Thurstone's law of comparative judgment applied to 5-point verbal annoyance responses from community noise studies

Matthew Boucher

NASA Langley Research Center

May 22, 2025

188th Meeting of the Acoustical Society of America joint with

25th International Congress on Acoustics

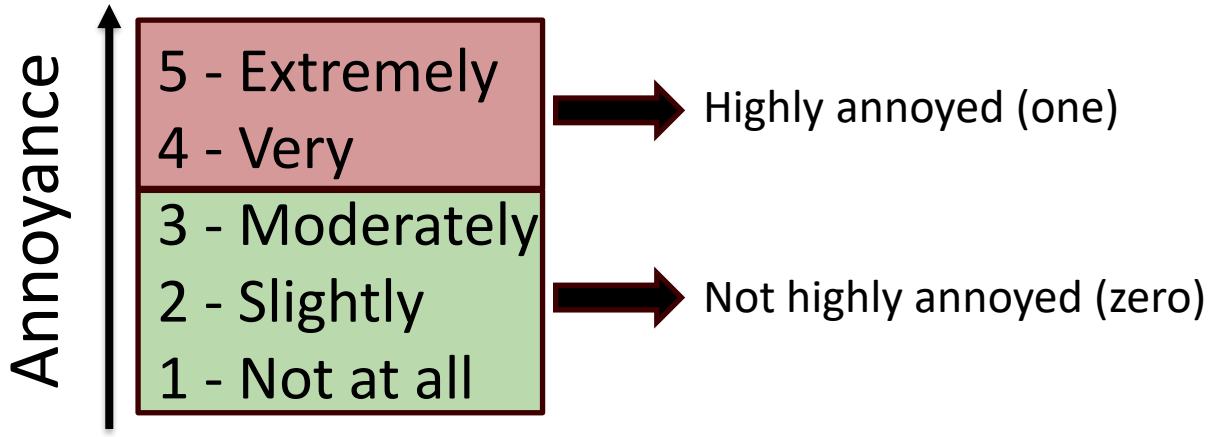
(Session 4pNS, Exposure Response and Community Tolerance Level)

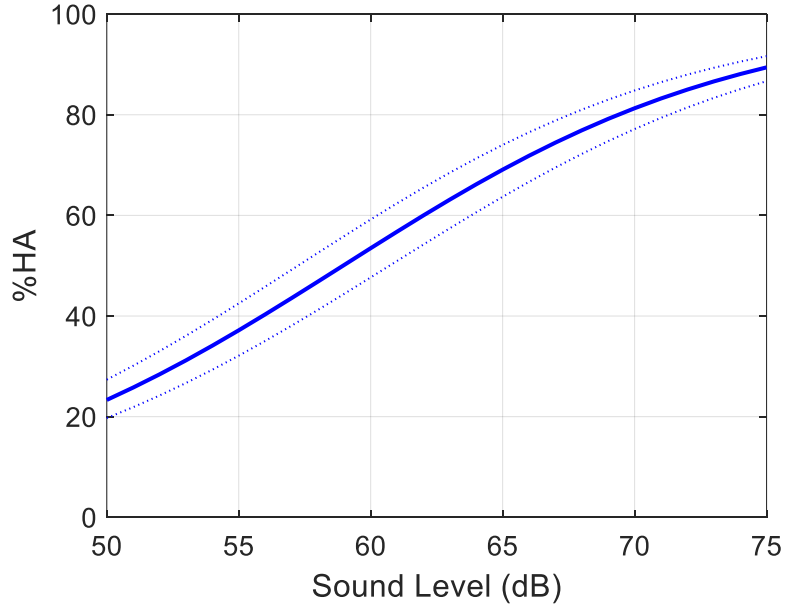
Overview

- Introduction
- Thurstone analysis and Law of Comparative Judgment
- Simulated community annoyance dataset
- Results

