Enhancing Auditory-Perceptual Skills and Optimizing Scaling Methods for Dysarthria Assessment

Mili Kuruvilla-Dugdale, PhD
Associate Professor
Speech Perception and Kinematics Lab
Department of Communication Sciences and Disorders
University of Iowa

A Historical Account

1817
Neurologists

Speech disorders as signs of neurologic disease

Distinction between speech vs. language loss

1872
Elocutionists

Visible Speech: Similar to IPA

Short forms to characterize consonants, vowels, and diphthongs.

1924
Speech Doctors

Subsystem assessment

Chewing method for dysarthria

Visuo-tactile methods for speech production

1949
Speech Doctors

Dysarthria features are a function of localization

Six dysarthria types based on anatomic localization

1969
SLP-Neuro Team

Classification and rating system for MSDs

Distinctive speech patterns of dysarthria types

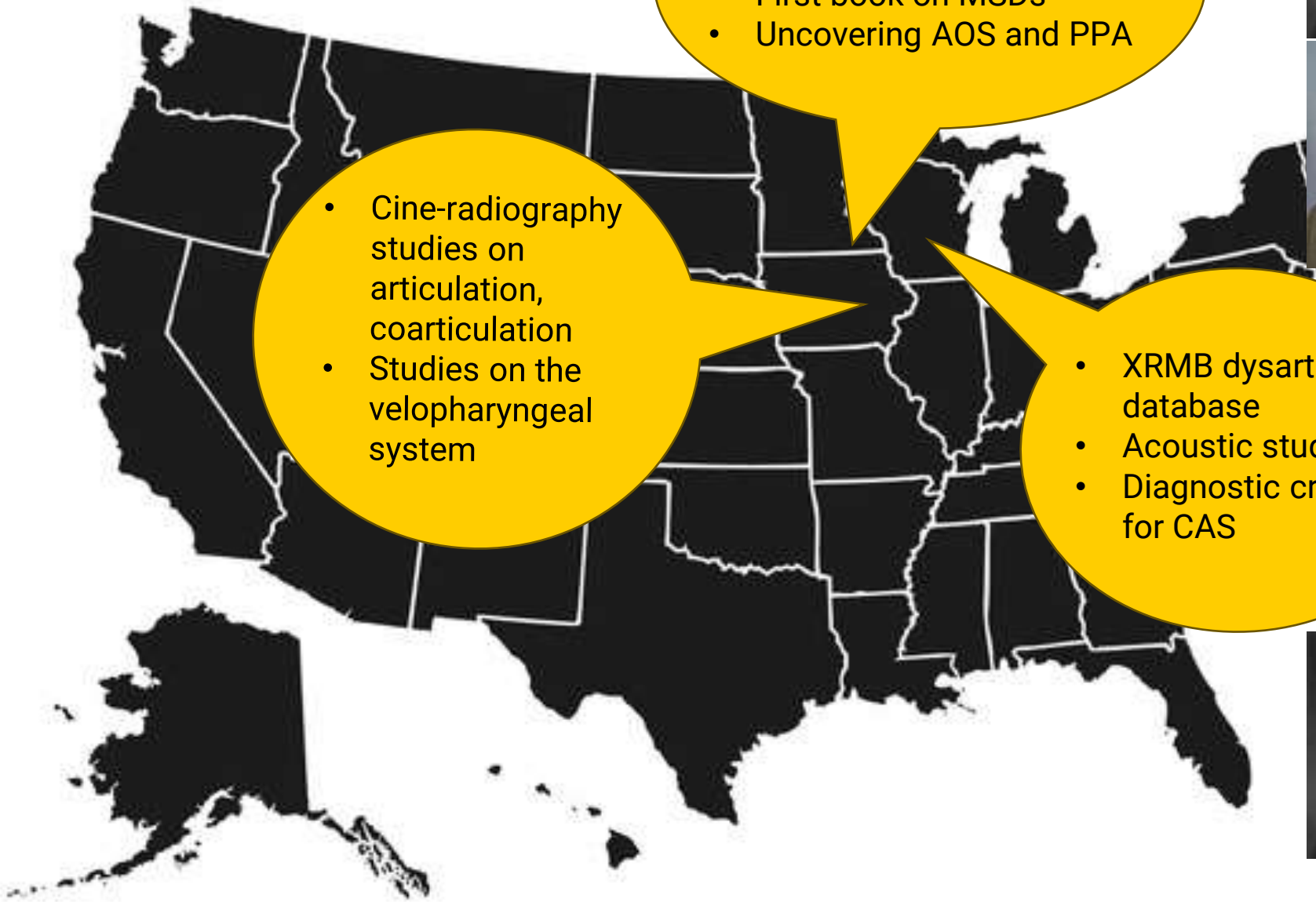
IOWA

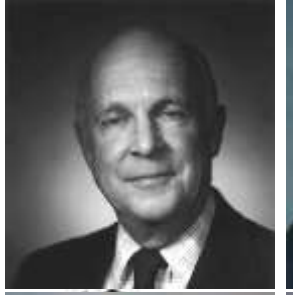
The OG Influencers

- Mayo Clinic Classification & Rating System
- First book on MSDs
- Uncovering AOS and PPA

