

# Investigation of Instrumental Timbres in a 2-D Emotional Space

Song Hui Chon, Eric Derr,  
James Palmer, Jordan Smith

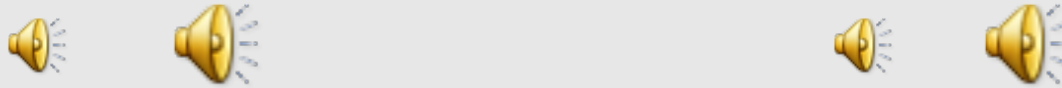
# Outline

- ❖ Introduction
- ❖ Historical Precedents
- ❖ Motivation
- ❖ Experiment
  - Interface
  - Hypotheses
  - Data Analysis/Interpretation
  - Possible Results
- ❖ Further Research/Conclusions

# Introduction

- ❖ MacDorman et. al (2007)
  - Emotion and loudness/pitch/rhythm
- ❖ Link between specific timbres and their propensity for expressing certain emotions
- ❖ TX802 timbre space.
  - Map results on timbre space

# Examples



- ❖ Interested in reaction to isolated timbre not entire passages
- ❖ Timbral qualities of notes in isolation may have an effect on emotional judgments made for an entire passage

# Historical Precedents

- ❖ Emotion and affect have been attached to music since music theory began
- ❖ Aristotle speaks of the Greek modes evoking certain emotions in *Politics* (c. 350BC):
  - Lydian: decorous, educative
  - Phrygian: enthusiastic, warlike
  - Mixolydian: sad, grave
  - Dorian: moderate, settled

# The Illusion of Instrument/Emotion Association

- ❖ Tunings and temperaments affected early instruments' sound
- ❖ Certain modes or keys could be strongly associated with certain instruments
- ❖ Composers in certain time periods or in certain types of music leaned toward certain instrument/emotion combinations

# Affects of modes and keys changed through time

❖ If even modes and keys depend on context, a certain emotion being attached to a certain instrument was surely an incidental or purely contextual phenomenon

# Biological motivation

❖ Strong links between emotional states and modes of vocal expression:

angry voice -> sharp attack, growling texture  
soothing voice -> soft attack, smooth texture



# Biological motivation

## ❖ Banse & Scherer 1996:

- people can recognize or discern the emotion of recorded speech, generally 4-5 times better than chance.
- acoustic correlates to emotion include pitch, intensity, spectral energy, and timing.

## ❖ Pierre-Yves 2003:

- good results for emotion detection based on timbral features of speech.

