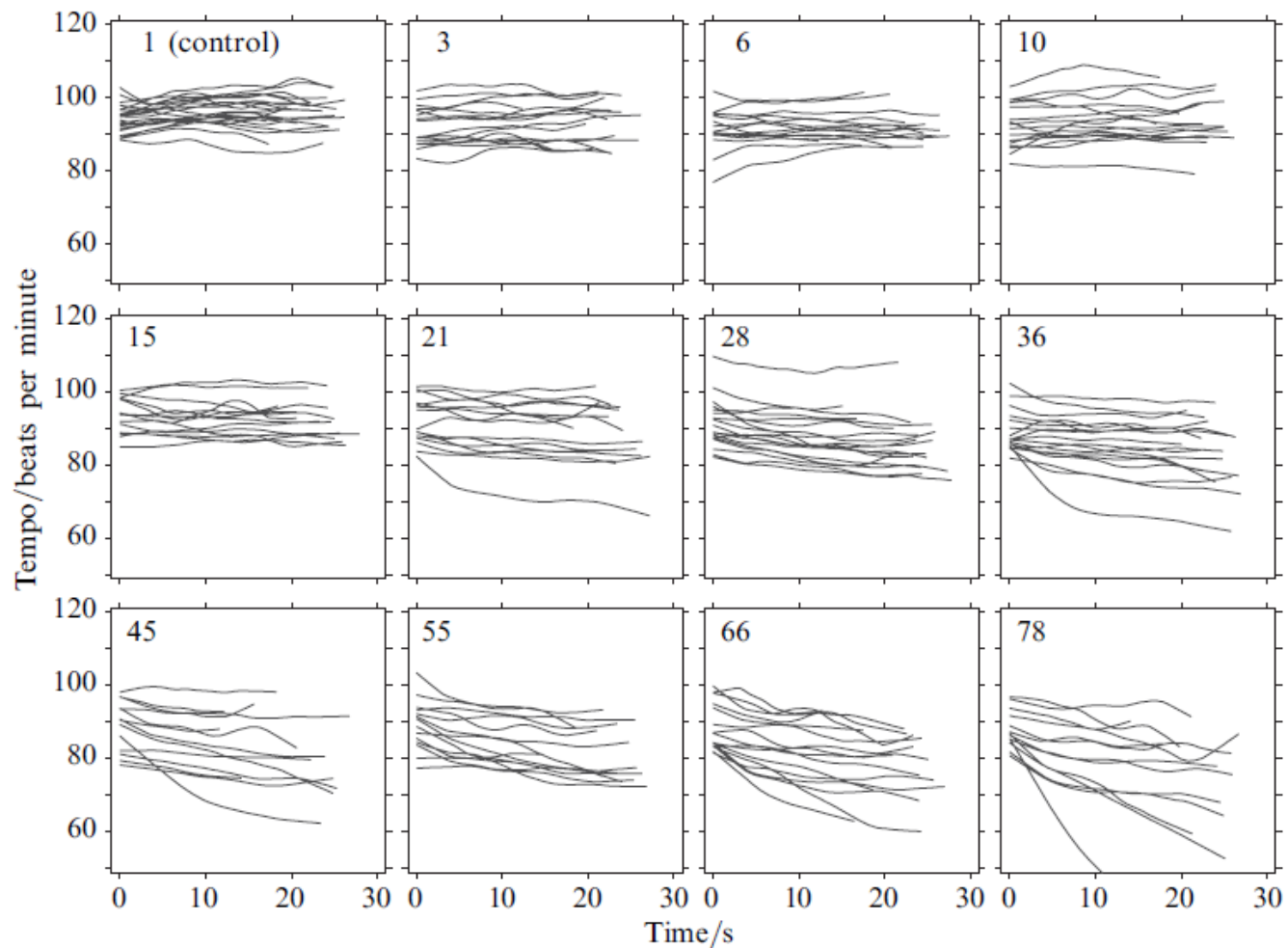
<https://ccrma.stanford.edu/~cc/vox/teleTalkMedia/wav/77.wav>



**Figure 1.** [In color online, see <http://dx.doi.org/10.1068/p6465>] Duo clapping rhythm used to test the effect of temporal separation. Subjects in separate rooms were asked to clap the rhythm together while hearing each other's sound delayed by a slight amount. Common beats in the duo clapping rhythm provide reference points for analysis of ensemble synchronization. Circles and squares represent synchronization points.



**Figure 4.** [In color online.] Lead/lag at different delays. Clapper in ‘San Francisco’ is green circles, clapper in ‘New York’ is red squares. Ideally, each vertically adjacent pair of events is simultaneous. Leading or lagging by one subject with respect to the other at these points is related to delay: leading at 3 ms; approximately synchronous at 15 ms; lagging at 78 ms. Lead/lag is measured with respect to measure-length periodicity. Odd-numbered events have inverted (antiphase) sign.



**Figure 5.** All trials' tempo curves grouped by delay. Tempo acceleration during a given performance is tracked by measuring inter-onset intervals as shown in figure 3. The delays (in ms) are shown in top left corner of each graph.

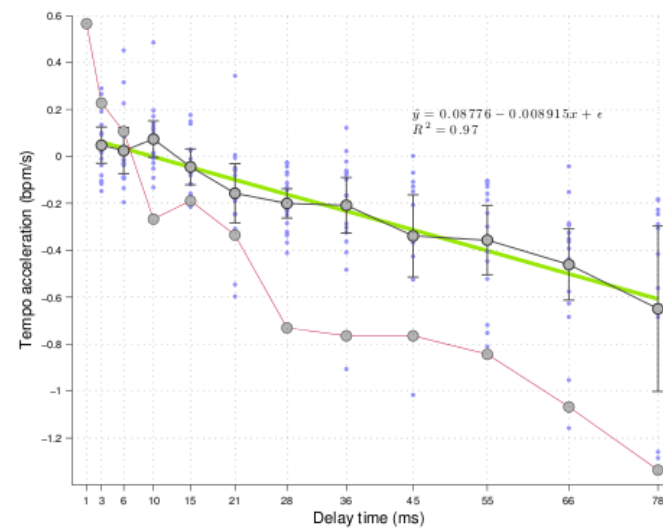


Figure 4.11: A single measure of tempo acceleration (its mean) is compared for all performances. Red line is the mean acceleration of the predicted tempos using the coupled oscillators.