- Cine-radiography studies on articulation, coarticulation
- Studies on the velopharyngeal system

- XRMB dysarthria database
- Acoustic studies
- Diagnostic criteria for CAS







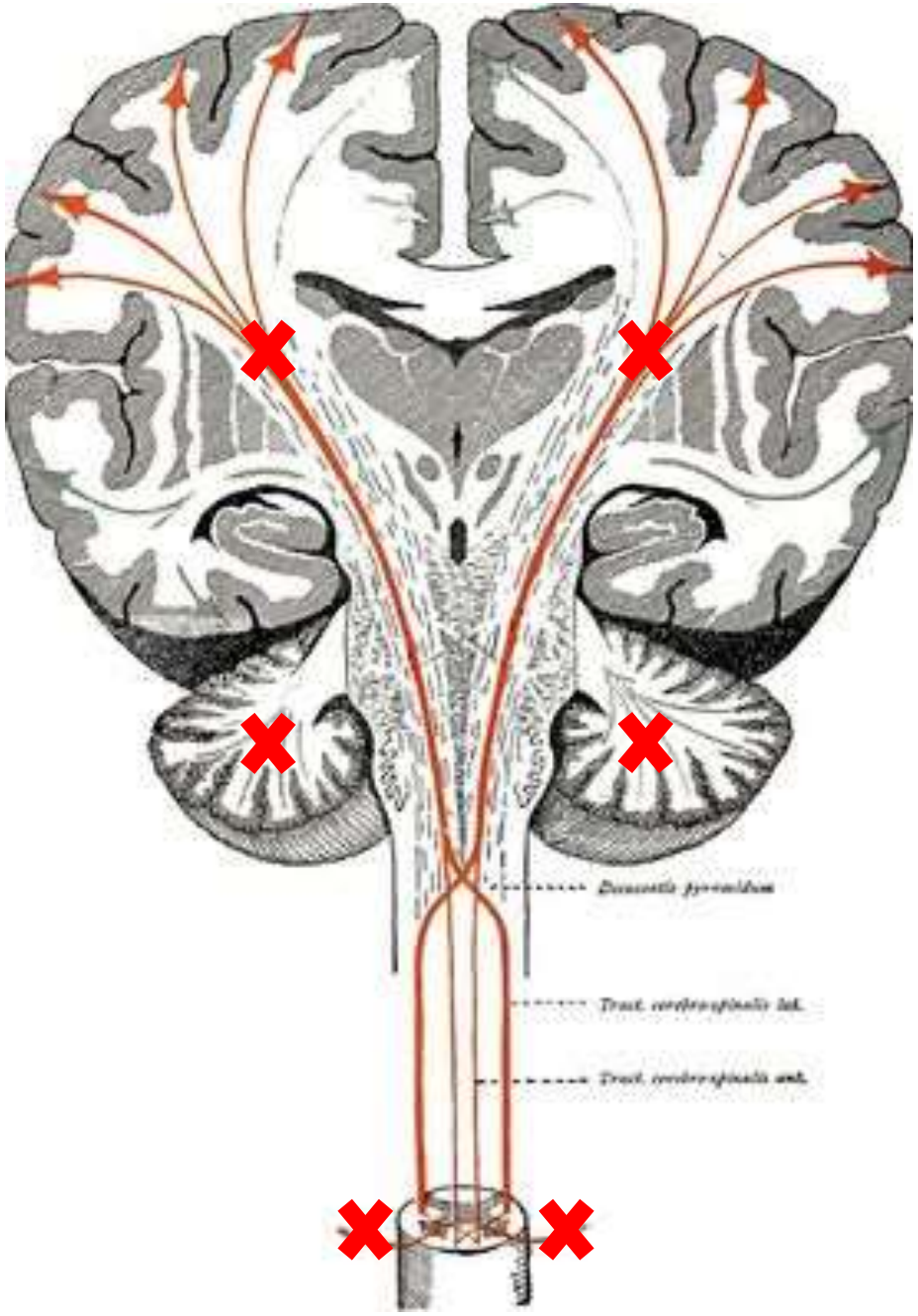






Dysarthria Classification System: Mayo Approach

- Spasticity = Spastic dysarthria
  - Localization = Bilateral UMN lesion
- Weakness = Flaccid dysarthria
  - Localization = Unilateral or bilateral LMN lesion
- Incoordination = Ataxic dysarthria
  - Localization = Unilateral or bilateral cerebellar lesion



Dysarthria: Is it Understudied?

