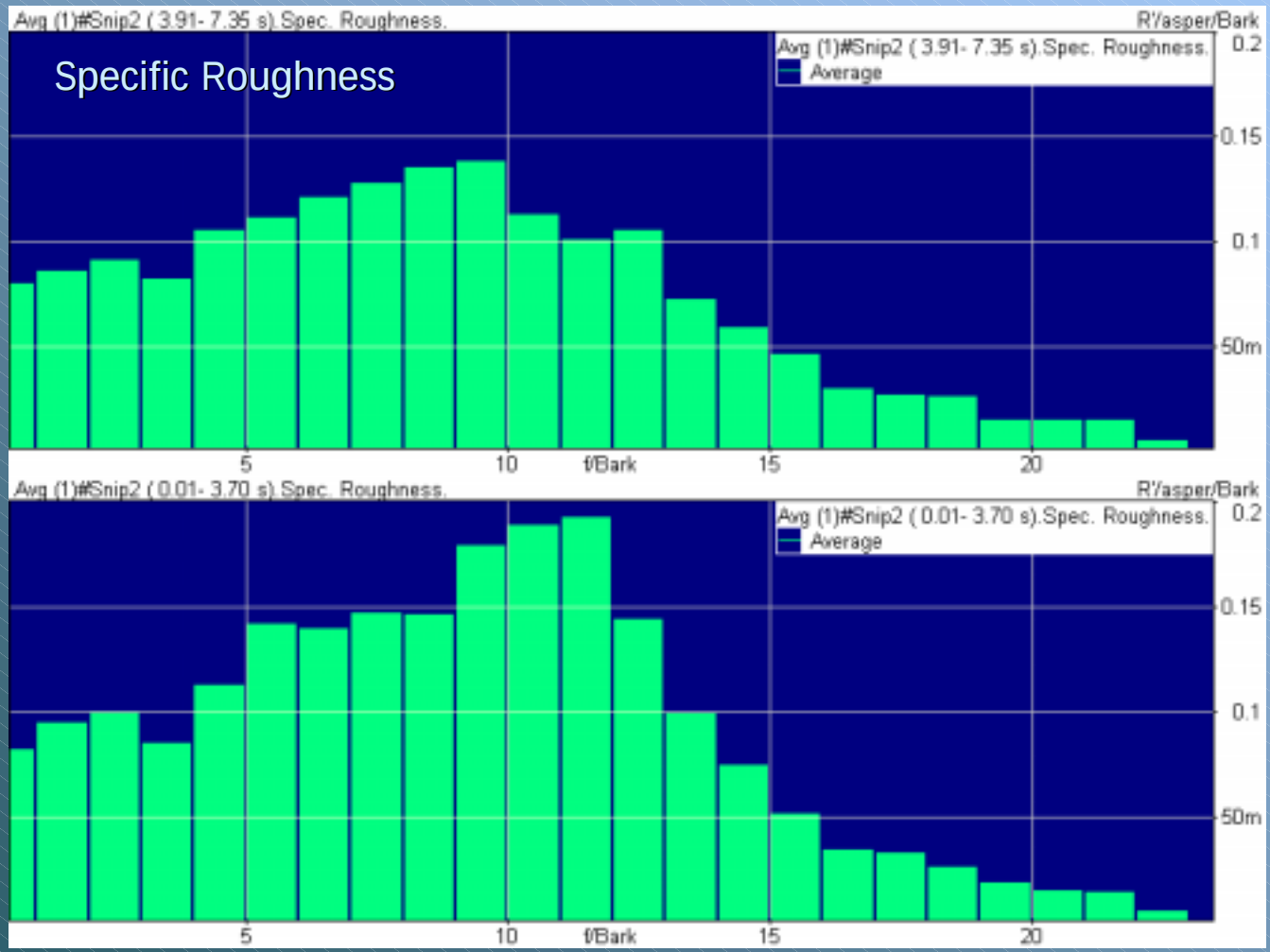