Perceptual response (dose-response) modeling

Common approach





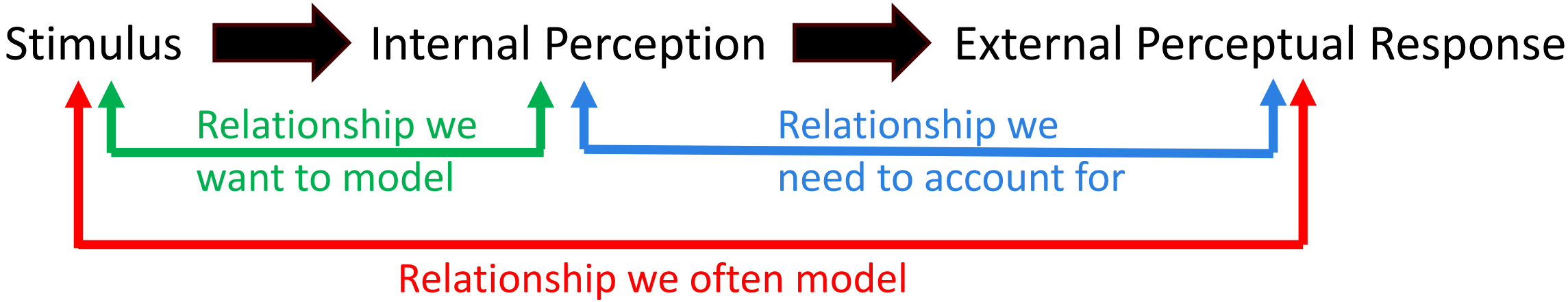
Potential limitations

- Dichotomization of 5-point responses
  - Highly annoyed or not
  - Discrete annoyance scale
- Equally spaced categories not guaranteed (ISO/TS 15666)

Motivation to use Thurstone analysis

- Continuous annoyance scale
- Just-noticeable-differences
- Evaluate spacing of response categories
- Internal annoyance perception

Focus on stimulus-to-internal perception relationship



- **Problem: people use response scales in different ways**[1]
  - Perception is an internal judgment
  - People differ in how they report internal judgments. They respond:
    - Differently for the same judgment
    - The same to different judgments

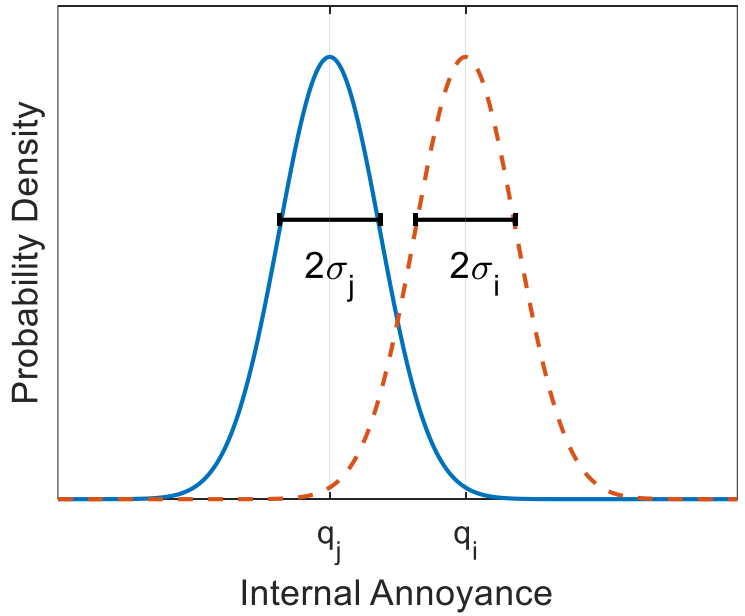
- **Possible solutions**
  - Multilevel analysis
  - Master scaling (Berglund, 1991)
  - Thurstone's Law of Comparative Judgment (1927)

[1] G.A. Gescheider, "Psychophysics", 1997

Basics of Thurstone's Law of Comparative Judgment[2]

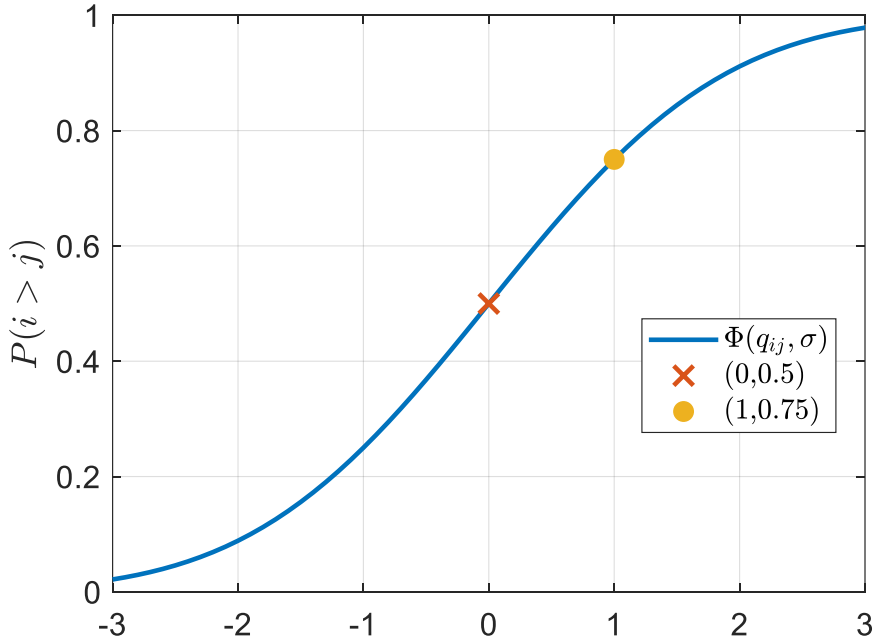
1. **Build preference matrix [C]**
   - **Entries c_{ij}: number of times sound i judged more annoying than sound j**
2. **Trial values for consensus annoyance opinion, q_i**
3. **Maximize probability**
4. **Output is annoyance value for each stimulus (Thurstone scale)**

Which sound was more annoying?