79, 300 

Dysarthria

16, 200 

Articulation & Dysarthria

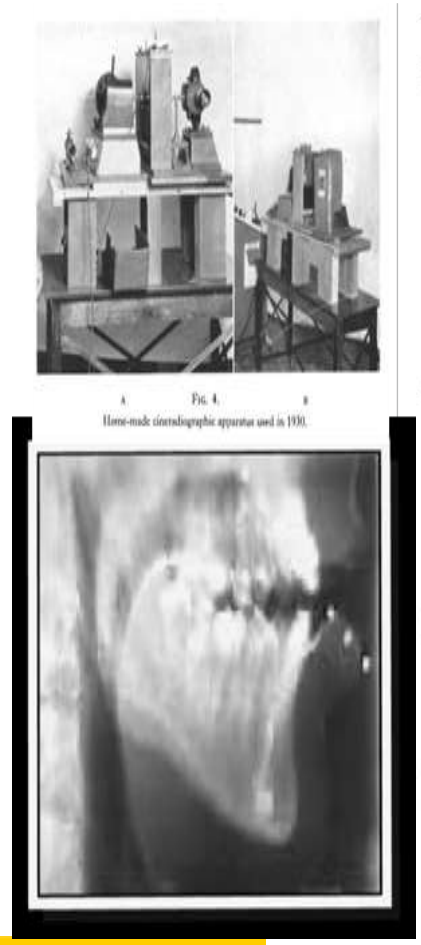
3, 050 

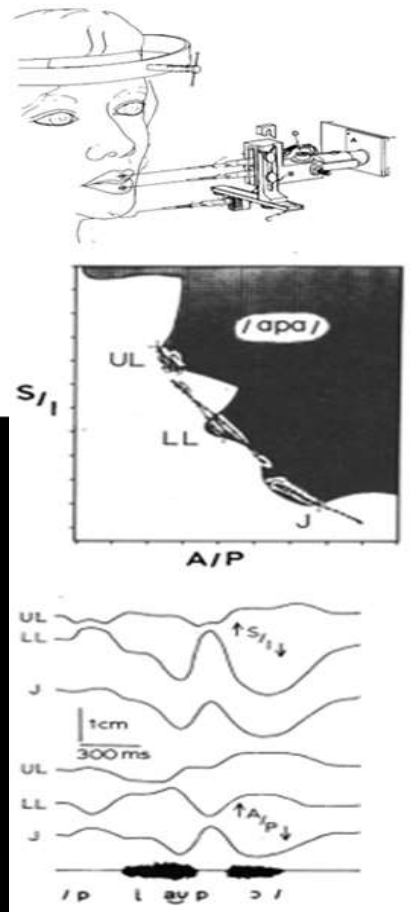
Tongue Kinematics & Dysarthria

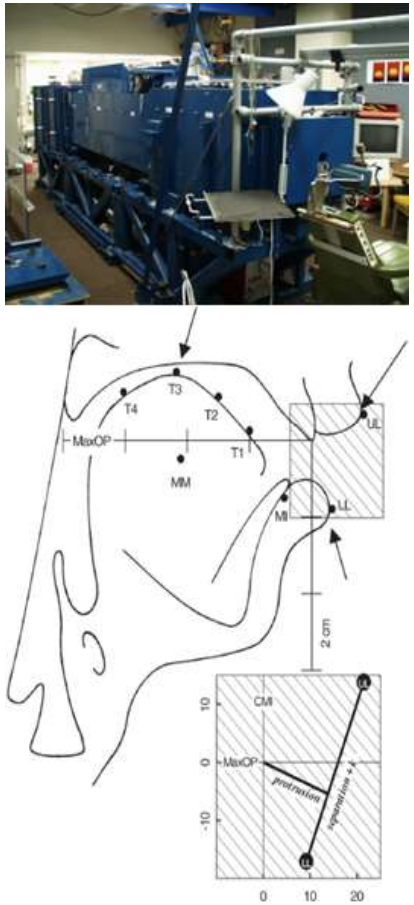
Keywords	Article Count
Dysphagia	381,000
Aphasia	265,000
Articulation & Speech Sound Disorders	20,500