# Why Our Experiment?

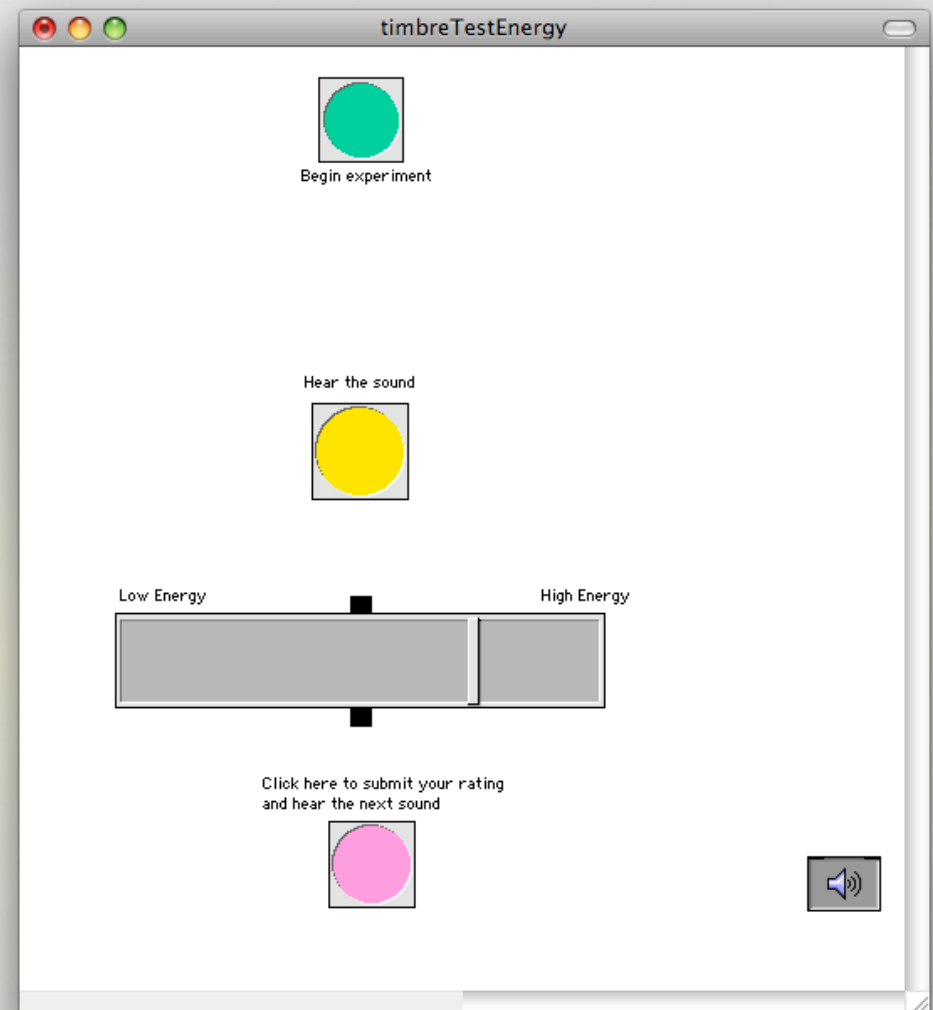
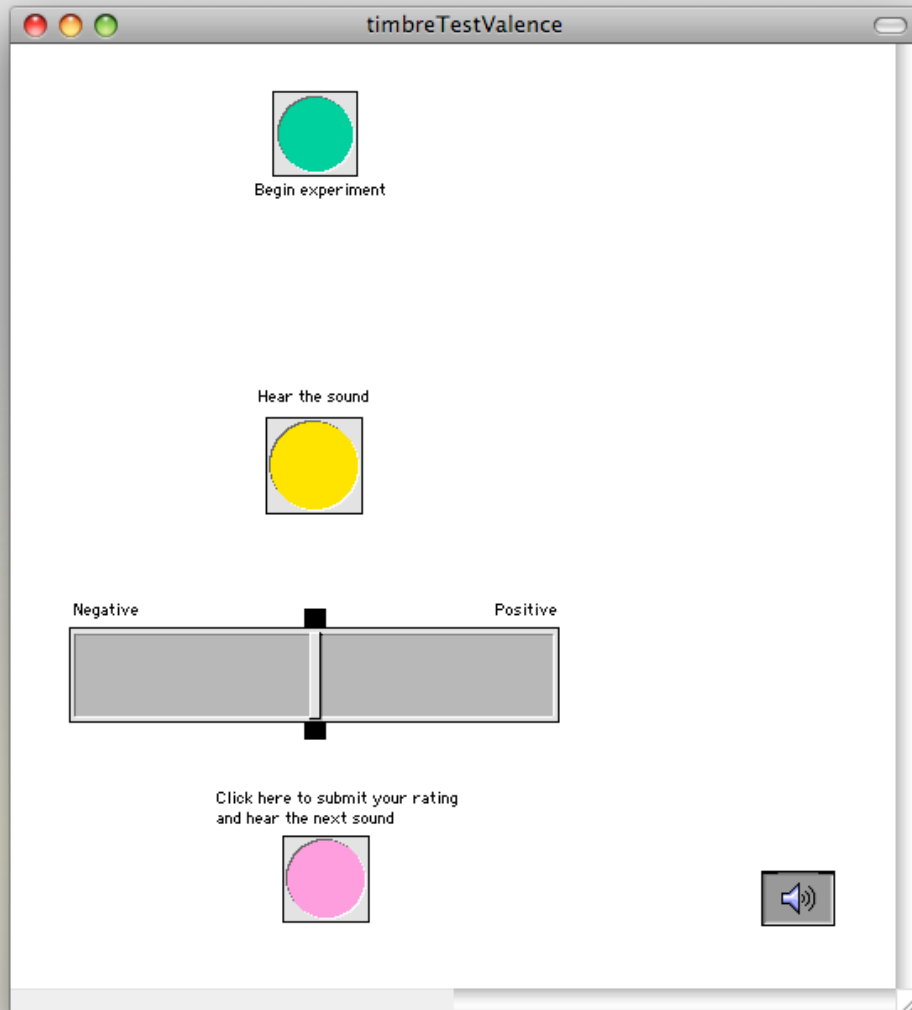
- ❖ There may be latent characteristics that give certain instruments a higher propensity for certain emotions
- ❖ We want to strip off the contextual and historiographical layers, as well as those of dynamics, pitch, rhythm, and pitch and timbral relationships
- ❖ Without immediately pinning down specific emotions, our experiment will check for similarities in increased or decreased energy and positive or negative valence

# The Experiment

- ❖ Using 2-D emotion model from Schubert 1996
  - Valence (Positive vs. Negative)
  - Energy (High vs. Low)
- ❖ Subjects
  - Group 1: Valence only
  - Group 2: Energy only
  - **Group 3: Both Valence and Energy**
- ❖ Stimuli
  - All timbre stimuli sets known so far
  - **21 TX-802 instrument tones including hybrids**
  - Different segments of sounds
    - attack only
    - steady-state only
    - **complete tone**