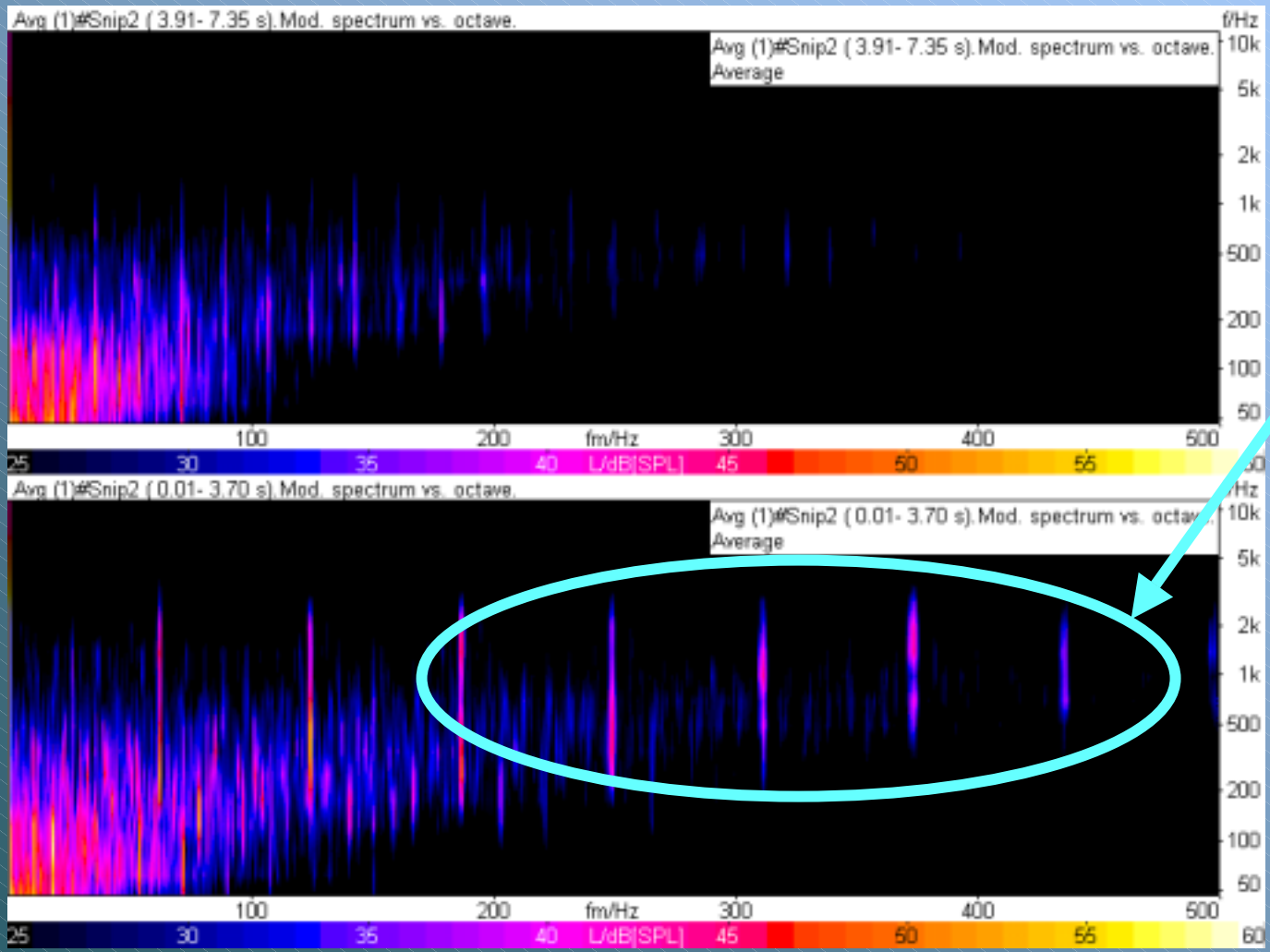
HVAC Compressor: Roughness?