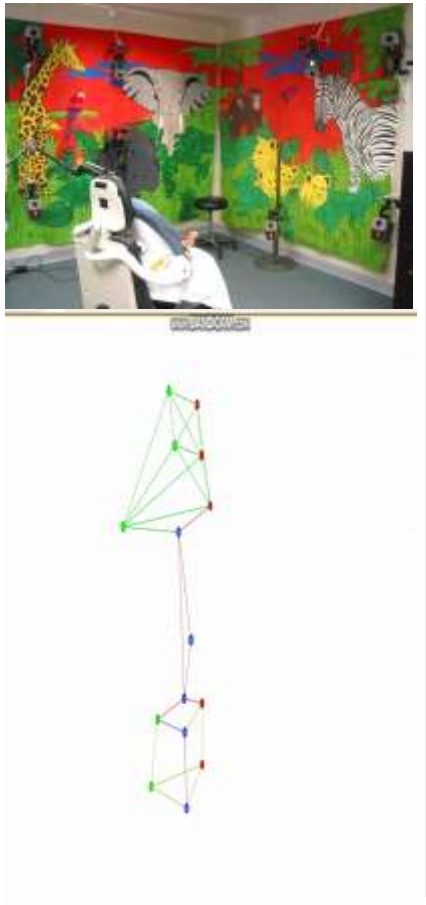
IOWA

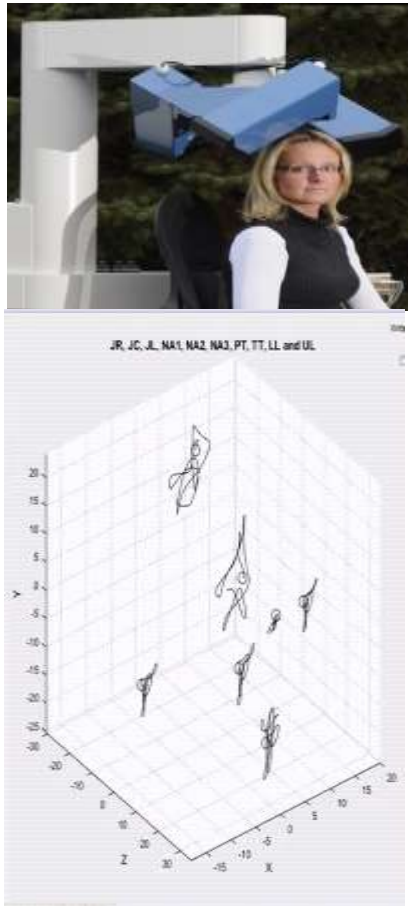
Despite Technological Advances Over 60 years…









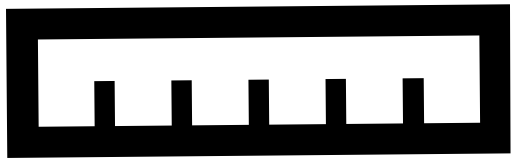


Gold Standard for Assessing Dysarthria

ADVANTAGES

LIMITATIONS

LISTENING & RATING PRACTICE



Orthographic Transcription

- Intelligibility or severity index
- High reliability
- Relatively insensitive to mild speech loss
- Time and resource costs

Scaling

- Severity of system and subsystem level changes
- Reliability depends on scale and feature
- Several scaling options to suit time and resource needs

Auditory-Perceptual Scaling

- Wrought with challenges
  - High clinician variability (Bunton et al., 2007)
  - Poor construct validity (Whitehill et al., 2002)
  - Response biases
  - Considerable flexibility in choice of rating scale and salient features

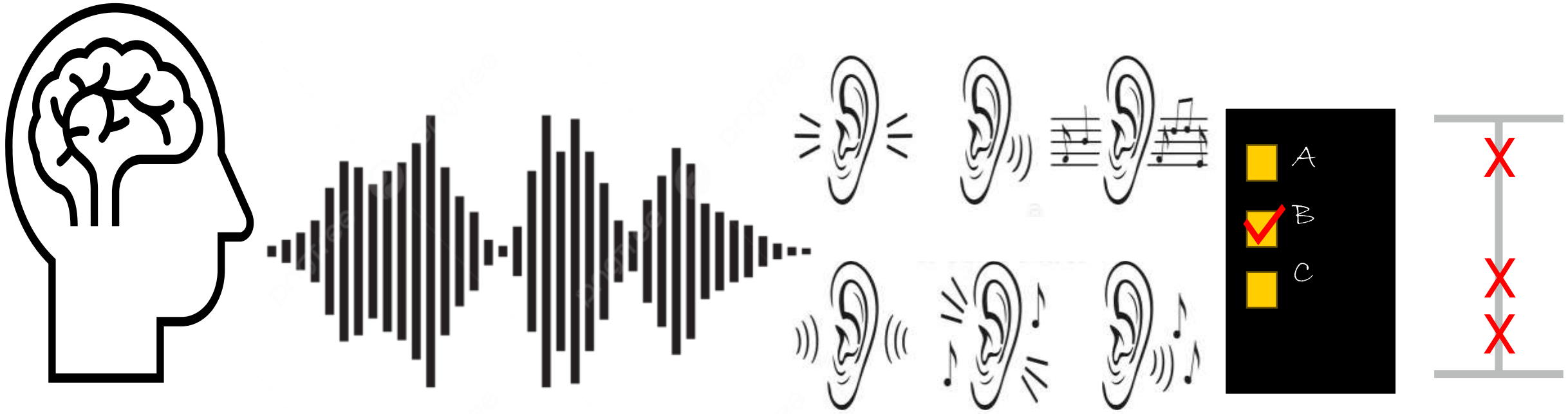
COVID-19: A Silver Lining

- Uptick in auditory-perceptual studies on dysarthria
- Circumvent resource limitations and research restrictions
- Improve scalability of assessment methods

But…

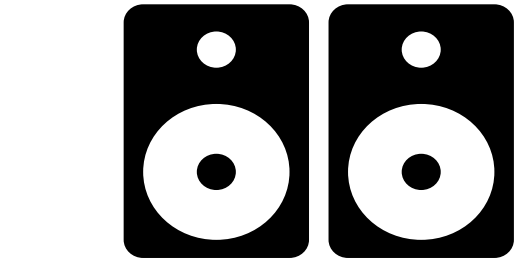
  – Exposed the many gaps

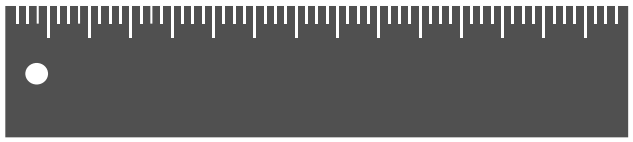
Central Problem: Poor Listener Reliability and Agreement