J.P. Cáceres, Synchronization in Rhythmic Performance with Delay, PhD Thesis, 2013

<https://ccrma.stanford.edu/~cc/vox/teleTalkMedia/avi/slowtw.avi>

## Slow Mommyheads

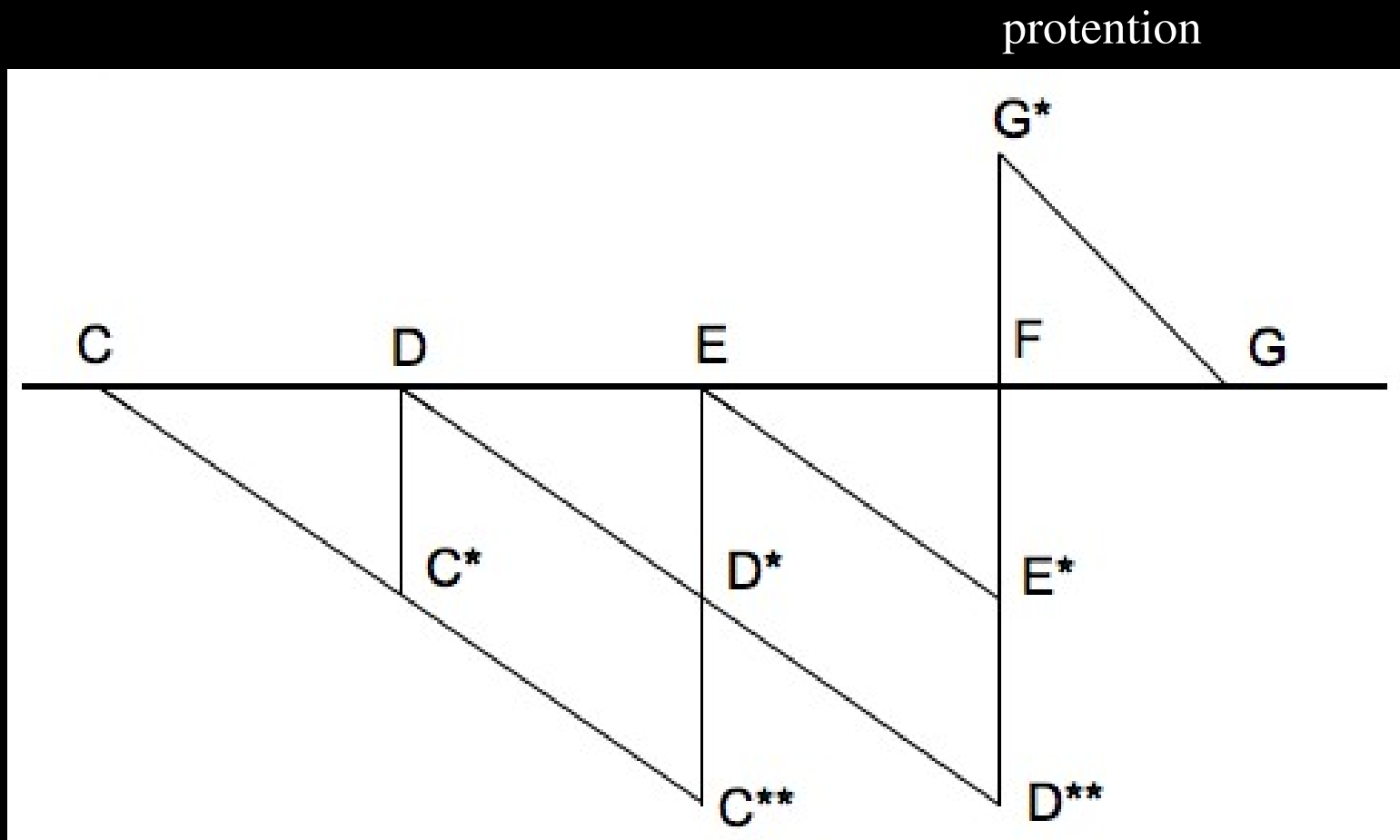


"We are constantly aware of a certain duration—the specious present—varying from a few seconds to probably not more than a minute, and this duration (with its content perceived as having one part earlier and another part later) is the original intuition of time."

W. James, *The Principles of Psychology* (1890)

"It belongs to the essence of perception not only that it has in view a punctual now and not only that it releases from its view something that has just been, while 'still intending' it in the original mode of 'just-having-been', but also that it passes over from now to now and, in anticipation, goes to meet the new now. The waking consciousness, the waking life, is a living-towards, a living that goes from the now towards the new now."

E. Husserl, *On the Phenomenology of the Consciousness of Internal Time* (1893–1917)

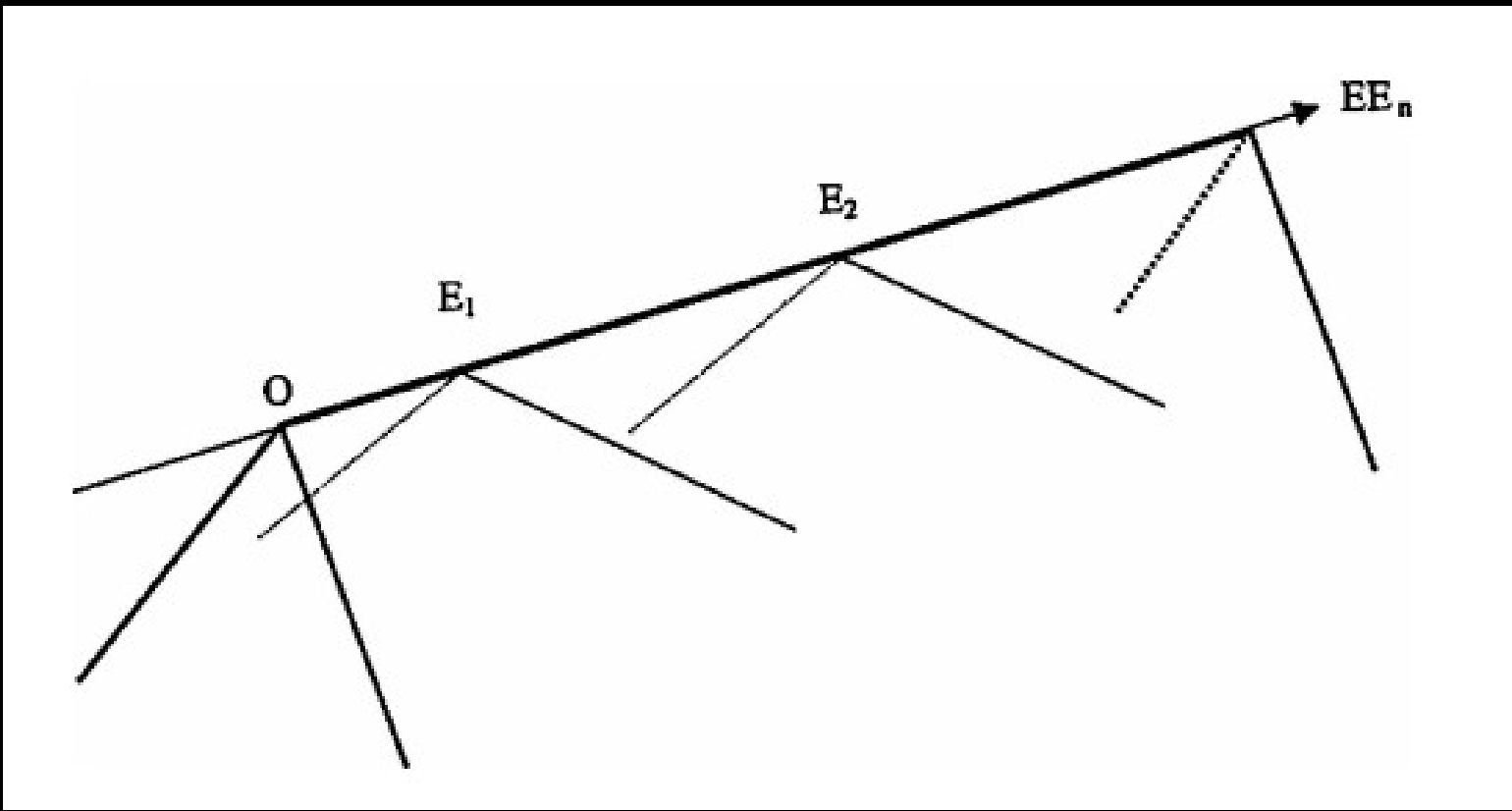


flow of primal impressions → retention

B. Dainton, *Temporal Consciousness*, Stanford Enc. Of Phil. (2010)

"...time presented itself to Husserl as the best candidate for the self-manifestation of consciousness: the ordering of past-present-future as a phase-continuum is not in and of itself the manifestness of time, but the lived tension of its passage, the upsurge of its movement, the being-lived-through (Erleben) [experience] of whatever order may otherwise belong to time."