As one would expect with so much amplitude modulation, the roughness psychoacoustic metric also shows a difference between these two sounds.

HVAC Compressor: Roughness?



None of this modulation content is considered by the roughness calculation.

However, this modulation content plays an important role in the timbre and quality of the sound.

Subjective Evaluation

- General methods
  - Threshold (will skip due to time constraints)
  - Supra-threshold (“Above threshold”)
- Determine test method
- Sound preparation
- Acquiring subjective data
- Analysis of subjective data
- Listening study demo

Reference:

N. Otto, S. Amman, C. Eaton, S. Lake,
“Guidelines for Jury Evaluations Of Automotive Sounds,”
1999 SAE N&V Conference 1999-01-1822, Traverse City, MI

Semantic Differential (SD) / Categorical Rating (CR)

- **SD:** scale defined by a pair of polar adjectives

QUIET **LOUD**

Extremely	Very	Somewhat	Neither	Somewhat	Very	Extremely
1	2	3	4	5	6	7

- **CR:** used when there's not a good pair of polar adjectives

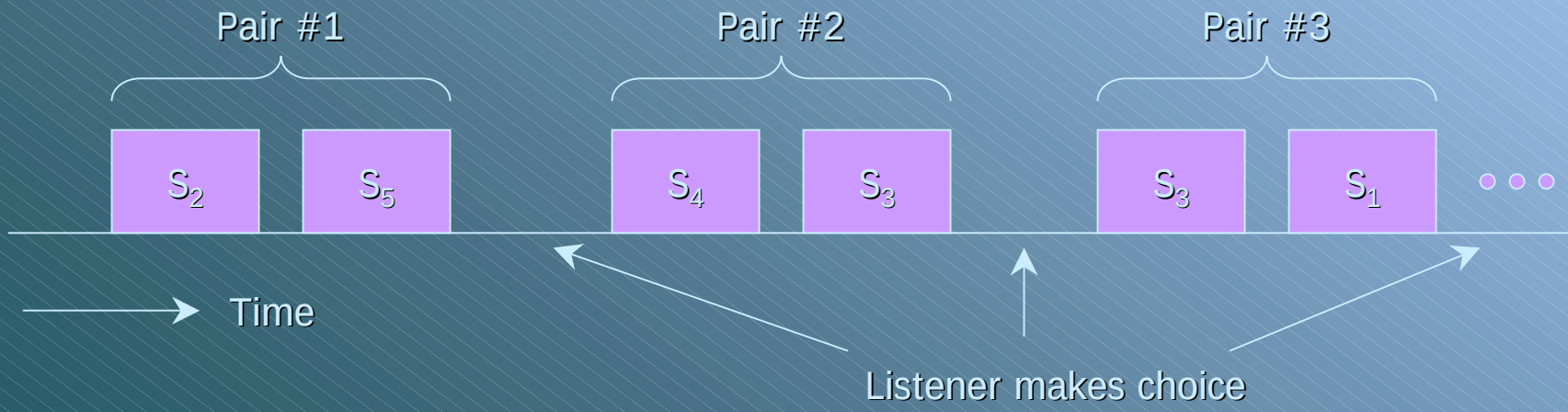
NOT SQUEAKING **SQUEAKING**

Extremely	Very	Somewhat	Neither	Somewhat	Very	Extremely
1	2	3	4	5	6	7

- Play each sound once and have listeners rate on M-point scale (e.g., M=7 above)
- Give verbal description with numbers
  - Listeners have more difficulty making absolute judgments
- Caveat: Scale/Range usage by different listeners

Paired-Comparison

- N sounds ⇒ N(N-1)/2 pairs ⇒ N(N-1) pairs with repeats
- Listener makes choice after each pair of sounds
  - Examples: “More Preferred”, “More Loud”
- Get rank order data from scores
  - score = number of times a sound is chosen over all pairs
- Bradley-Terry or Thurstone methods to convert pair probabilities to linear scale (“Merit Values”)
- Caveat: Provides only relative ratings



General Methods – Supra Threshold

See references for details on other methods

- Stevens' Methods
  - Magnitude Estimation/Scaling
  - Magnitude Production
- Rank Order
- Multi-Dimensional Scaling

Determine Test Method

Usually determined by

- What's the desired result? (Relative ratings, Absolute ratings, Rank order, ...)
- Knowledge about which characteristics vary in your sounds

Loudness **Roughness**

Fluctuation

Impulsiveness **Sharpness**

Other Characteristics ...

Example: SQ of Power Seats

Problem:

- Customer satisfaction issue with loudness and modulation of power seat motor sounds.

Goal:

- Determine SQ metrics that describe subjective **preferences** in a *relative* sense
- Determine levels of **acceptability** in an *absolute* sense

Determine Test Method:

- Paired-comparison of preference to get relative ratings
- Categorical Ratings to get acceptance in absolute sense

Note: Previous study shows correlation of N50 and Fluctuation Strength to perceived loudness and fluctuation, respectively

Sound Preparation

- Obtain recordings/sounds that span the range of the effect you're trying to measure
- Sounds should be free of "distractions"
  - Noise from vehicle instrumentation
  - Unwanted squeaks, rattles from recording environment
  - Talking/Breathing/Moving of person making the recording
  - Other undesired characteristics that are not being investigated
- Duration of sounds
  - Long enough to capture the "event"
  - Sounds should be same duration, if duration is not a factor
- Apply onset/offset (fade in and out) to each sound to prevent "popping" on playback (25-50 msec is typical)

Check Total Test Time

- Maximum test duration should be 30-45 minutes
  - Longer tests result in fatigue/vigilance issues => poor subjective results
- Determine total test time given the following:
  - Number of sounds
  - Duration of sounds
  - Types of tests
  - Time for study description/introduction/context
- If duration is too long, consider the following:
  - Reduce number of sounds
  - Reduce duration of sounds
  - Split study into two separate studies

Example: SQ of Power Seats

- Obtain recordings/sounds that span the range of the effect you're trying to measure

- Sounds should be free of "extraneous events"
  - Sound contains clicks:  Clicks removed: 
  - Grinding characteristic (very different from other sounds): 
- Duration of sounds:
  - 2.4 seconds is long enough to capture modulation effect
- Apply onset/offset:
  - 25 msec

Example: SQ of Power Seats

Check Test Time

- Test description/introduction/context: **~10 min**
- Paired Comparison (with repeat)
  - N=7 sounds ⇒ N(N-1) = 42 pairs
  - Sound duration: 2.4 sec
  - Inter-stimulus interval: 0.5 sec
  - Inter-pair interval: 5 sec
  - Total: [42*((2*2.4)+0.5) + 41*5]/60 = **7.12 min**
- Categorical Rating of "Acceptance"
  - Introduction of Categorical Rating process: **~3 min**
  - Play each sound and allow 5 seconds to rate
  - Total: [7*(2.4+5)]/60 = **0.86 min**
- Total Test Time: **~21 min**