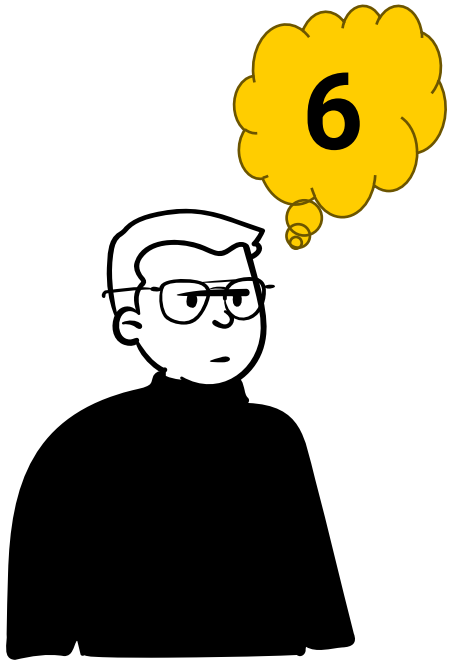
Auditory Perceptual Rating: Challenges

1 SIGNAL





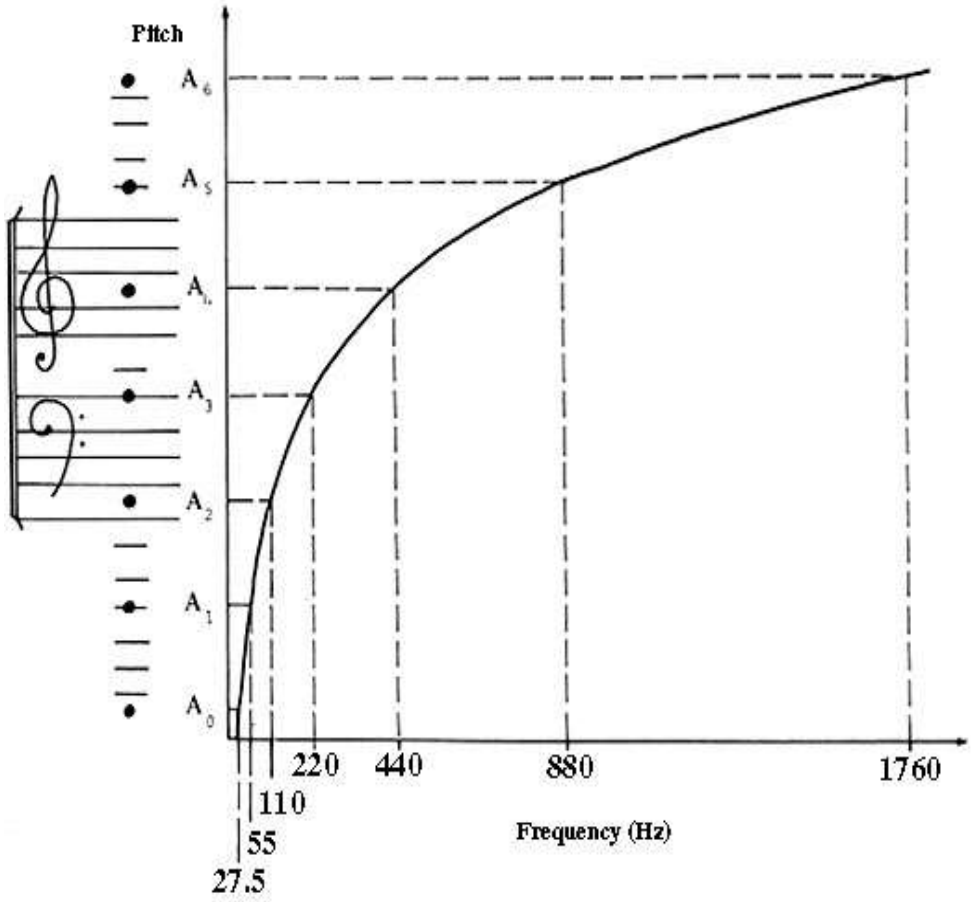
3 TASK



2 RATER

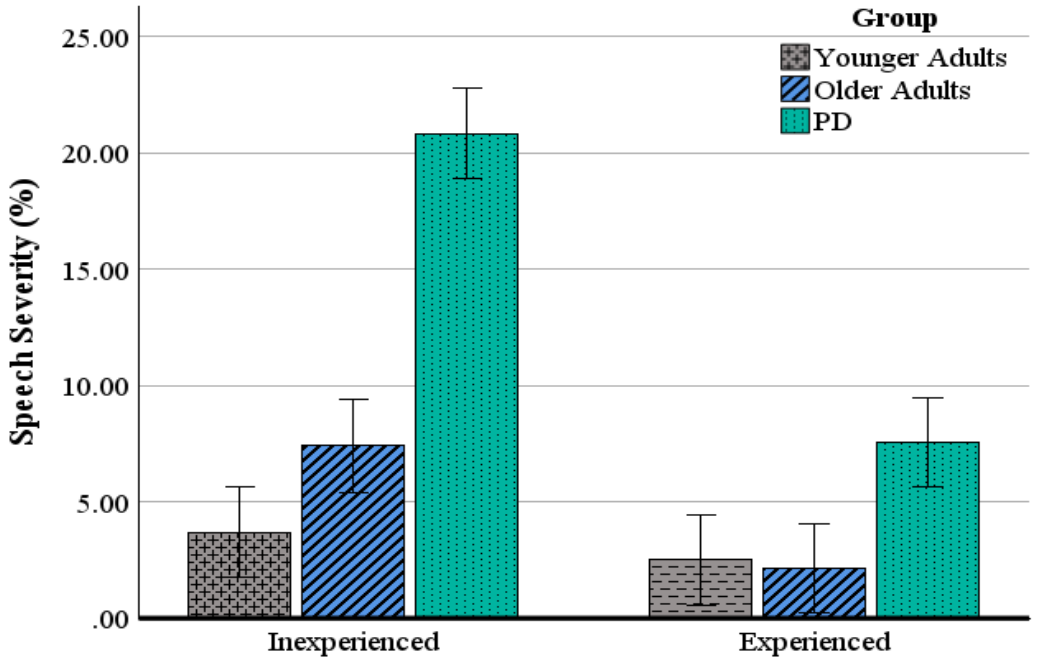
Signal-Related Challenges

- Multidimensionality
  - ❑ Reduced loudness
  - ❑ Monopitch/Monoloudness
  - ❑ Short rushes of speech
- Mapping physical units
  - ❑ Pitch = Frequency
  - ❑ Loudness = Intensity