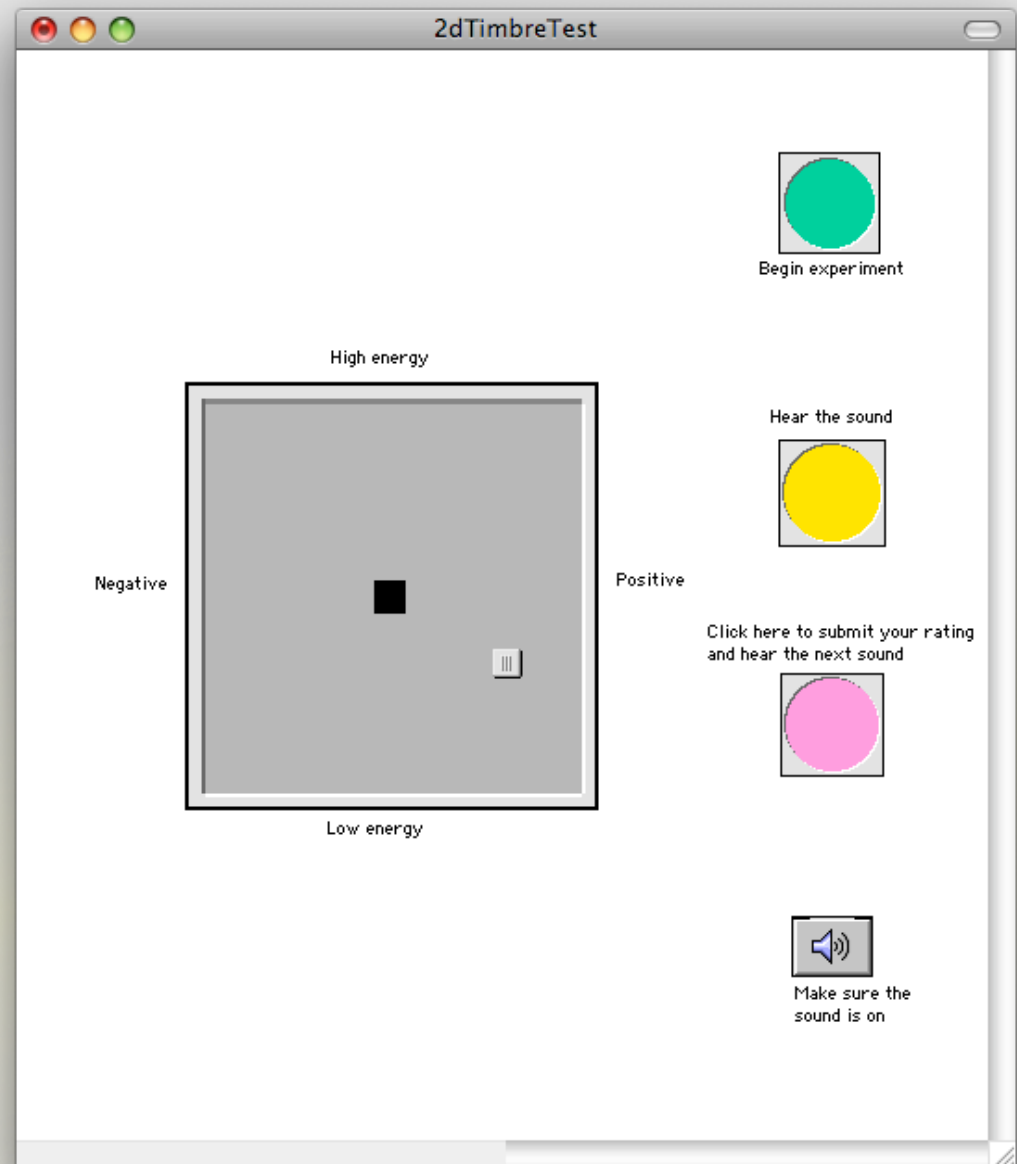
# Interface

Valence or energy judgments alone:



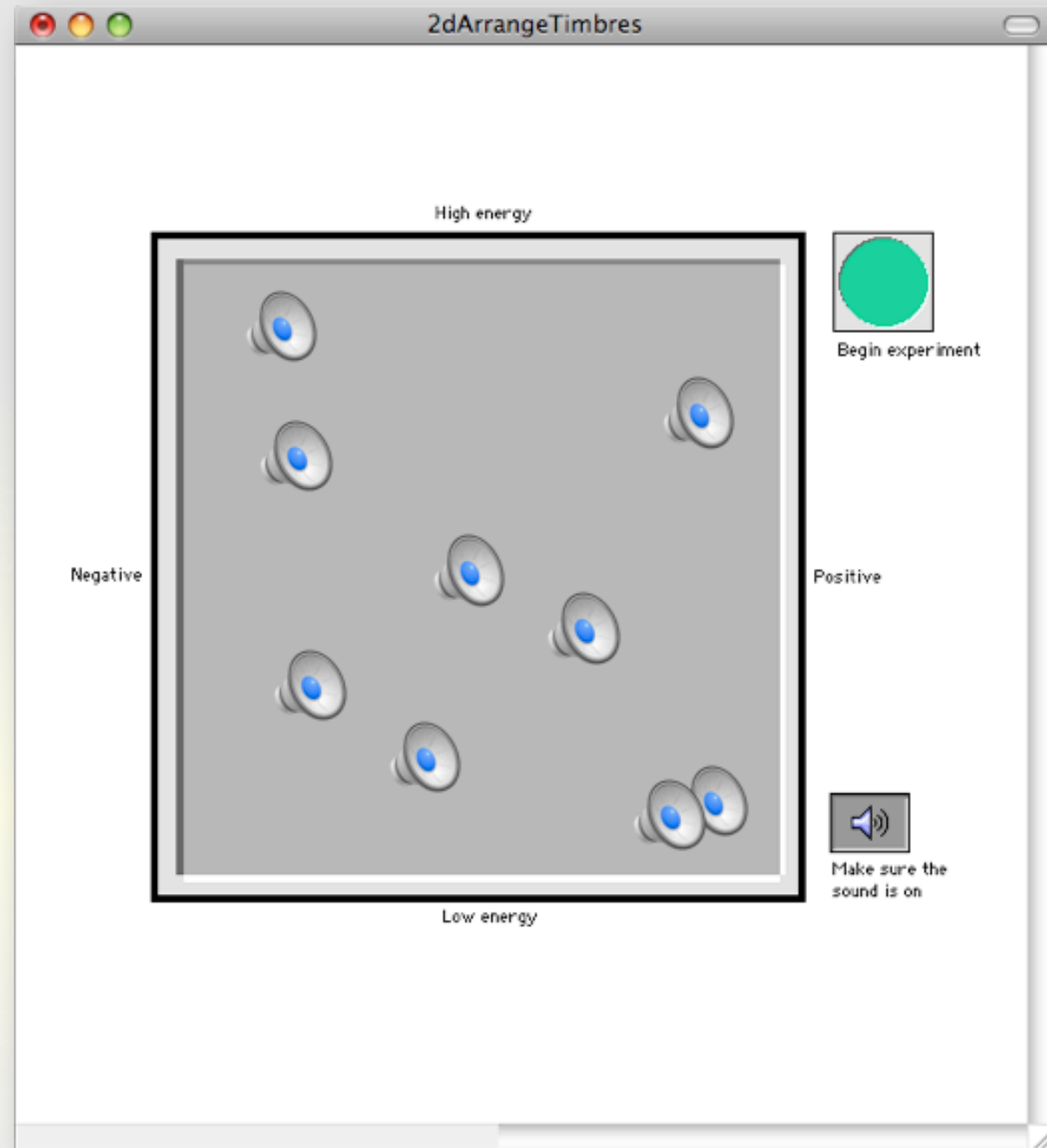
# Interface

Valence and energy judgments together:



# Interface

Proposed interface  
for arranging timbres:



# Hypotheses

1. There is no significant difference among male vs. female or musicians vs. non-musicians.
  - If difference among male vs. female
    - Investigate if frequency (or spectral centroid) dependent
  - If difference among musicians vs. non-musicians
    - Investigate if amount of experience is significant
    - Investigate if specialty (e.g., a brass player) is significant

# Hypotheses

2. The results will be consistent between

- Valence-only vs. Valence-Energy
- Energy-only vs. Valence-Energy
- Euclidean distance between Valence-only and Energy-only vs. Valence-Energy

❖ If not consistent, investigate if warped relationship exists



# Hypotheses

3. The portion of sounds presented is not a significant factor.

- Experiments with attack-only and steady-state portion only will exhibit similar behaviors as those with whole-tone sounds.
- The results with whole tones will be independent of sound duration (greater than 300 ms).

# Hypotheses

4. There will be a cloud of observations for one instrument sound (and not scattered everywhere in the 2-D emotion space) from which we can estimate the center of mass and standard deviation.
  - Centroids that are closer in distance may be mapped to instruments that are closer in 3-D timbre space.
  - Standard deviation of a stimulus cloud may reflect the "confusability" and may be reciprocal to timbral salience.

# Hypotheses

5. It is possible to map a distinctive emotion to an instrument tone.

- There will be a consistency in response regardless of gender, cultural background, and musical training.