Difference limen

$q_{ij} = 1$ ➡ **75% consensus**



Difference in annoyance

$$q_{ij} = q_i - q_j$$

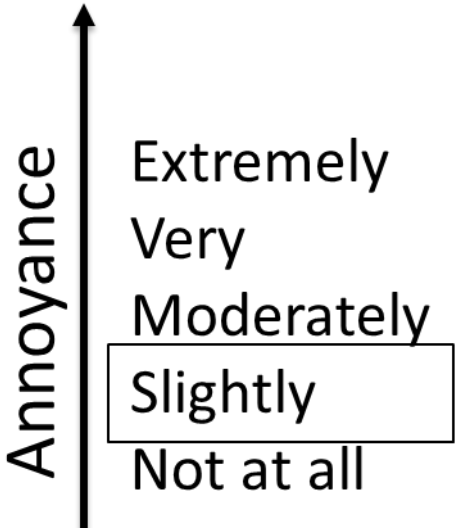
[2] Thurstone, "A law of comparative judgment," Psychological Review, 34, 273-286 (1927)

5-point responses and Law of Comparative Judgment[2]

- Community noise tests are not based on paired comparisons
- However, 5-point verbal responses can be interpreted in a paired comparison context[3]
- Apply Law of Comparative Judgment to get internal perception[4]

➢ **Law of Comparative Judgment applied to 5-point verbal annoyance scale**

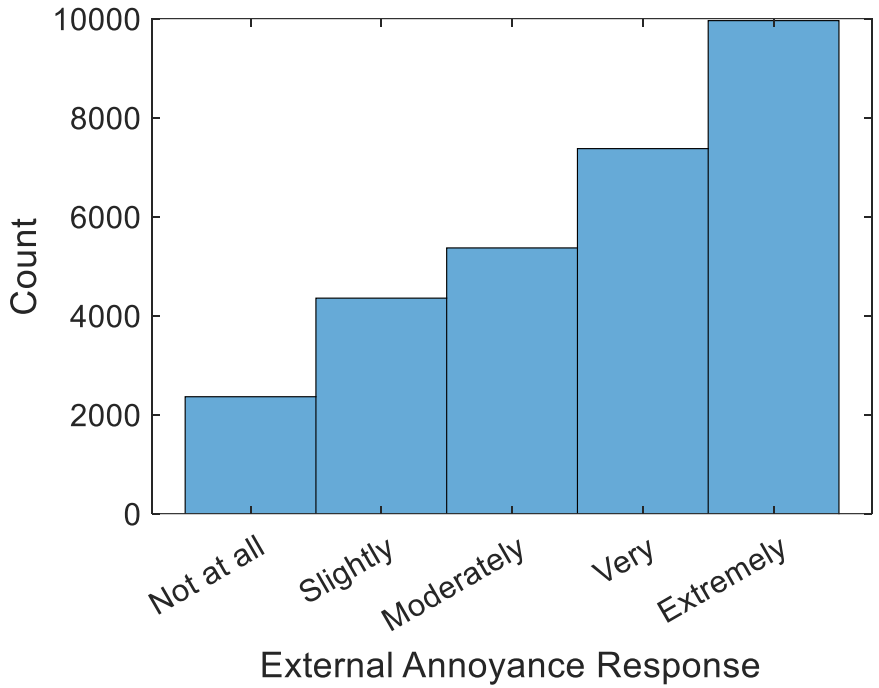
- Example response: "slightly"
- Judgment between what was heard and other verbal categories[3]
- What was heard:
  - More annoying than "not at all"
  - Less annoying than other categories