  - ❑ Overall Severity = ?
  - ❑ Voice Quality = ?



Rater-Related Challenges

- Internal standards
  - Experts: Unstable and idiosyncratic
  - Non-experts: Blank slate
- Training and experience
  - Uniform training
  - Components of training
  - Years and type of experience



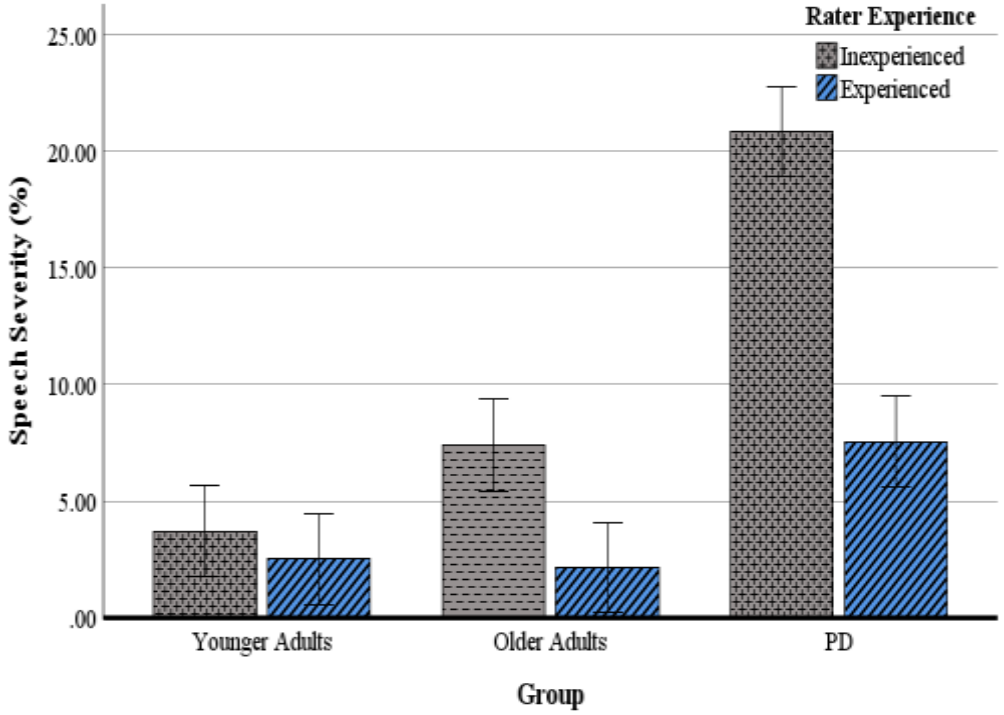
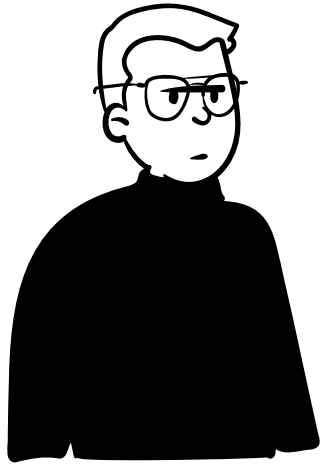
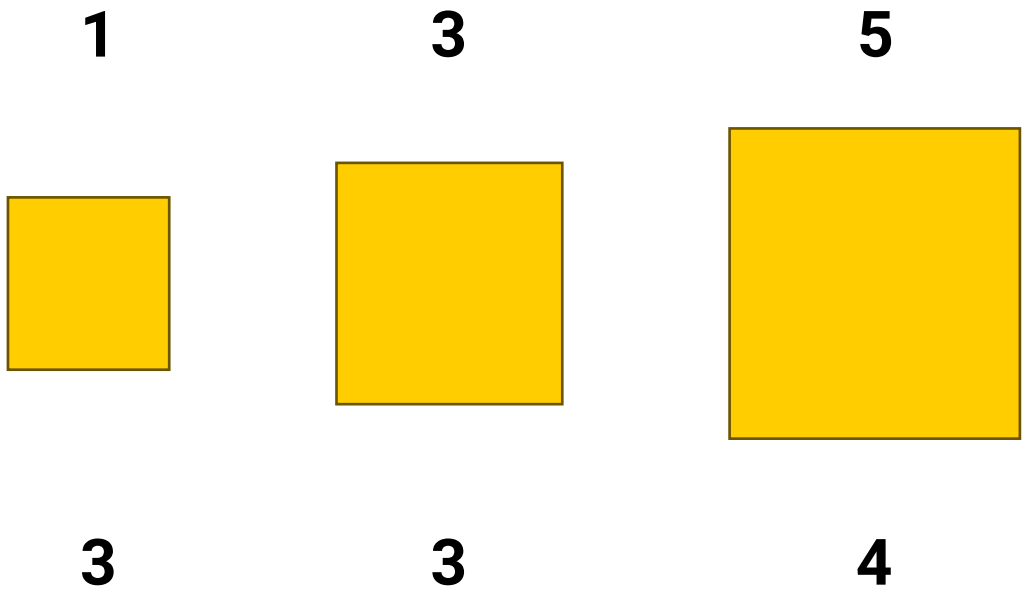