# Hypotheses

6. There is a close relationship between the 2-D emotion space and the 3-D timbre space.

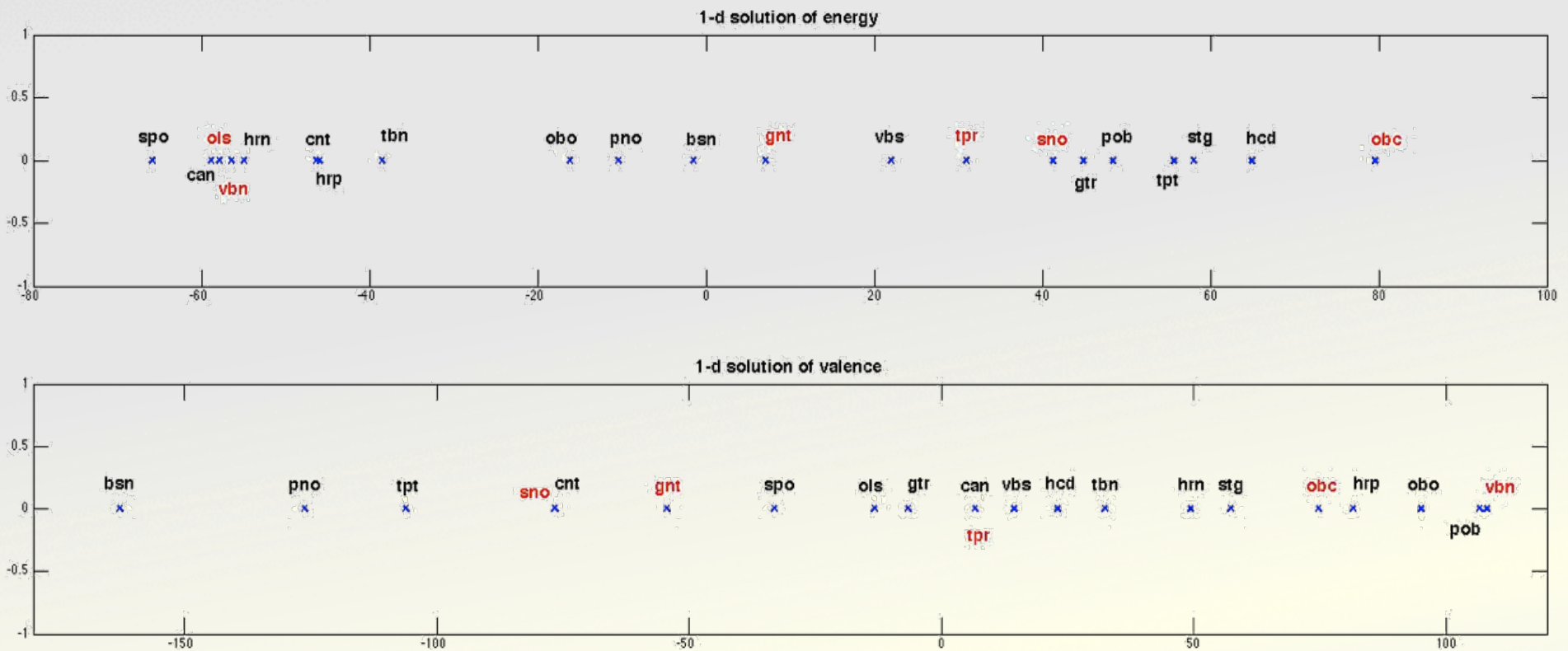
- Instruments that are closely located in the 2-D emotion space will be neighbours in the 3-D timbre space.
- Energy and/or Valence may be highly correlated with timbral dimensions (attack time, spectral centroid, spectral flux).

# Pilot Experiment

- ❖ 2-D (both Energy and Valence)
- ❖ 21 TX-802 tones including hybrids
- ❖ 4 subjects
- ❖ Data analyzed in 1-, 2-, 3-dimensional MDS methods
  - Energy-only
  - Valence-only
  - Euclidean distance of Energy and Valence