Acquiring Subjective Data

- Demographics of listeners
  - Gender, Age
  - Vehicle Segment (e.g., luxury, truck, C-class)
  - Expert vs. Non-expert Evaluators
- Number of listeners
  - Paired-comparison, Categorical Ratings
    - 30-60 when using “company employees”
    - 75-100 when using customers
- Listening environment
- Description of test
- Provide context, if needed
- Provide feedback about progress

Example: SQ of Power Seats

- Demographics
  - Combination of expert and non-expert evaluators
- Number of listeners
  - Used “company employees”: 75 evaluators
- Listening environment
  - Quiet listening room
  - Sounds played back over headphones
- Description of test
  - Evaluating SQ of power seat sounds
  - Recordings made in driver’s seat of the vehicle
- Provide feedback about progress
  - Paired-comparison: “Pair X of 42”

Analysis Of Subjective Data

Paired-Comparison (PC)

- Repeatability
  - Consider eliminating listeners with repeatability less than 60%
- Consistency
  - Evaluate all triads of sounds (A,B,C)
  - If A>B and B>C, then consistent if A>C
  - Consider eliminating listeners with consistency less than 60%
- Convert pair probabilities to scale data
  - Bradley-Terry or Thurstonian Models
- Re-predict pair probabilities from Scale Data and correlate to actual pair probabilities

Analysis Of Subjective Data

Semantic Differential (SD)/Categorical Ratings (CR)

- Standard Statistics: Mean, Std. Dev., Confidence Intervals, etc.

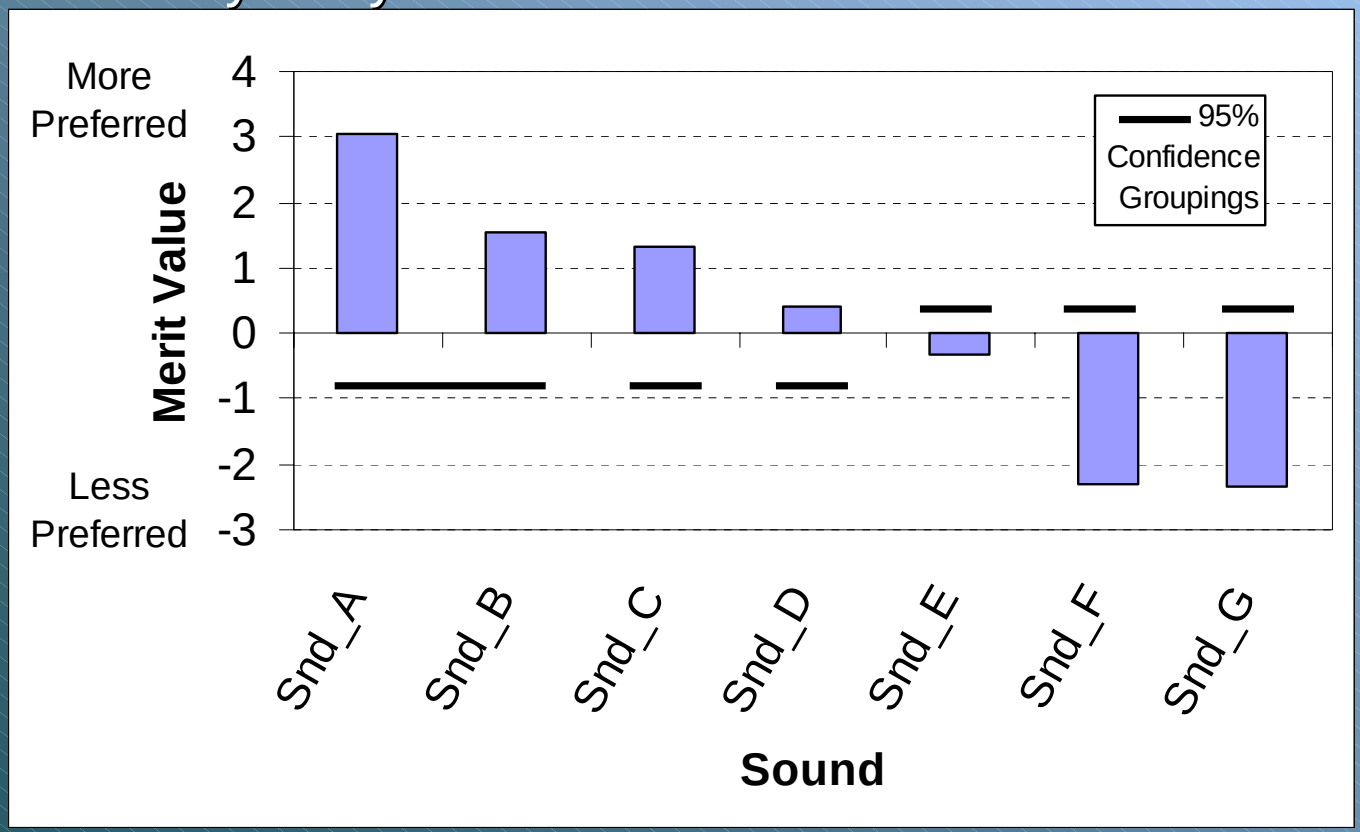
Fit of SD/CR to Paired-Comparison

- Multi-variable linear regression
- Often used when PC is based on “preference” and SD/CR is based on loudness, sharpness, etc.
- Can indicate which dimensions make up the preference