Figure 2 in: Kuruvilla-Dugdale et al. (2019). A comparative study of auditory-perceptual speech measures for the early detection of mild speech Impairments. *Seminars in Speech and Language*, *40*, 394–406.

Task-Related Challenges

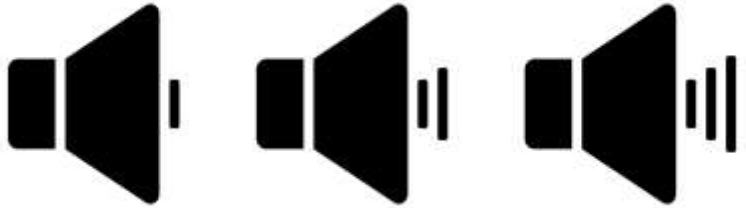
- Instructions
- Scale type
  - Nominal
  - Ordinal
  - Interval
  - Ratio





Task-Related Challenges

- Feature type
  - Prothetic



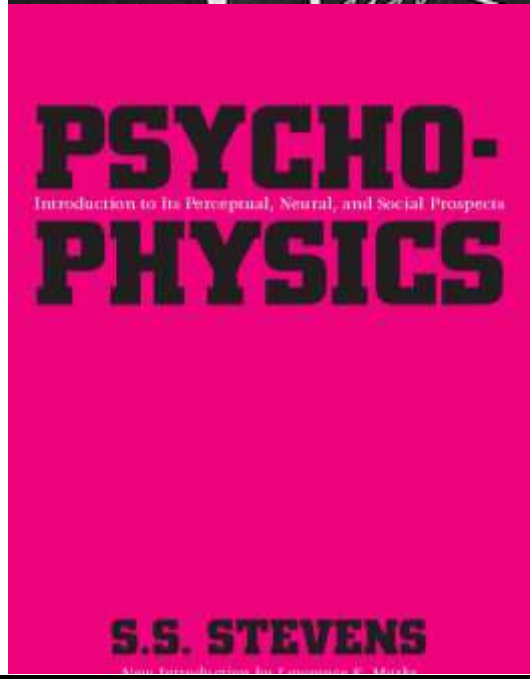
  - Metathetic

Optimizing Auditory-Perceptual Scaling: Task

Measurement Level **Psychophysical Continua** **Scale Fit**

Optimizing Auditory-Perceptual Scaling

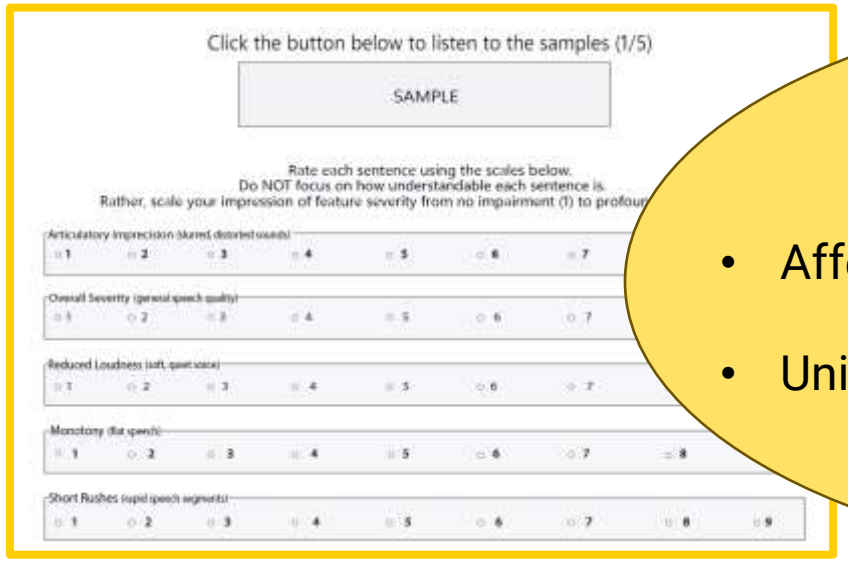
- Measurement level
  - Nominal
  - Ordinal
  - Interval
  - Ratio
- Psychophysical continua to be rated
  - Prothetic
  - Metathetic
- Construct validity
  - Scale fit



Measurement Level

INTERVAL SCALE

Ordinal or interval level data?