J. Dodd, *Reading Husserl's Time-Diagrams from 1917/18* (2005)



E. Husserl, *Die 'Bernauer Manuskripte' über das Zeitbewußtsein* (1917/18)

Reading ahead  
Planning ahead  
Vamping automatically

...imply “thinking in sound”

“thinking in sound”

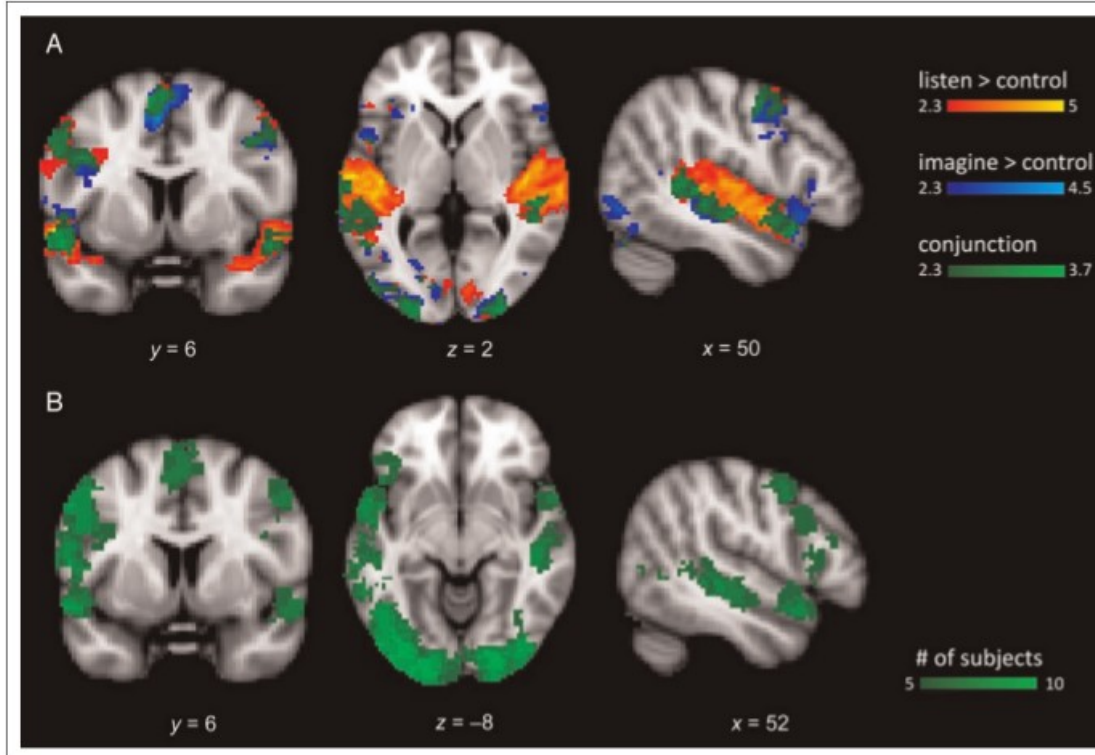
"audiation"

"audialization"

"imaginary sound"

"auditory imagery"

"audiles"



**Figure 2.** (A) Activity during the encoding phase of the experiment during listening to (red) and imagination (blue) of familiar melodies compared with baseline (cluster threshold in FSL with  $z > 2.3$ ,  $p < .05$ ). Green areas show the overlap of the activation maps. (B) Overlap of activity during imagining and listening compared with baseline was also computed on an individual level. Green areas show voxels where at least five participants showed significant activity in both conditions compared with baseline.

S. Herholz, et al. *Neuronal Correlates of Perception, Imagery, and Memory for Familiar Tunes*, J. Cog. Neurosci. (2012)

Consider loudness as an acoustical dimension to isolate

"quasi-loudness"  
phenomenological agreement as  
method

D. Schmicking, *Is there imaginary loudness? Reconsidering phenomenological method*, Phenomenology and the Cognitive Sciences, (2005)

Behavioral methods  
Imaging methods

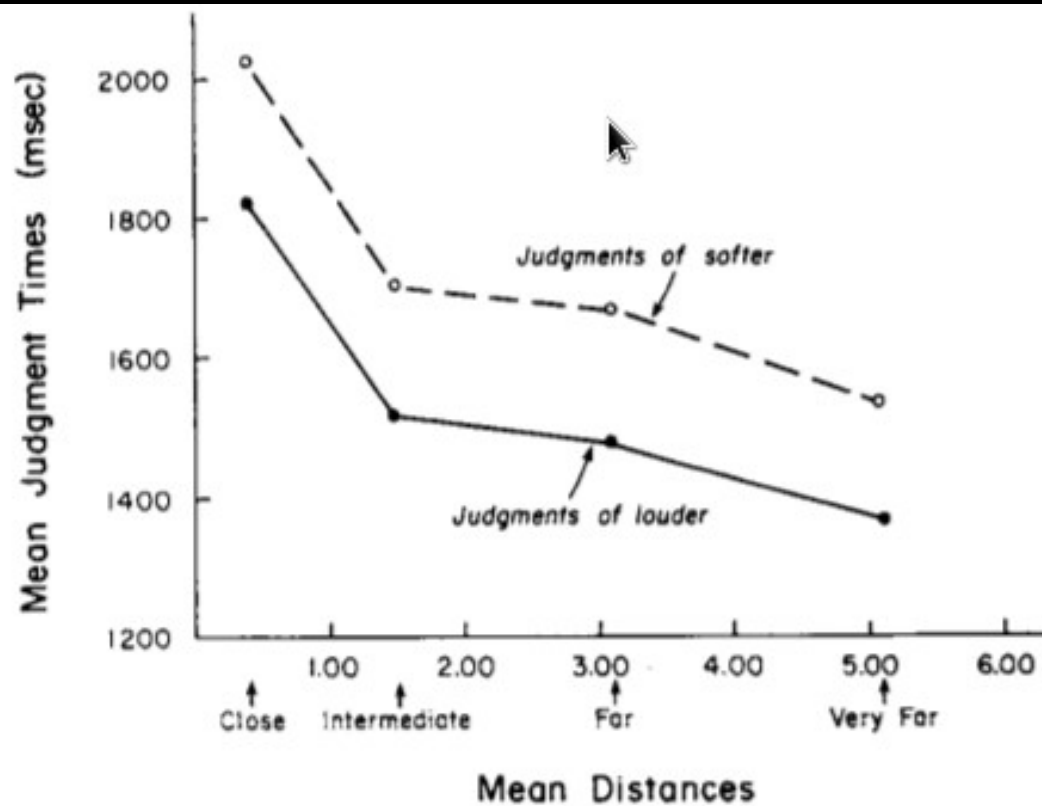
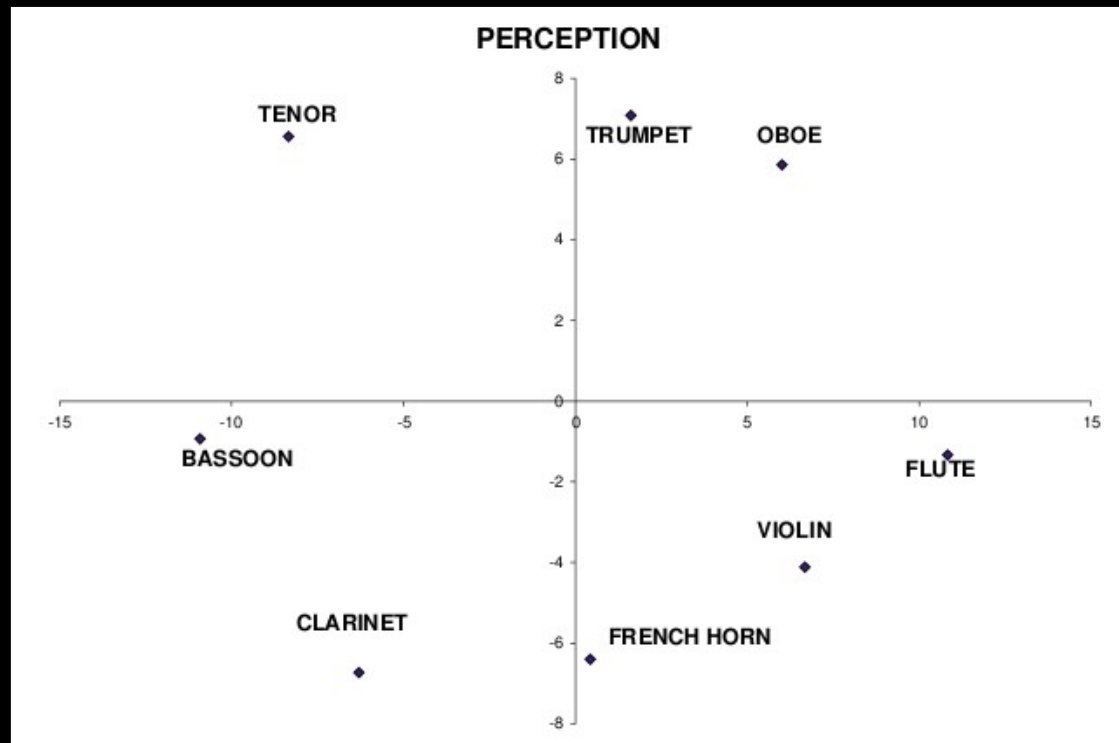