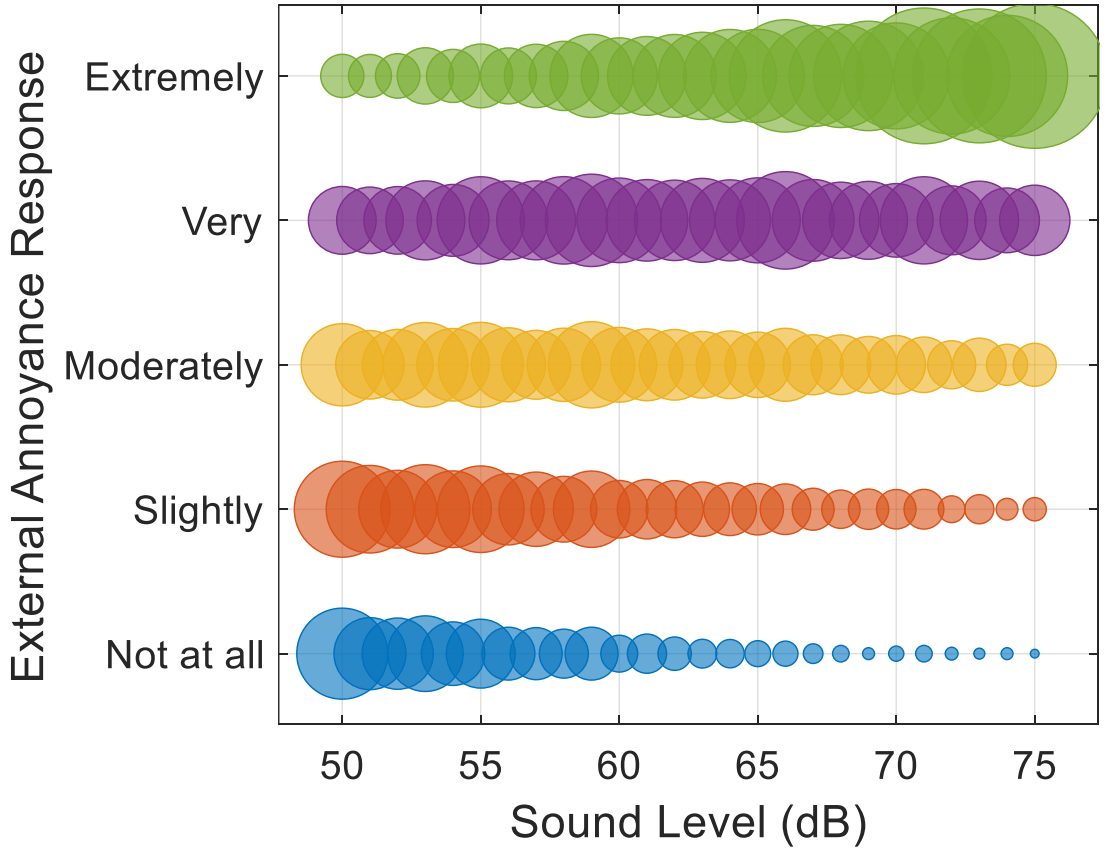


[2] Thurstone, "A law of comparative judgment," Psychological Review, 34, 273-286 (1927)
[3] Rimoldi et al., "The law of comparative judgment in the successive intervals and…," Educational Testing Service RB-54-5, 1954
[4] Perez-Ortiz and Mantiuk, "A practical guide and software for analysing paired comparison experiments," arXiv, 2017.

Application on a simulated dataset[5,6]

- **Dataset parameters***
  - 86 single events
  - 742 participants
  - 47% contribution rate





*Emphasis on simulated data. Demonstration purposes only. Interpretation for specific noise source not recommended.

[5] Doebler et al., "Dose error impacts on a collection of realistic dose-response curves…", InterNoise23

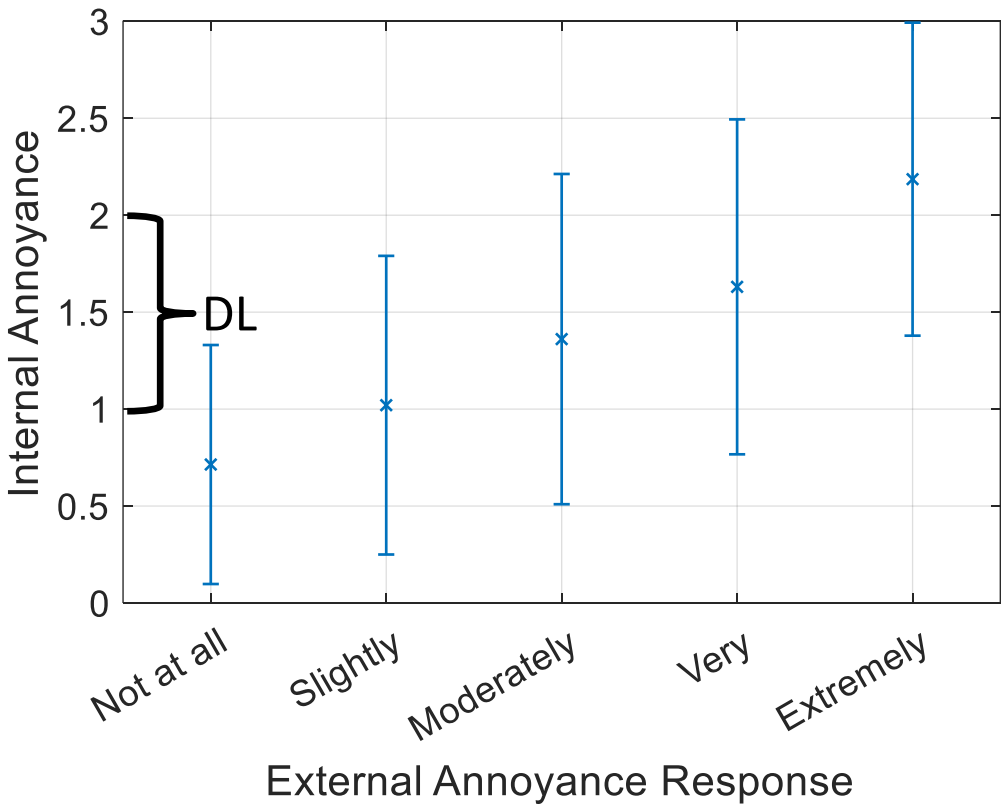
[6] Lee et al., "Bayesian statistical models for community annoyance survey data," J. Acoust. Soc. Am., 2020