Example: SQ of Power Seats

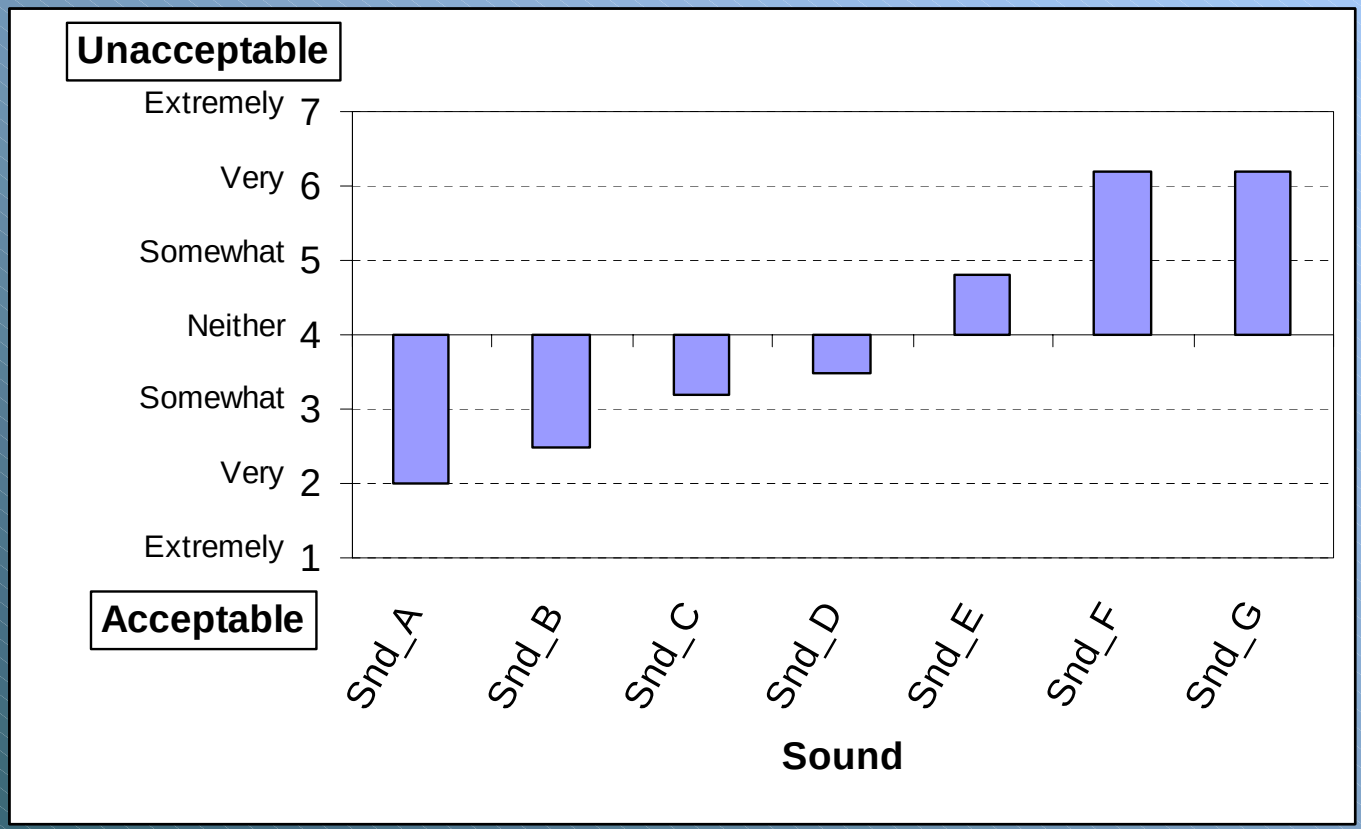
Paired-Comparison

- 3 listeners removed for poor repeatability/consistency
- Average Repeatability: 83%
- Average Consistency: 94%
- Bradley-Terry Model: $R^2 = 0.97$



Example: SQ of Power Seats

Average Acceptance Ratings



Subjective – Objective Correlation

Goal: Replace subjective testing with a highly correlated objective measure of the sound

- Linear Regression

$$y = A_0 + A_1 x_1 + A_2 x_2 + \mathrm{L} + \varepsilon$$

  - y: Merit Value, Average SD/CR Rating, ...
  - A_i:Coefficients from linear regression
  - x_i:Objective measures (loudness, sharpness, ...)
  - ε: Error in prediction
- Caveats/Issues
  - Use Common Sense – even when you get a large R^2 fit
    - Do the objective measures make sense?
    - Do the +/- signs of the coefficients make sense?
  - Number of objective measures vs. number of sounds
  - Poor R^2 fit for SD/CR ratings because of scale use by listeners