Figure 2. Symbolic distance and preference effects, Experiment 4.

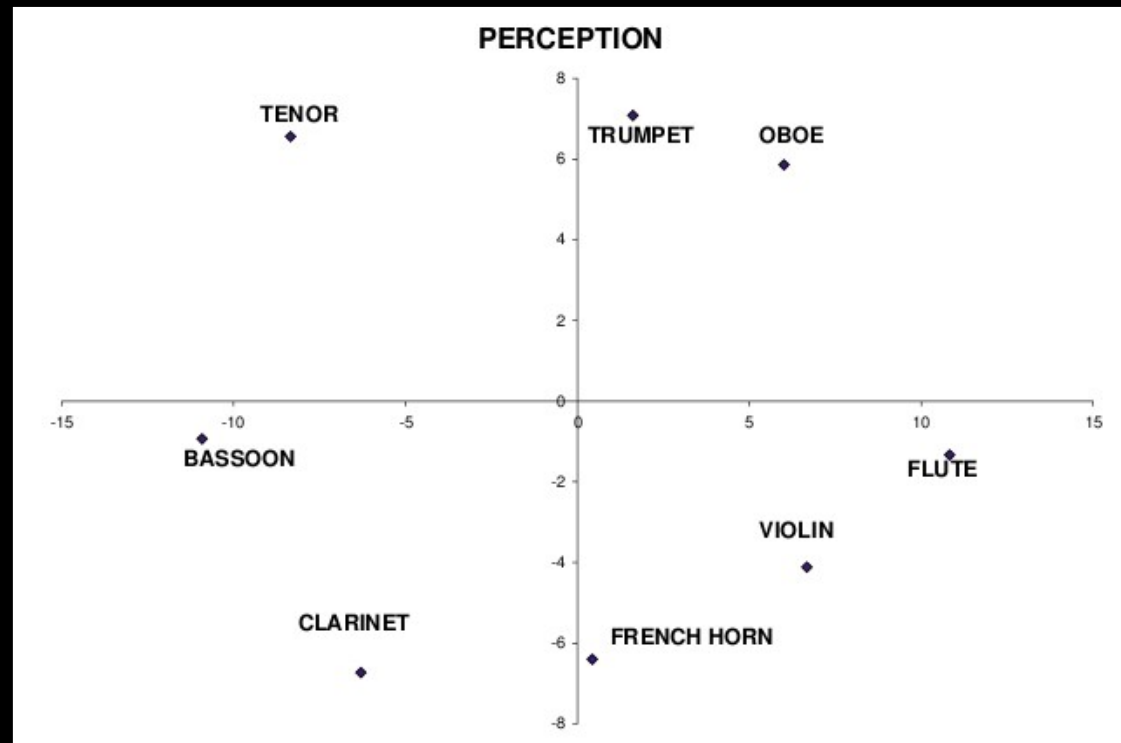
M. Intons-Peterson, *The role of loudness in auditory imagery*, Memory & Cognition (1980)

Consider "quasi-timbre"



A. Halpern, et al. *Behavioral and neural correlates of perceived and imagined musical timbre*, *Neuropsychologia* (2004)