Thurstone analysis gives internal annoyance perception

- **Each single event (SE) gets unique annoyance value**
  - Categorical based on SE ID
  - Sound level ignored
- **75% difference limen (DL) built-in[4]**
  - Δ(Internal Annoyance) = 1
- **Higher internal annoyance for higher 5-point verbal response**
- **Although relationship appears linear, spacing of categories is unknown**

- **Relationship we need to account for**

Internal Perception ➡ External Response



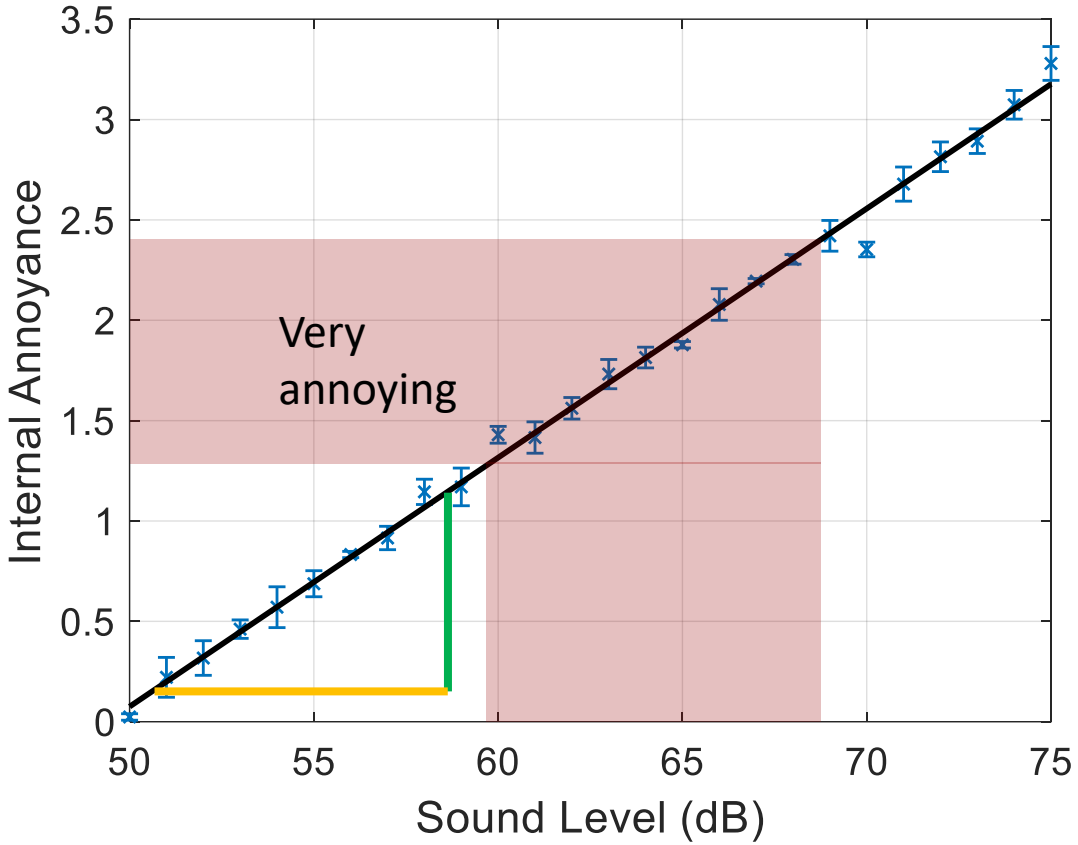
[4] Perez-Ortiz and Mantiuk, "A practical guide and software for analysing pairwise comparison experiments," arXiv Stat.AP, 2017