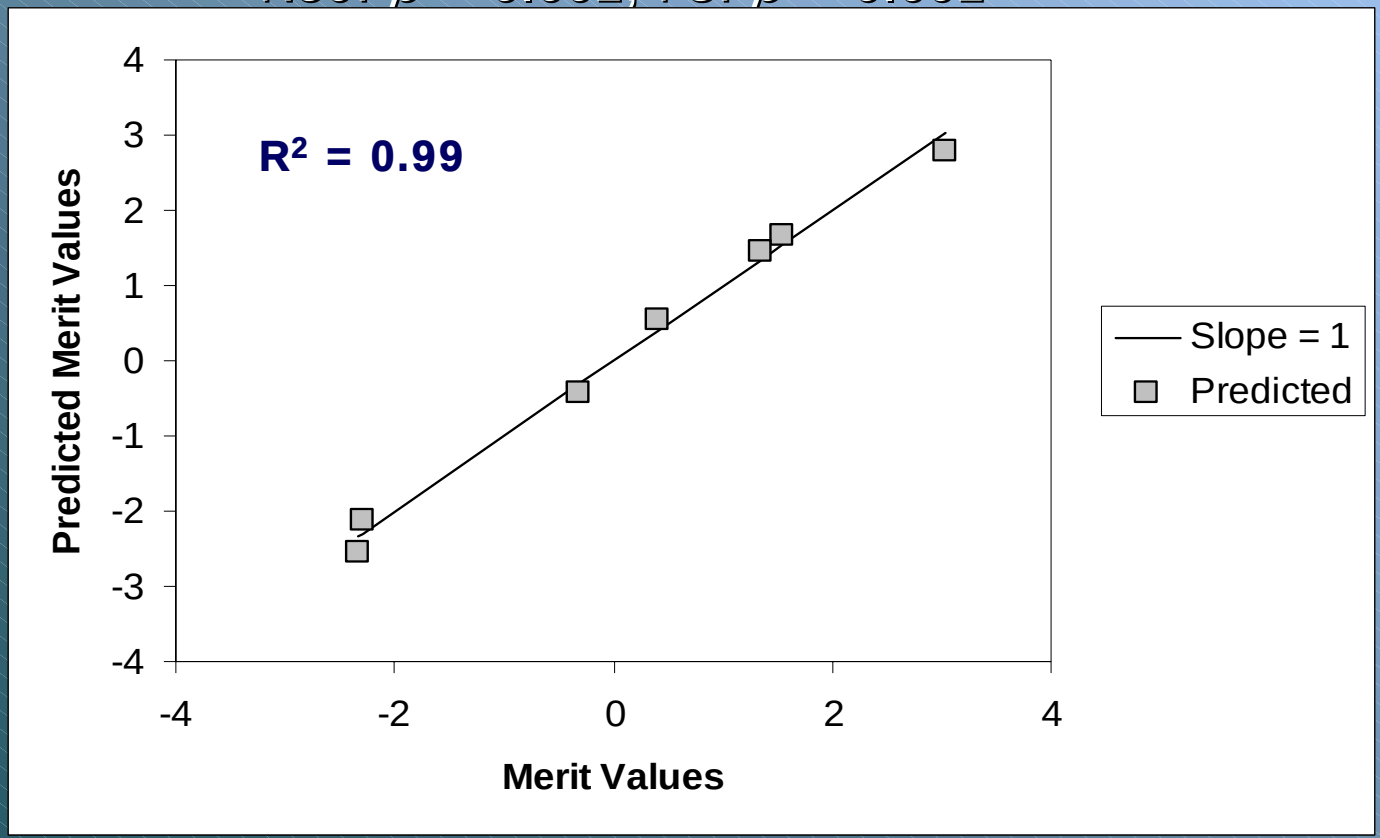
Example: SQ of Power Seats

Predicted Merit = $A_0 + A_1*(N50) + A_2*(FS)$

- A_1, A_2 are negative
- Statistical Significance
  - N50: $p = 0.001$; FS: $p < 0.001$