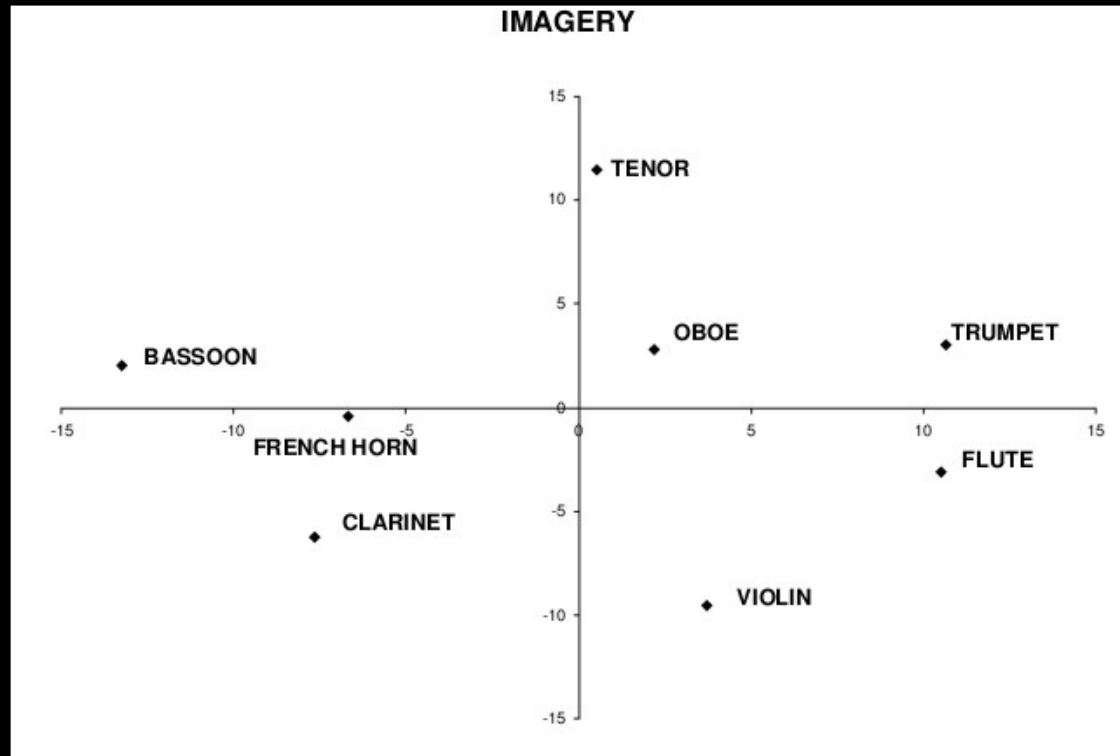
↑  
nasality,  
spectral  
centroid



brilliant →

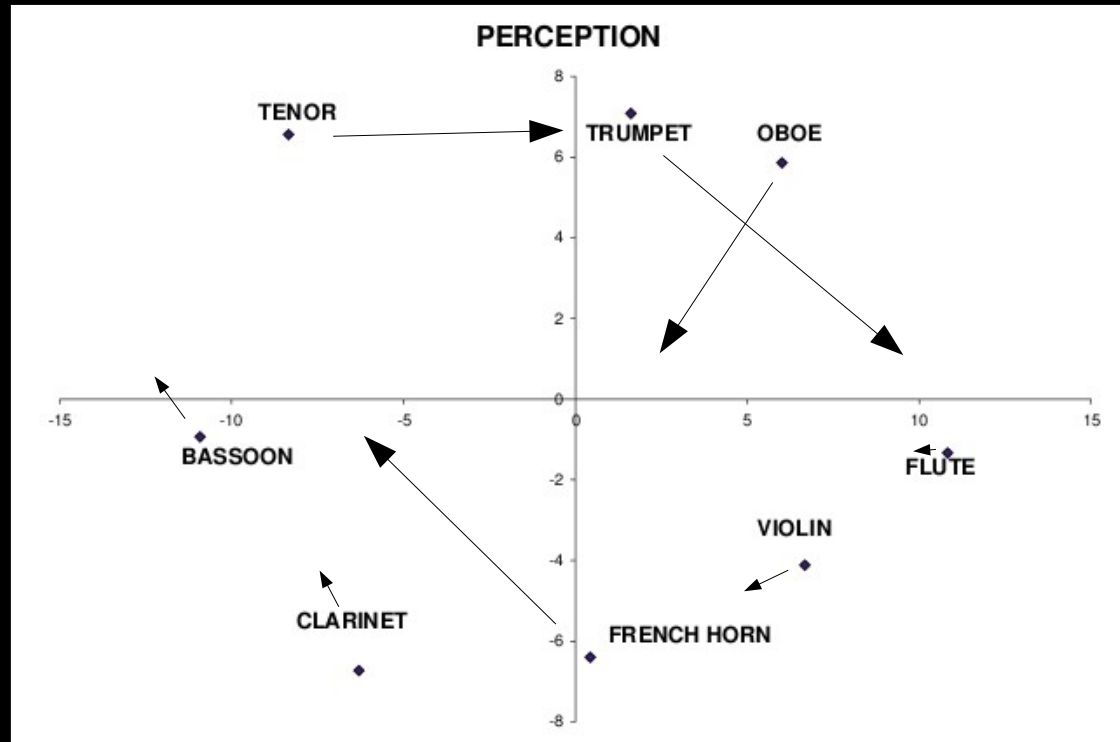
cf. R. Kendall, et al. *Perceptual and Acoustical Features of Natural and Synthetic Orchestral Instrument Tones*, Music Perception (1999)

↑  
nasality,  
spectral  
centroid



brilliant →

↑  
nasality,  
spectral  
centroid



brilliant →

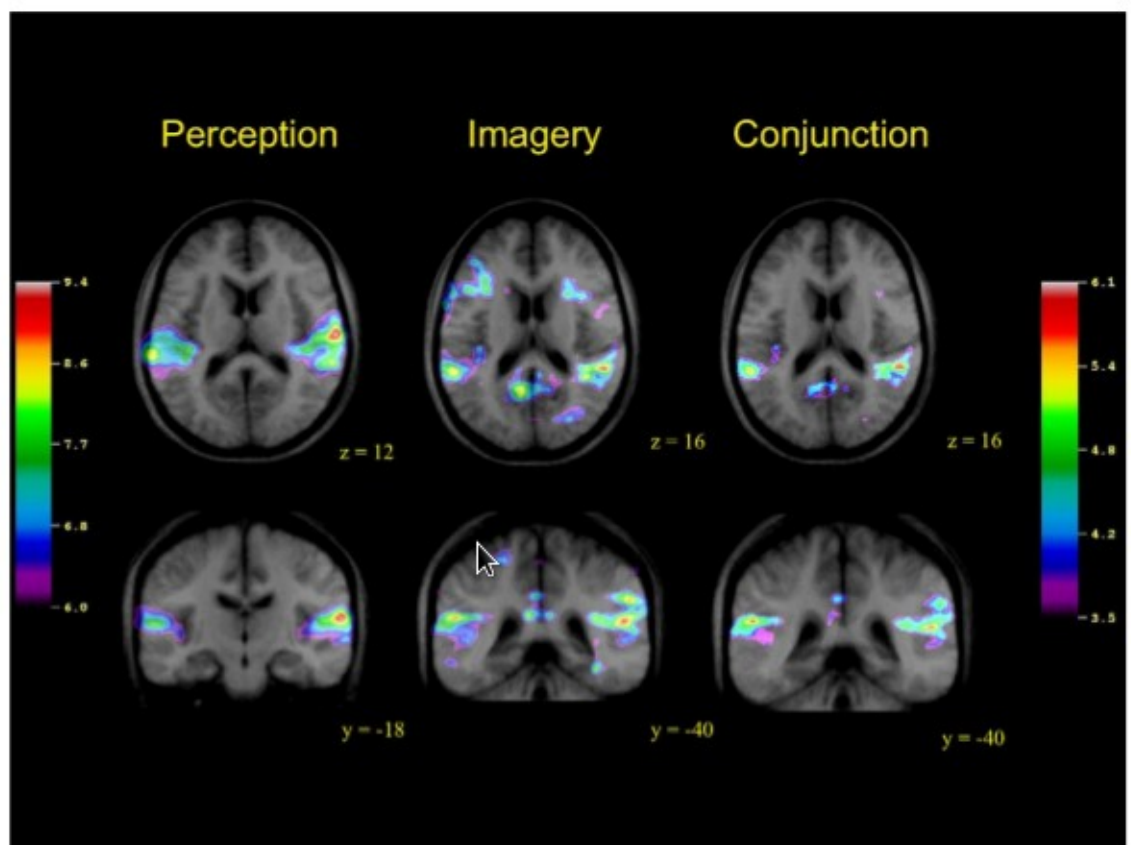


Fig. 3. Statistical parametric maps of fMRI signal changes ( $t$ -values, in color scale) superimposed upon group-average anatomical MR image at selected locations. In all panels, the top image illustrates a horizontal section taken through the temporal lobes (position indicated by stereotaxic  $z$ -coordinate), and the bottom panel illustrates a coronal section taken close to the peak location of signal change (position indicated by stereotaxic  $y$ -coordinate). Left panel: data from the timbre perception vs. visual imagery control condition (data referred to left color bar). Significant BOLD signal changes are evident in primary and adjacent auditory cortices, with an asymmetric response in the right planum temporale. Middle panel: data from the timbre imagery vs. visual imagery control condition (data referred to right color bar). BOLD signal changes are evident within posterior portions of secondary auditory cortices. Right panel: data from the conjunction analysis of perception and timbre imagery (data referred to right color bar). Significant overlap in posterior auditory areas is seen.