Model internal annoyance perception vs. sound level

- **Fechner's Law[1]**
  - Linear relationship between sensation and logarithm of stimulus
  - Applied to perception here
- **Internal annoyance highly correlated with sound level**
- **Just-noticeable-difference**
  - What is **ΔSPL** when **Δ(Annoyance) = 1**
  - 1 JND ≈ 8 dB *
- **Map verbal categories to sound level (e.g., "very")**

- **Relationship we want to model**

Stimulus → Internal Perception

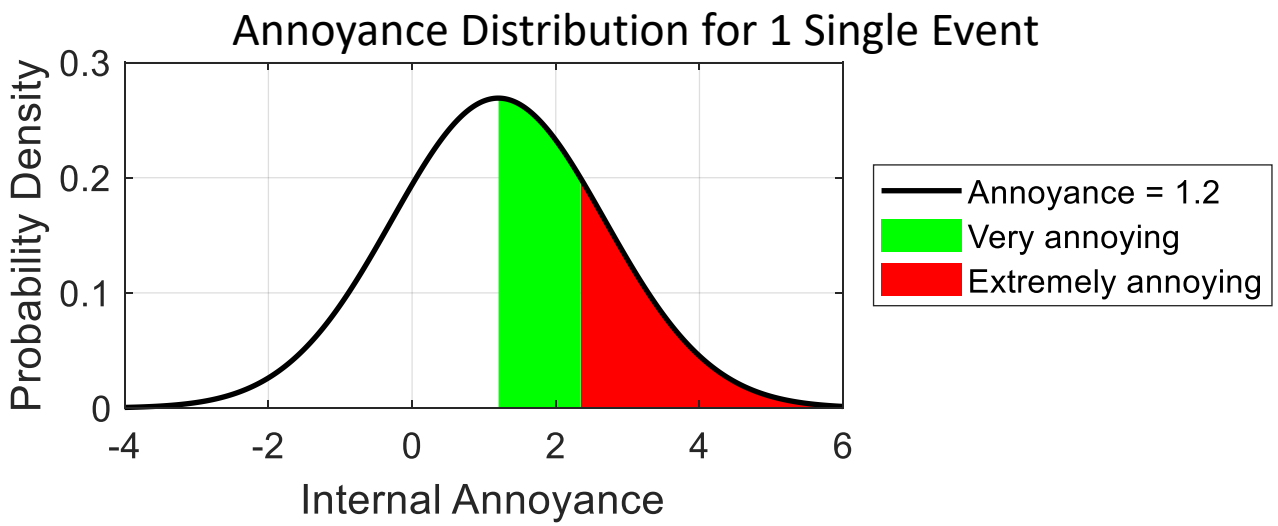


[1] G.A. Gescheider, "Psychophysics", 1997

Comparison of standard and Thurstone approaches

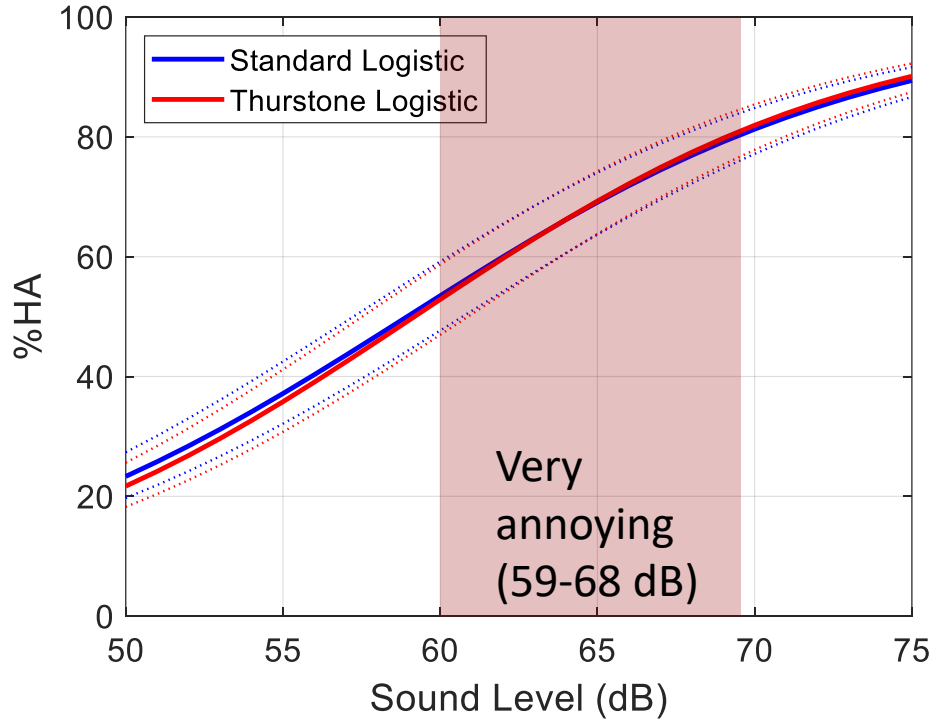
Thurstone Logistic Approach

- Probability density of annoyance for each single event