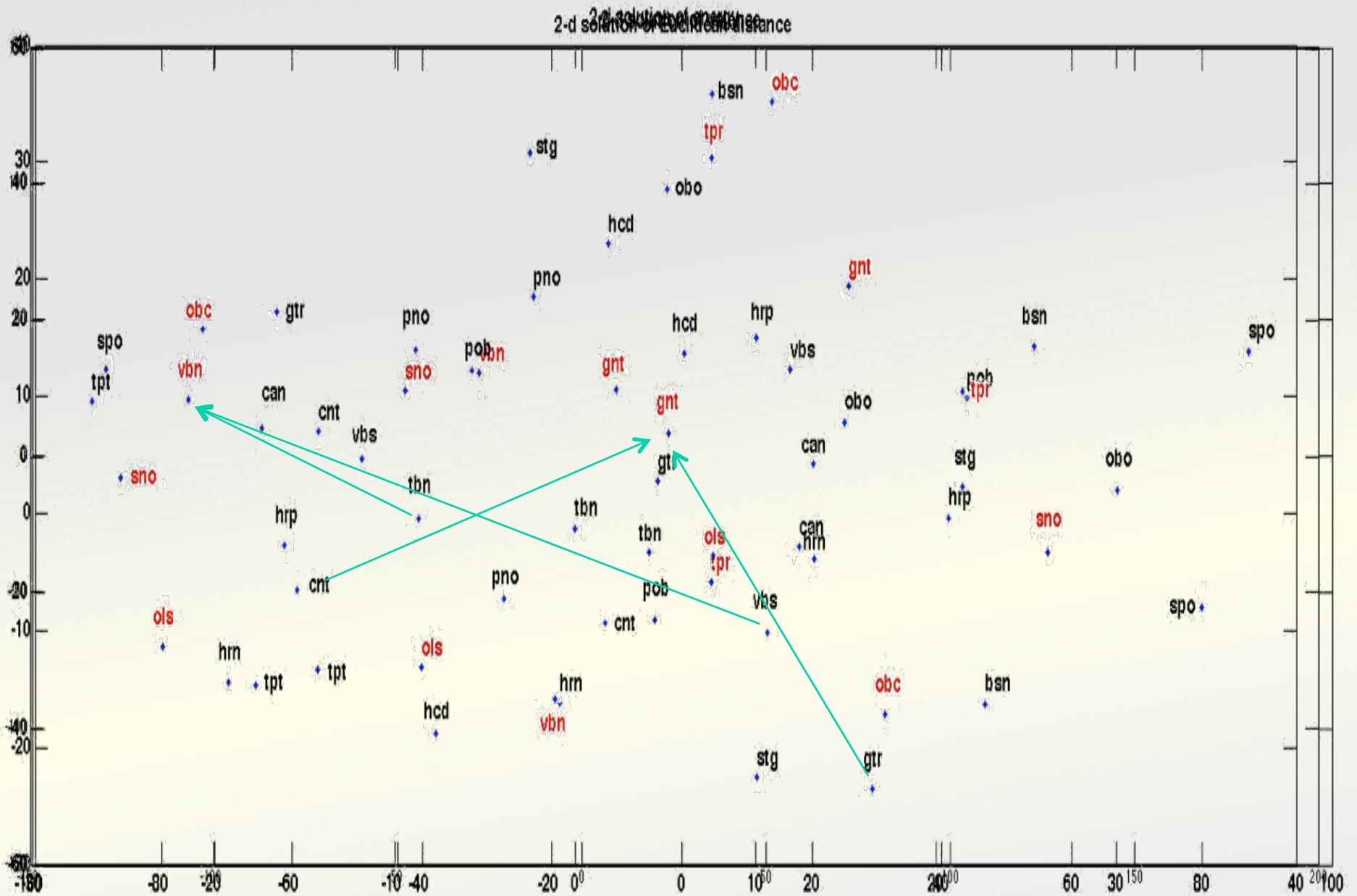
# Data Analysis 1-D



## Correlation coefficients

- Energy: 0.1873 **-0.6546** 0.3213
- Valence: 0.3091 0.1387 -0.2234

# Data Analysis 2-D

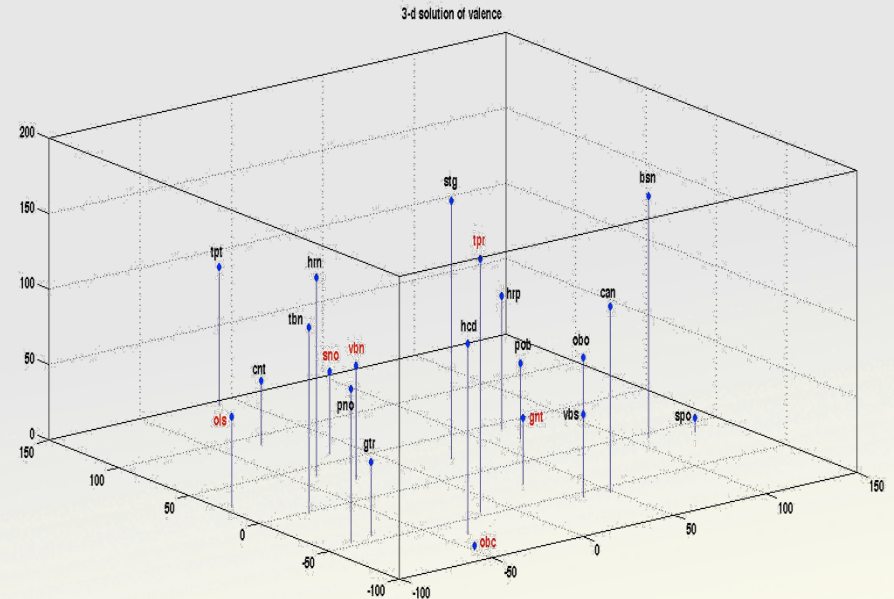
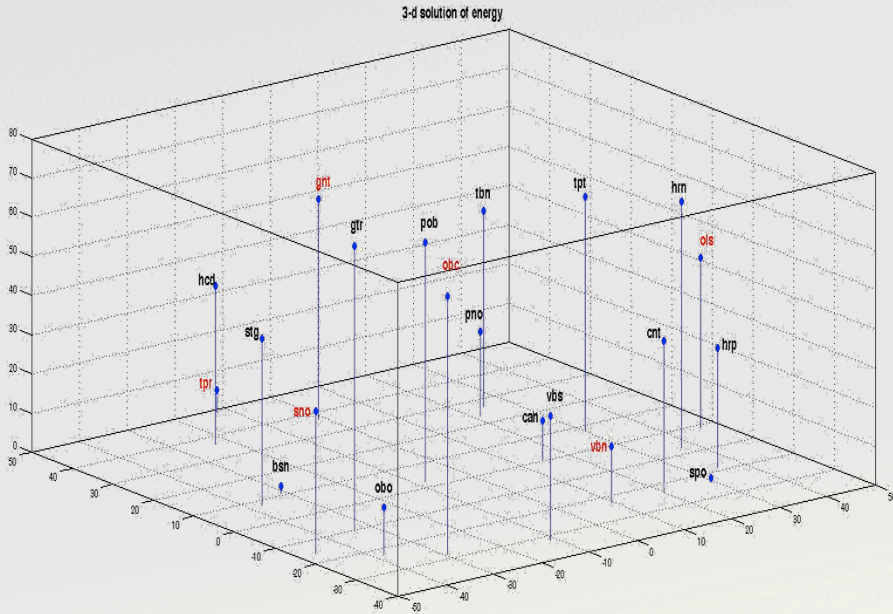