A. Halpern, et al. *Behavioral and neural correlates of perceived and imagined musical timbre*, *Neuropsychologia* (2004)

Consider "quasi-location"

– no studies apparently





Chris will now stop talking.

A different voice will be used as we continue the presentation.

It is the sound of your inner voice  
which is what you are hearing as you  
read these words.

How would you rate the clarity of the  
imagined sound?

Choose:

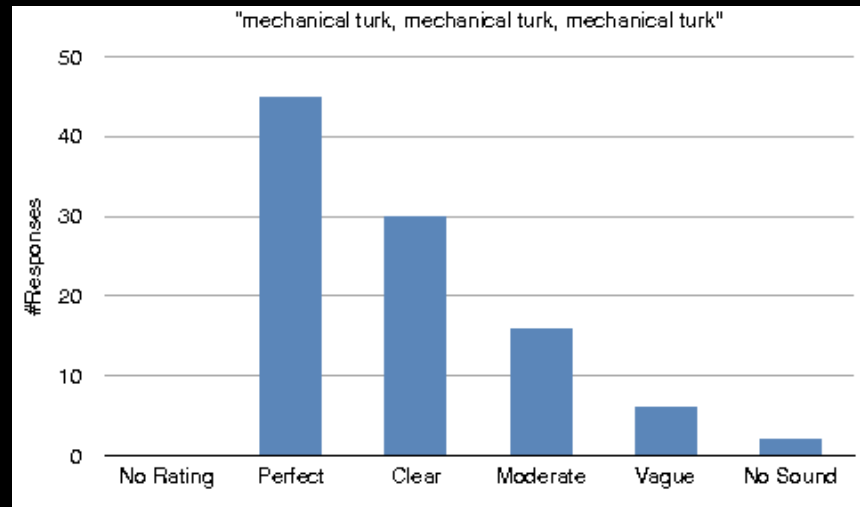
Perfect, Clear, Moderate, Vague, No image

adopted from Vividness of Visual  
Imagery Questionnaire (VVIQ Scale)

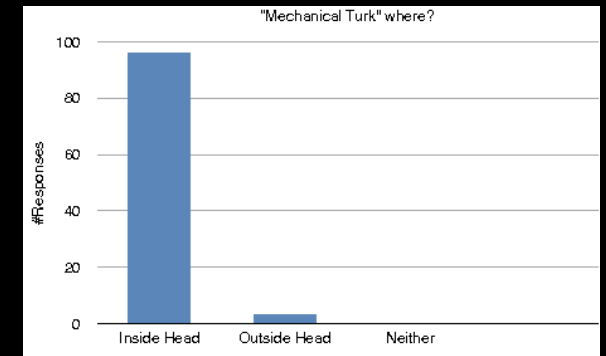
D. Marks, *Visual Imagery Differences in the Recall of Pictures*  
Brit. J. Psych (1973)

Is it located inside your head?

# How 100 Mechanical Turk workers rated clarity:



...and location inside head:



Imagine and rate for clarity

-- a friend or relative saying anything

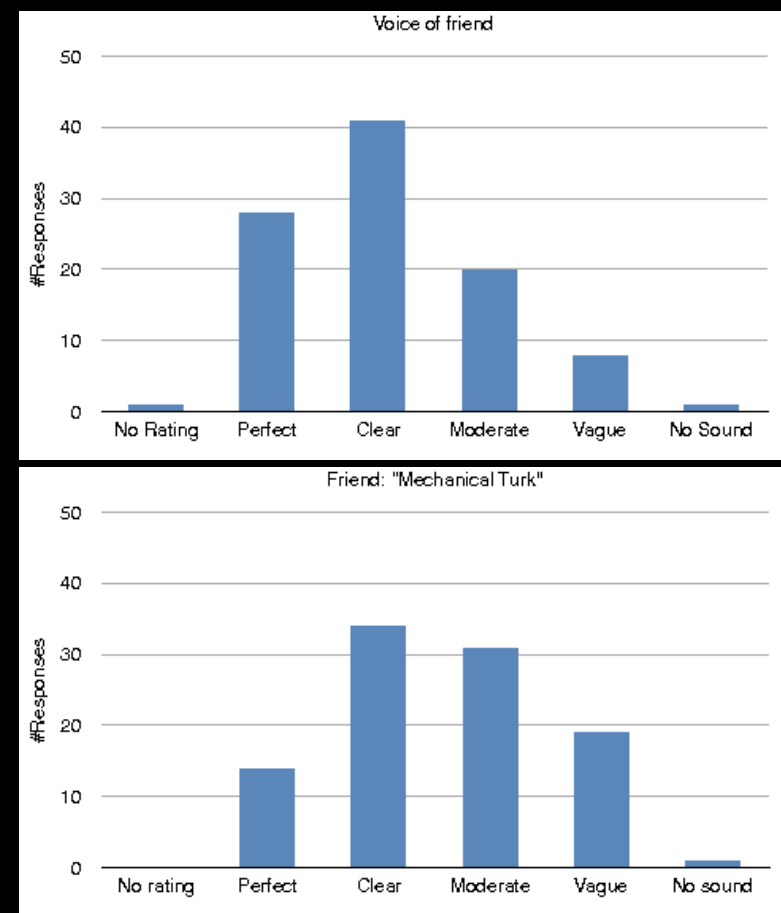
Imagine and rate for clarity

- a friend or relative saying anything
- them saying “mechanical turk”

clarity of them  
saying anything

saying “mechanical turk”