- Calculate percent highly annoyed
- Generate new binary response and do logistic regression



Relationship we often model

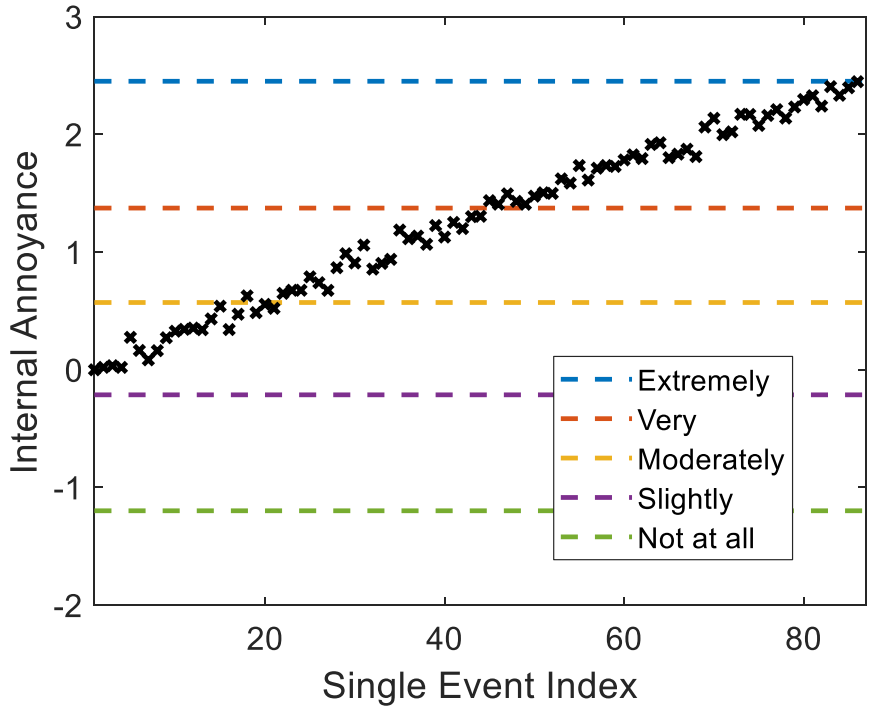
Stimulus ➡ External Response



- Both approaches agree
- Thurstone gives range of categories

Evaluating distance between verbal categories

- **Spacing may be sensitive to study demographics (ISO/TS 15666)**
  - Indicated by Lee et al. using ordinal regression (JASA, 2020)
  - Fidell and Teffeteller (JSV, 1981)



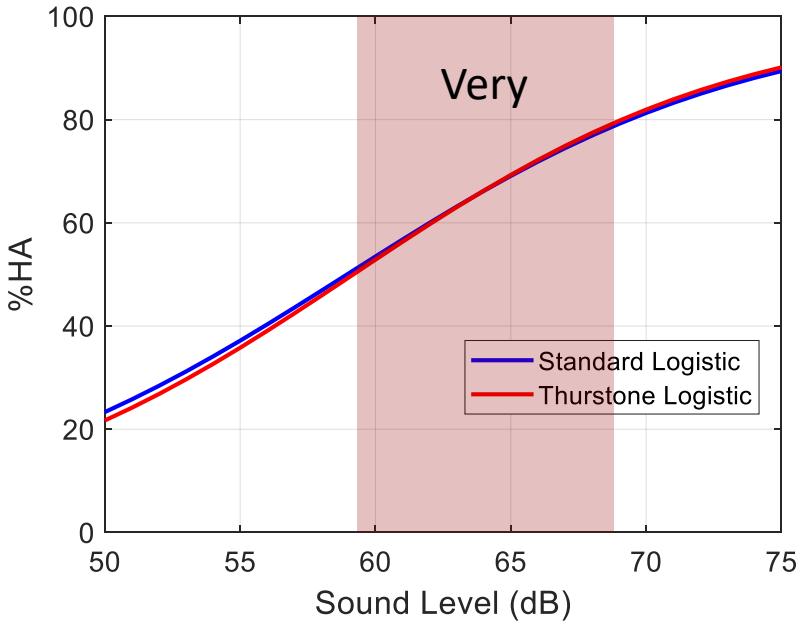
1 JND ≈ 8 dB

Interval	Difference in Annoyance	Change in Sound Level (dB)
'Extremely' – 'Very'	1.08	8.7
'Very' – 'Moderately'	0.80	6.5
'Moderately' – 'Slightly'	0.78	6.3
'Slightly' – 'Not at all'	0.99	7.9