Avg. Fluct. Strength

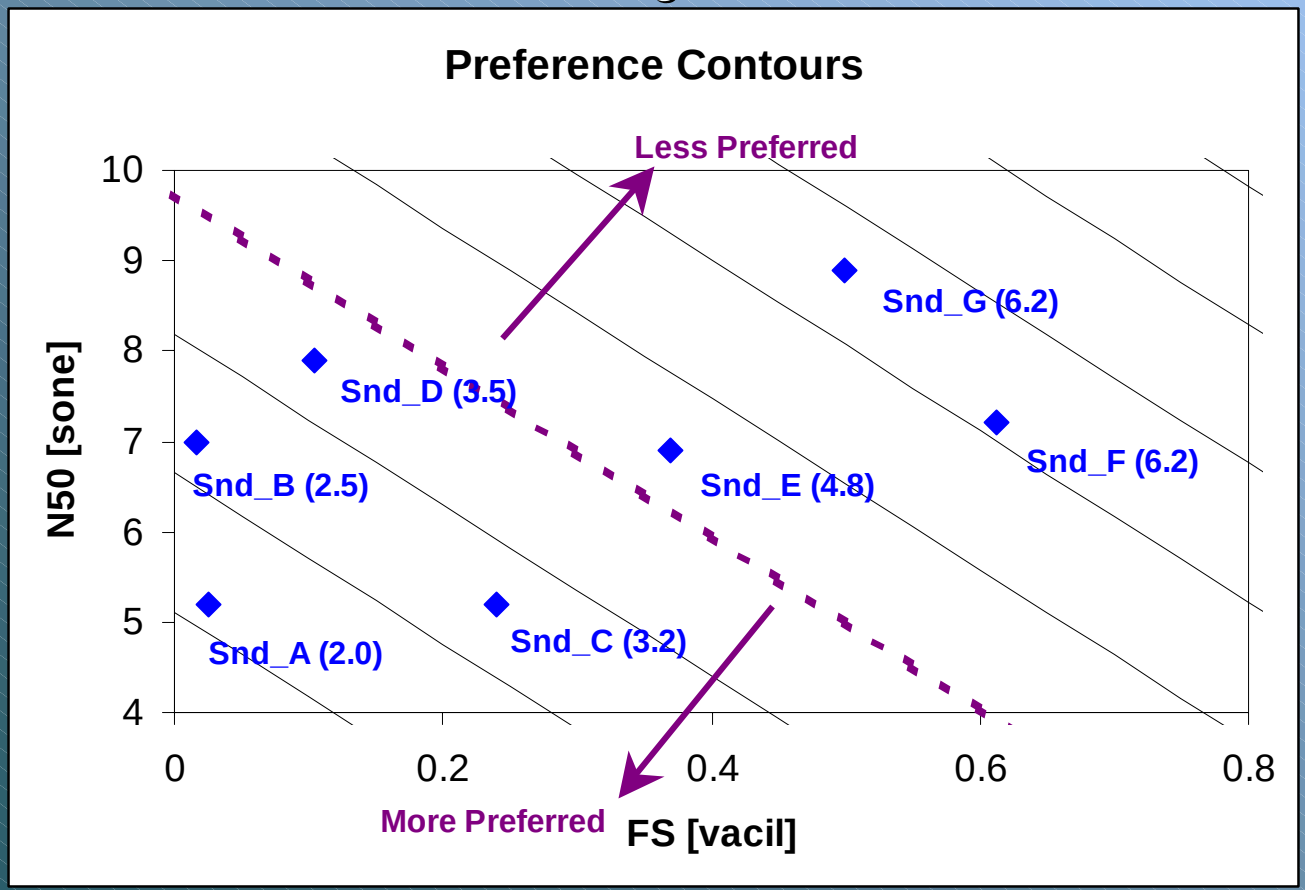
Median Loudness



Example: SQ of Power Seats

Set Acceptance levels for N50, FS

- Diagonal Lines: lines of constant merit value
- Numbers in parenthesis: Avg. Acceptance Ratings
- Use these results to set targets for N50, FS



Conclusion

Sound engineering has become very important.

- A lot of automotive products are becoming more refined.
- Product differentiation is no longer just about “being quietest.”

Sound quality cannot be measured solely with conventional measurement technique.

Artificial Head, psychoacoustics, conventional metrics, multi-channel analysis, all guided by listening and employing techniques like filtering, are helpful tools for sound engineering…

Listening studies provide the “voice of the customer” to validate subjective/objective correlation.

But:

For sound design we need the human factor, the expert to create a good sound.

Questions?

Contact Information

Wade Bray, HEAD acoustics, Inc., Brighton, MI
Telephone: (248) 486-0099 ext. 207
e-mail: wbray@headacoustics.com

Mike Blommer, Ford Motor Company. Dearborn, MI
Telephone: (313) 621-8197
e-mail: mblommer@ford.com

Scott Lake, General Motors Corp. Milford Proving Ground, MI
Telephone: (248) 685-4163
e-mail: scott.a.lake@gm.com

Tools & Methods - References

Psychophysics and Psychoacoustics:

- *Psychophysics* - S. S. Stevens; ISBN: 0887386431
- *Psychoacoustics: Facts and Models* - Zwicker & Fastl; ISBN: 3540650636
- *Hearing* - Moore, B.C.J (Editor); ISBN: 0125056265
- *Spatial Hearing* - Blauert, J.; ISBN 0262024136
- Journals: *Acustica*, *Journal of ASA*
- CD: *Auditory Demonstrations*, available from Acoustical Society of America (ASA) website at: http://asa.aip.org/discs.html

Subjective Evaluation

- "Guidelines for Jury Evaluations Of Automotive Sounds" N. Otto, S. Amman, C. Eaton, S. Lake, *1999 SAE N&V Conf. 1999-01-1822*, Traverse City, MI
- *The Method of Paired Comparisons*, David, H. A.; ISBN: 0195206169