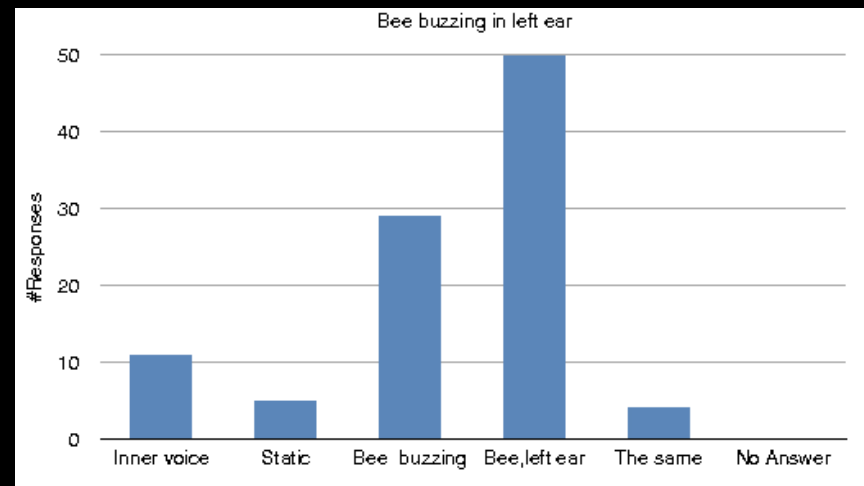
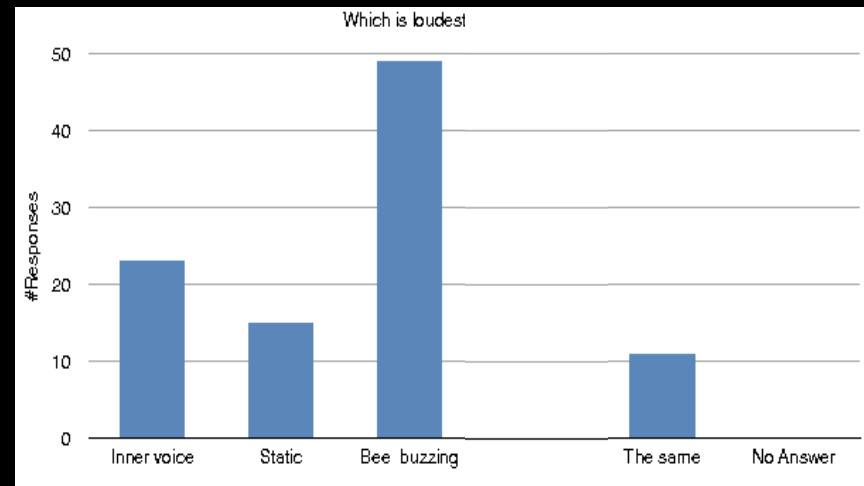
(a potential control)



Effect of location  
on loudness

implicitly inside head

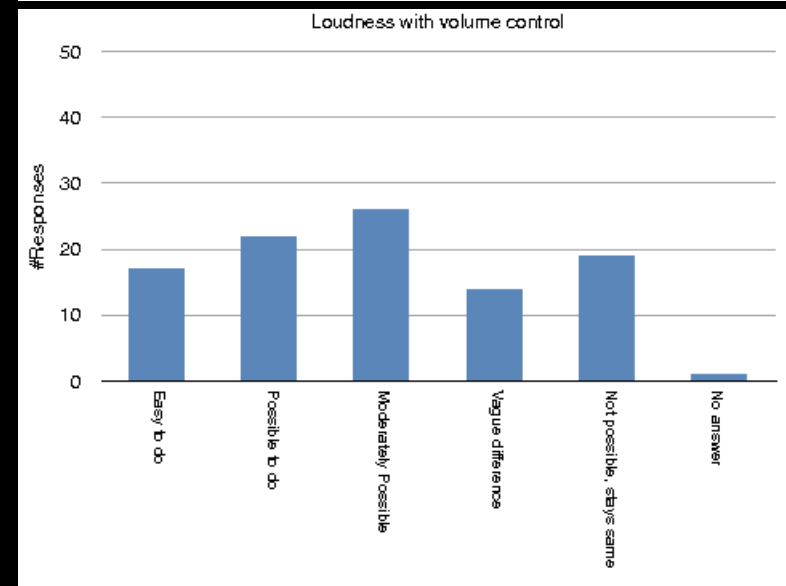
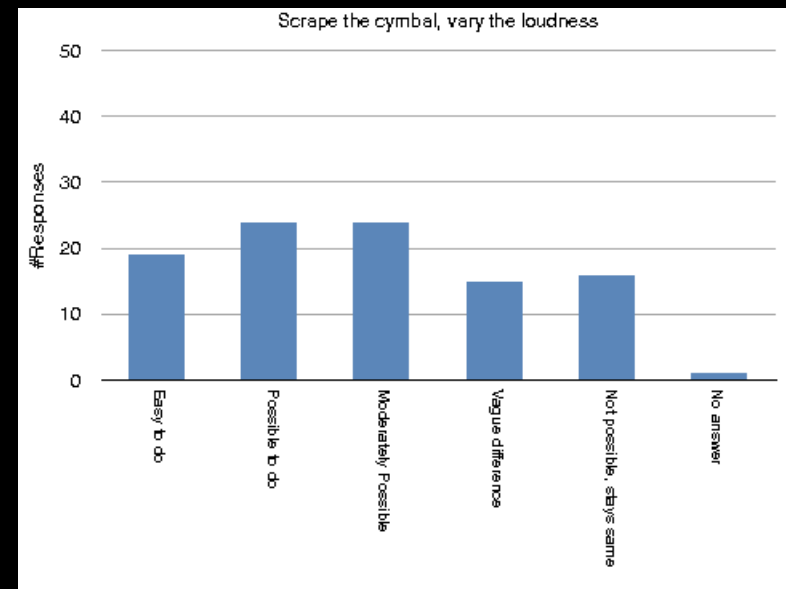
explicitly outside



Imagine scraping a stick across the surface of the cymbal

Can the cymbal sound be louder or softer?  
(no instruction how to do it)

Can it be louder or softer  
the way a volume control  
changes it?

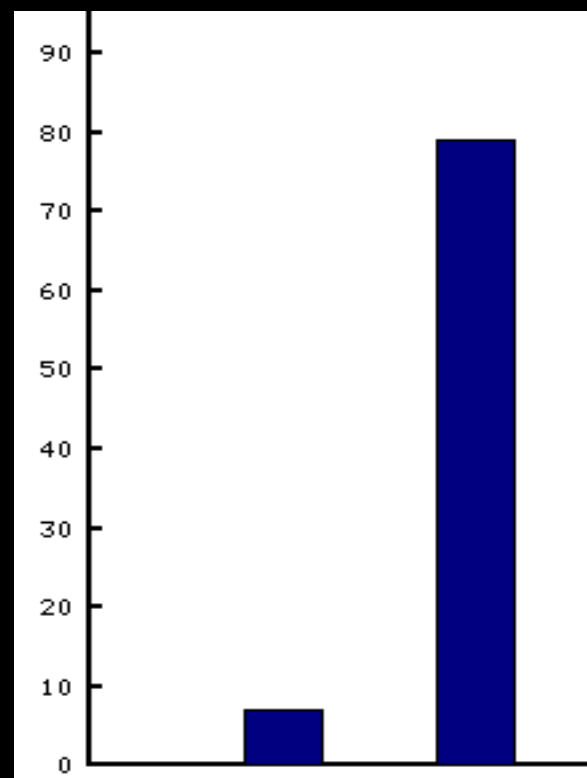


imagine a large ball like a soccer ball  
dropped onto grass from waist height

imagine a smaller, harder ball, for  
example a hard baseball also dropped  
onto pavement from waist height

Try to isolate just one bounce for the  
two different balls.