- **Thurstone analysis**
  - Can quantify distance between 5-point verbal response categories
  - Use JND to determine change in sound level between response categories

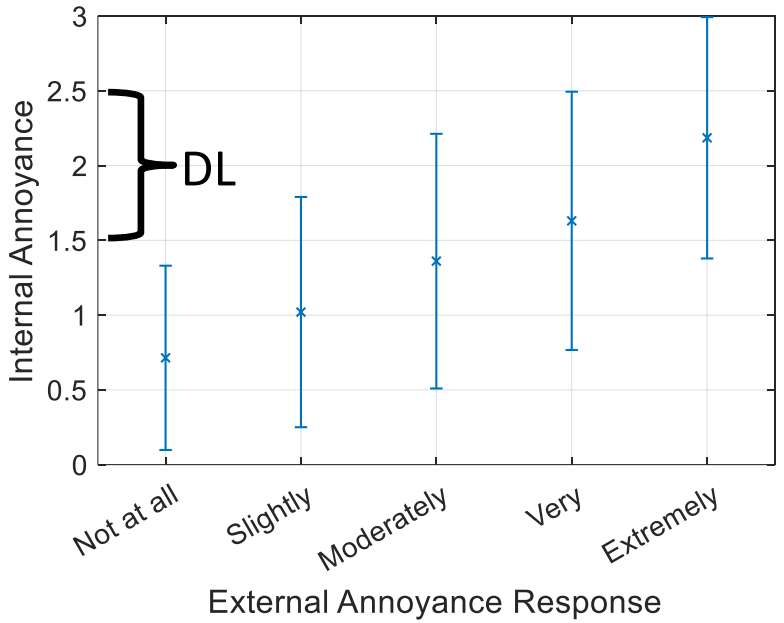
Summary

Relationship we often model



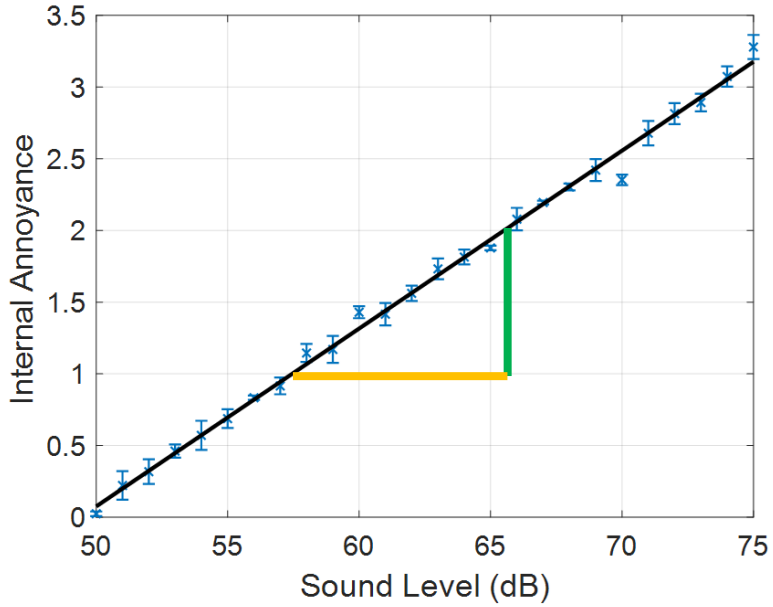
- Thurstone analysis agrees with standard approach
- Range of response categories found

Relationship we need to account for



- Continuous annoyance scale with built-in 75% difference limen (DL)

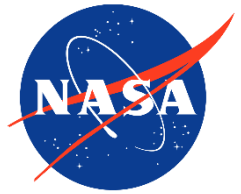
Relationship we want to model



- Determine JND for noise metrics

Additional insight: spacing of verbal categories can be quantified

Thank You



- Work supported by Commercial Supersonic Technology Project.

- Thank you to Will Doebler (NASA Langley) for generating the dataset used in this presentation.

- Acknowledge Jonathan Rathsam, Pete Parker, Aaron Vaughn, Nathan Cruze, Will Doebler and Kathryn Ballard for useful discussions.

For questions: matthew.a.boucher@nasa.gov