Tools & Methods - References

Signal Processing:

- *Digital Signal Processing* - Proakis and Manolakis; ISBN: 0133737624

Statistics for Subjective Evaluation:

- www.statsoft.com / Electronic Statistics Handbook
- Quantitative Applications in the Social Sciences: Sage Publications
  - *Magnitude Estimation, Factor Analysis, Multi-dimensional Scaling*, etc.
- *Introduction to Regression Analysis*: Montgomery and Peck, ISBN: 0471533874

Of Related Interest:

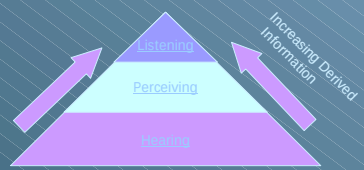
- *Music, the Brain & Ecstasy* – Jourdain, R.; ISBN: 038078209X

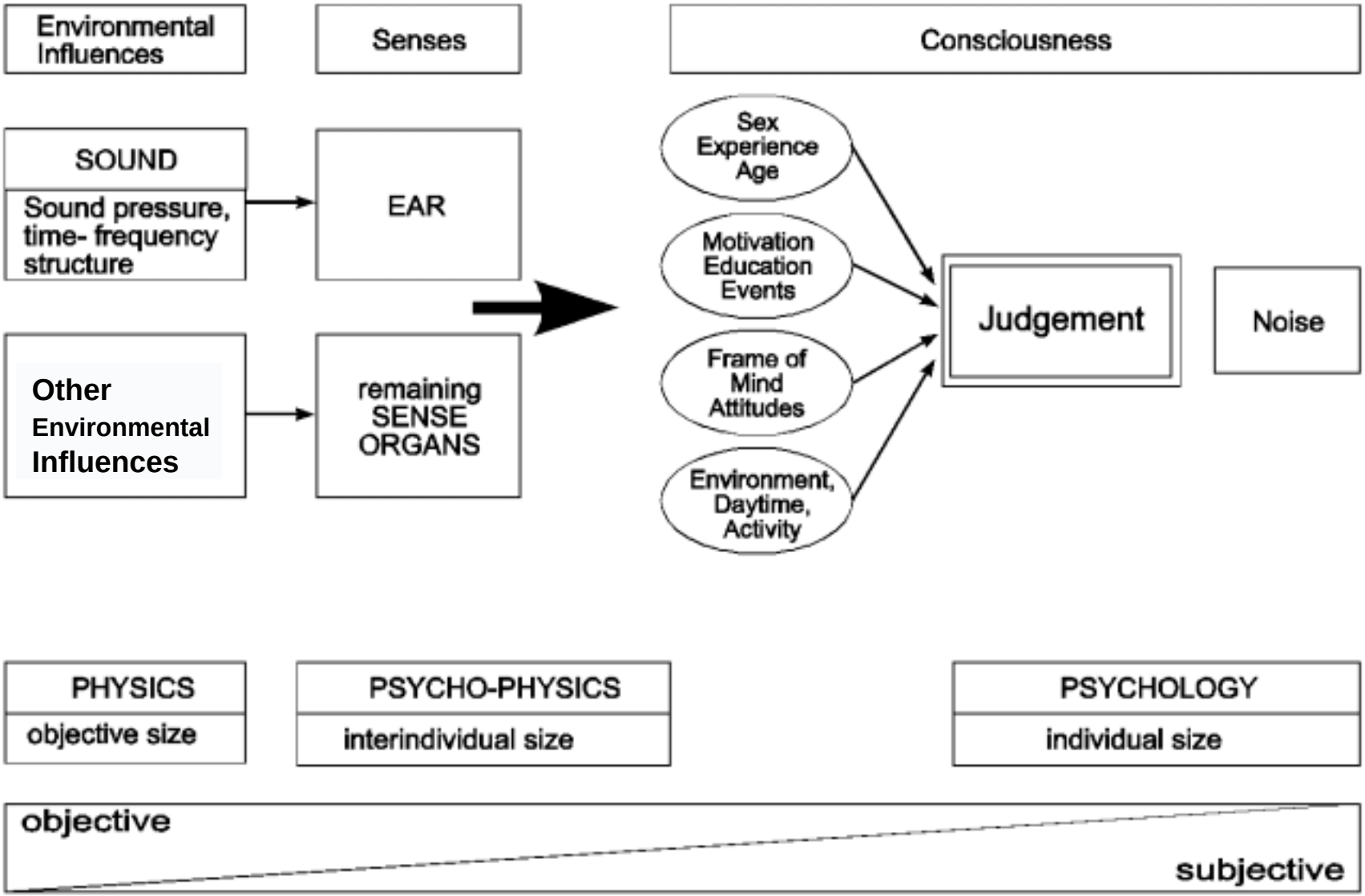
Listening - Definition

Listening goes one step further than perception, introducing our expectations of the information as part of the processing of the information.

Demonstration of two examples:

- Famous Quotation
- What are you hearing?





No <u>one</u> measure can satisfy this continuum.