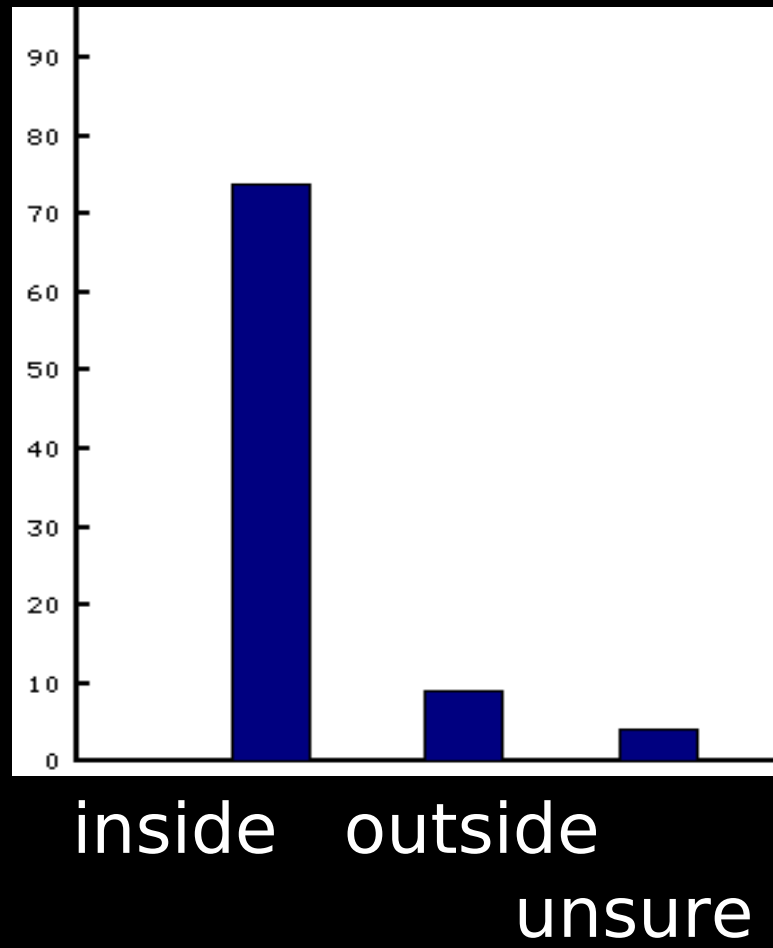
ball sounds are

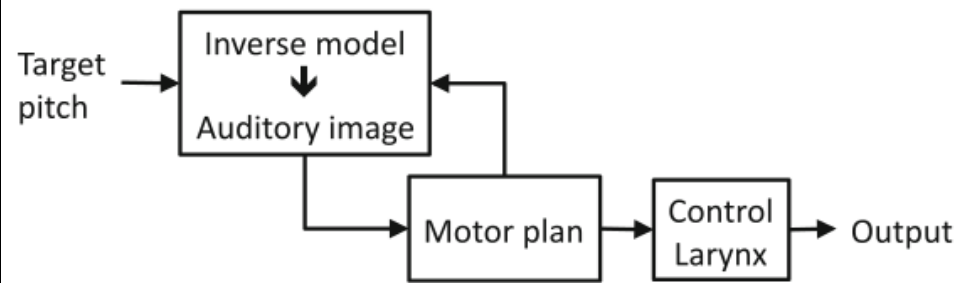


same

different

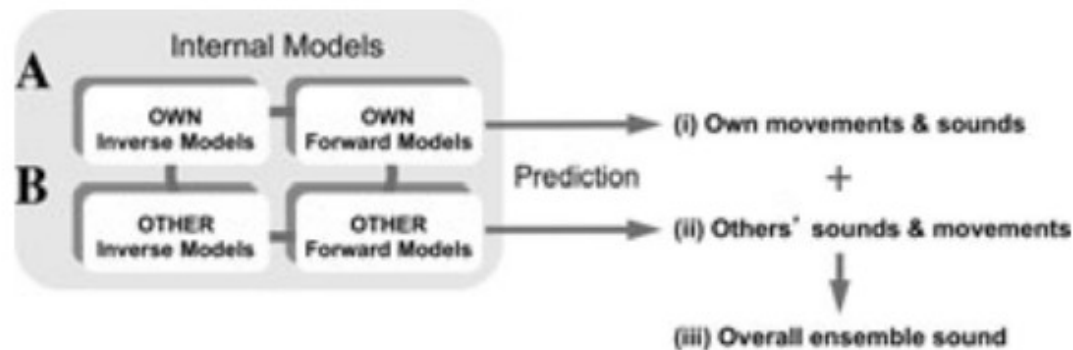
ball sounds are





**Fig. 1** Representation of the vocal pitch imitation system, including the hypothesized inverse internal model. (Adapted from Hickok, Houde, & Rong, 2011)

P. Pfordresher, A. Halpern, *Auditory imagery and the poor-pitch singer*, *Psychon. Bull. Rev.* 2013)



**Figure 2.** Forward and inverse internal models representing sensorimotor transformations related to (A) one's own actions and (B) others' actions. Together, these models generate predictions that can facilitate (i) anticipatory control of a performer's movements and sounds; (ii) anticipation of coperformers' sounds and movements; and (iii) anticipation and partial control of the overall ensemble sound.

P. Keller, *Mental magery in music performance: underlying mechanisms and potential benefits*, Ann. N.Y. Acad. Sci. (2012)

Can imagined sound on "the private proscenium" traverse into echoic memory or are these completely differentiable presentations?

Edward Casey (*Imagining*, 1976) proposes qualities which lend themselves to scaled-up "armchair introspections" via laptops, networks, and a wide, eager pool of subjects.

Thank you to:

John Granzow, June Holtz, Pauline Oliveros, Andrea Halpern, David Huron, Jonathan Berger, Jieun Oh  
Andy Stuhl, Music 220 classes

good study, very interesting

This was a very interesting hit. It is one that really makes you listen and concentrate, thanks.  
It was very interesting to take this survey!

Is there any way to find out what this study is for? This was interesting to do.

This was an interesting task. I've never been asked to do anything quite like it before. This was fun to complete.

I found it interesting to not be able to imagine my voice coming from a spot outside myself or any voice saying mechanical Turk including that of my mother. I could hear her voice in my mind but could not imagine or hear her saying mechanical Turk. Odd and I wonder why that is.

interesting

very interesting study... a subject I've never thought about but makes complete sense

Definitely a different type survey. It was fun. Thanks for being creative!

IF IT IS MENTIONED TO SEE A PROMINENT OBJECT, ITS EASIER TO IMAGINE. LIKEWISE, ITS EASIER TO IMAGINE A PERSON IF IT IS MENTIONED TO IMAGINE A PERSON WHO IS CLOSE TO YOU.

This was a weird but interesting hit. I had never thought about this aspect of hearing or inner voices. I will work with it a bit and see what happens. I realized while doing it that depending on where my focus is, I can hear the sound/voice in different manners- in my head- my voice or a different voice, outside my head like someone else is speaking, in my head like someone else is speaking, in my head with the voice I usually hear speaking to me- whatever I think I can make physically happen.

I did my best, but some of the questions were not entirely clear to me.

As I began the task, I was able to imagine hearing my voice. Then as the task went on, it became more difficult to hear anything, because there was no sound.

This was interesting. I have very good hearing and often hear things other people don't. However, I wasn't very good at imagining a friend's voice or a sound coming from outside my head.

This was an unusual but very interesting study. It is almost like separating body and soul.

It's a fascinating topic and I found it a little difficult to describe in words, so having checkboxes was really helpful.

This HIT was very interesting. I enjoyed completing it. I heard and felt what "mechanical turk" sounds like inside my head and I was able to visualize what "mechanical turk" looks like with my eyes closed.

It never seems as if the sound is coming from outside my head.

Weird, but fun! It was easier to focus left and right when I was able to look in that direction.

I had hard a time imagining the voices being heard from either side for section 4 and 5. In fact, I had tell everyone to leave the room so I could have an easier time focusing. I was then able to complete the sections with some credible results.

It's an interesting experiment. I never realized that in order for me to imagine a sound coming from a particular direction my eyes needed to be pointed in that direction. Of course, it's not really something I tend to think about at all.

Very interesting stuff...

---