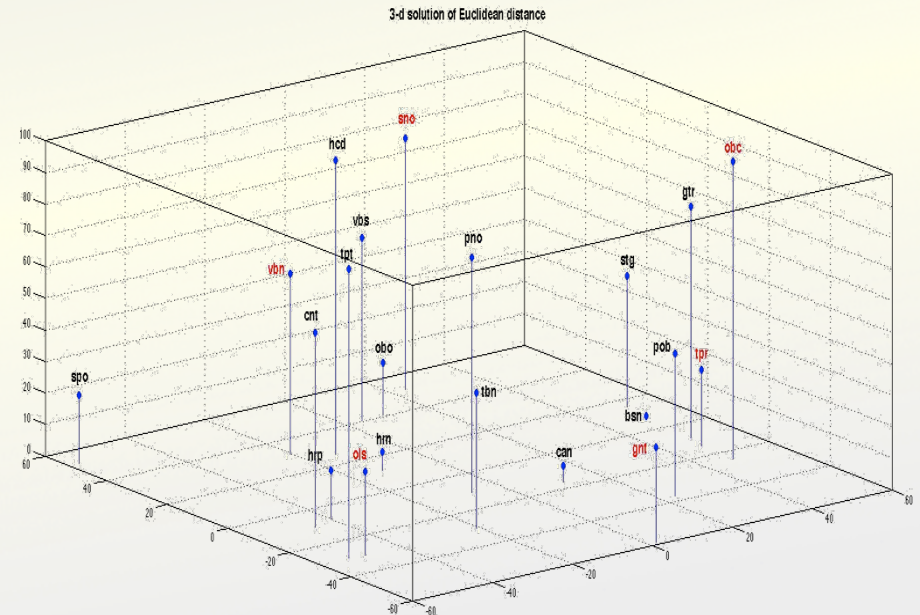


# Data Analysis 3-D



## ❖ Correlation coefficients

Energy:	-0.0898	<b>0.5352</b>	-0.2248
	-0.2103	-0.3473	0.1351
	0.0447	0.0993	0.2810
Valence:	-0.2218	-0.2215	-0.1432
	<b>-0.4683</b>	0.2116	0.1277
	-0.3762	-0.3345	-0.1785
Euclidean Distance:	-0.0204	<b>-0.5141</b>	-0.0028
	0.1485	-0.0417	0.0522
	0.3251	-0.1824	0.2717



# Data Interpretation

## 1) MDS analysis shows:

- Valence: 1-D similar to 2-D, which are slightly better than 3-D
- Energy: 1-D slightly worse than 2-D, which is slightly better than 3-D
- Euclidean-Distance: 2-D similar to 3-D

## Which means:

- No significant differences between 1-D, 2-D and 3-D solutions for Valence-only and Energy-only
- No significant differences among 2-D and 3-D solutions for Euclidean-distance observations

# Data Interpretation

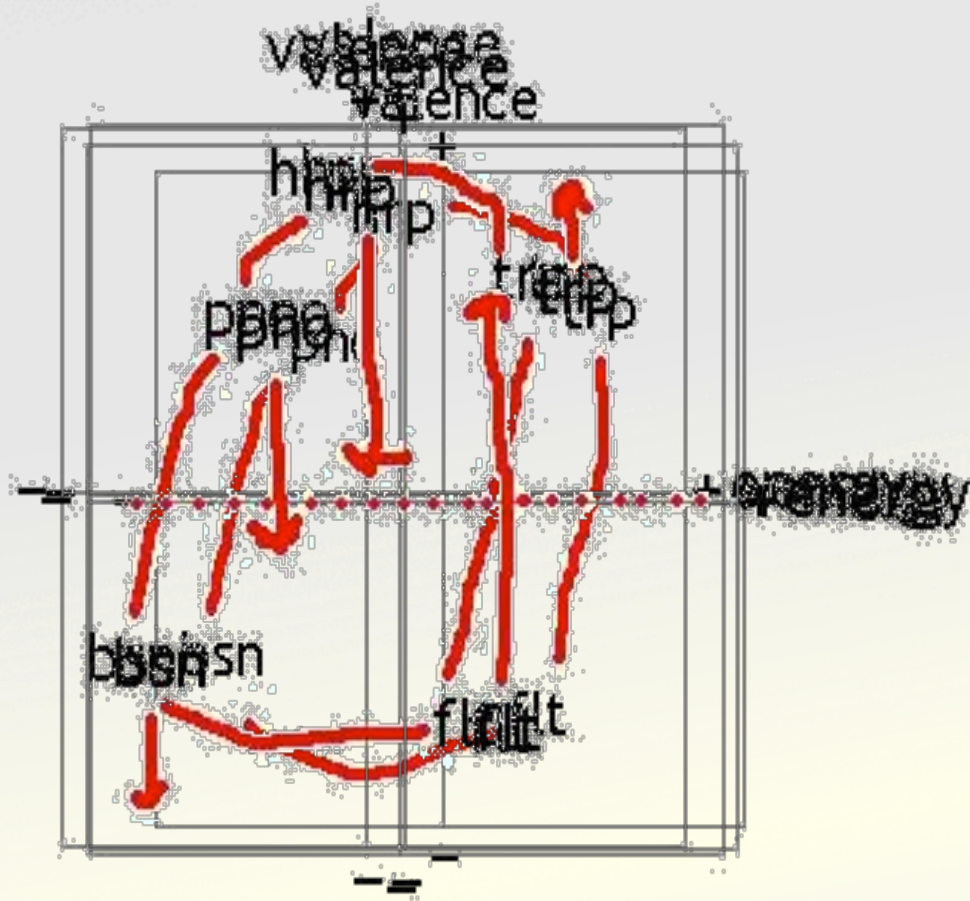
2) Correlation coefficients show:

- No significant correlations between timbre space dimensions and MDS-emotion dimensions
  - Exception: Energy and Spectral centroid

# Validity of Hypotheses

- ❖ Pilot study shows no significant correlations between MDS-emotion dimensions and timbre space dimensions.
  - But only 4 subjects were tested
  
- What can we expect with a million subjects for our study?

# Possible Results



# Possible Results

- 1) We will perform MDS on the respondents' data, which may show that:
  - valence ratings can be mapped in fewer than 3 dimensions
  - energy ratings can be mapped in fewer than 3 dimensions
  - both can be mapped in fewer than 3 dimensions
  
- ❖ Thus, for each axis in the emotion space, it may or may not be the case that respondents judge sounds consistently according to that criterion.



# Possible Results

2) Then we will compare these data to the known timbre spaces where we may find that:

- the valence axis of the emotional space maps to an axis in the timbre space
- the energy axis maps similarly
- both axes map to the timbre space

❖ If only one of the axes of the emotion space successfully maps to the timbre space, we could perform the experiment again with a substitute axis (e.g., "dominance")

# Further Avenues of Research

- ❖ Include percussion sounds
- ❖ Consider repeating this experiment on all musical notes to find trajectory of emotion space for entire range of an instrument
  - Does the emotional "labeling" change over the frequency range?
- ❖ Survey normative perceptions of instruments' sounds in memory (no sound stimulus)
  - Compare result with emotion judgment of the "most typical"
    - note (range)
    - loudness
    - extended technique

# Conclusions

How are we contributing to the general theory of timbre?

- Discover if latent emotional characteristics are inherent in various instrumental timbres
- Identify these characteristics
- Relate these findings to predetermined timbre spaces, thus trying to correlate valence and energy with other established timbral characteristics

# Bibliography

- Banse, R., and K. Scherer. 1996. Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*. 70(3): 614-36.
- Bigand, E., S. Vieillard, F. Madurell, J. Marozeau, and A. Dacquet. 2005. Multidimensional scaling of emotional responses to music: The effect of musical expertise and of the duration of the excerpts. *Cognition and Emotion* 19(8): 1113-39.
- Caclin, A., S. McAdams, B. Smith, and S. Winsberg. 2005. Acoustic correlates of timbre space dimensions: A confirmatory study using synthetic tones. *Journal of the Acoustical Society of America* 118(1): 471-82.
- Grey, J. M., and J. W. Gordon. 1978. Perceptual effects of spectral modifications on musical timbres. *Journal of the Acoustical Society of America* 63(5): 1493-500.
- McAdams, S., S. Winsberg, S. Donnadieu, G. De Soete, and J. Krimphoff. 1995. Perceptual Scaling of Synthesized Musical Timbres: Common Dimensions, Specificities, and Latent Subject Classes. *Psychological Research* 58(3): 177-192.
- Pierre-Yves, O. 2003. The production and recognition of emotions in speech: features and algorithms. *International Journal of Human-Computer Studies*. 59: 157-83.
- Schubert, E. 1996. Continuous Response to Music Using the Two Dimensional Emotion Space. *Proceedings of the 4th International Conference on Music, Perception and Cognition*. 269-274.

# Discussion Questions

- ❖ Do you think instruments carry inherent emotional information?
  - Would it be related to the timbre of the instrument?
  - Would it affect (or be affected by) the mood of a piece of music?
  
- ❖ Would the emotional characteristics be additive?
  - For example, if a harpsichord is consistently judged "high" in valence and a cello is consistently judged "low" in valence, would a harpsichord-cello fusion result in a relatively neutral valence?