DIRECT MAGNITUDE ESTIMATION

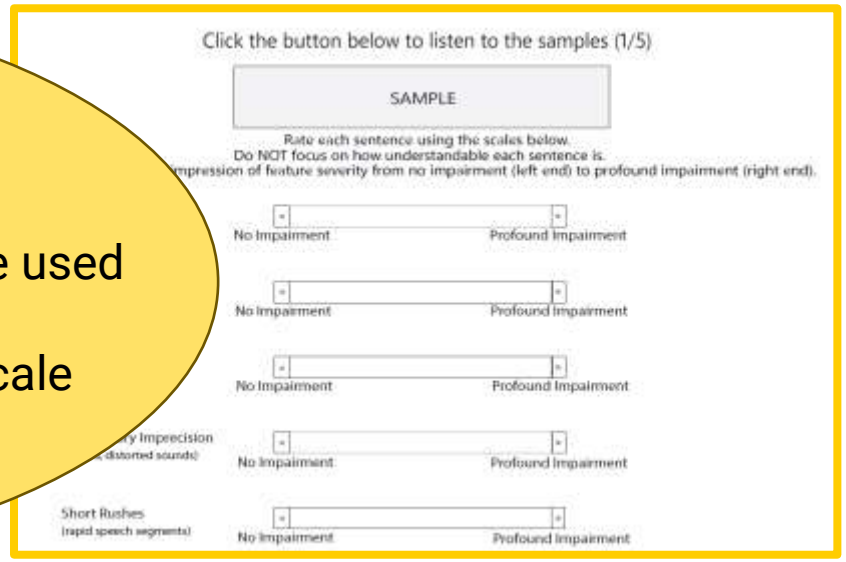
Ratio level data

Why Does It Matter?

- Affects the statistical tests that can be used
- Unique biases associated with each scale

VISUAL ANALOG SCALE

Measurement level is unknown



IOWA

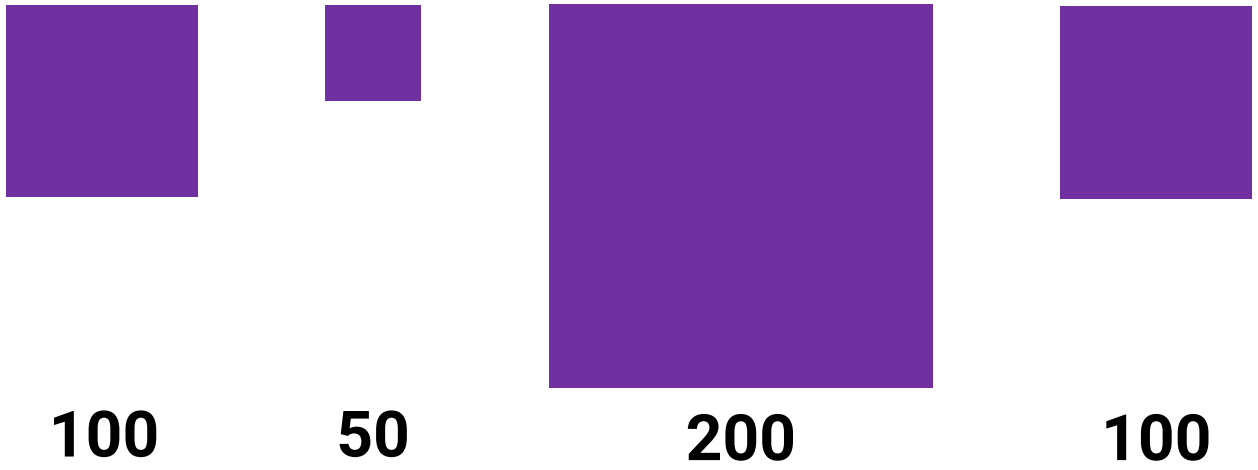
Construct Validity: Scale Fit

Scale Type + Feature Type

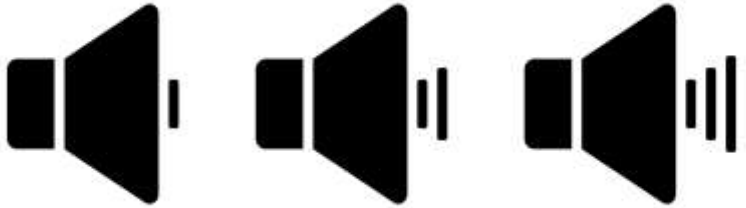
Interval/Ratio + Prothetic/Metathetic

1

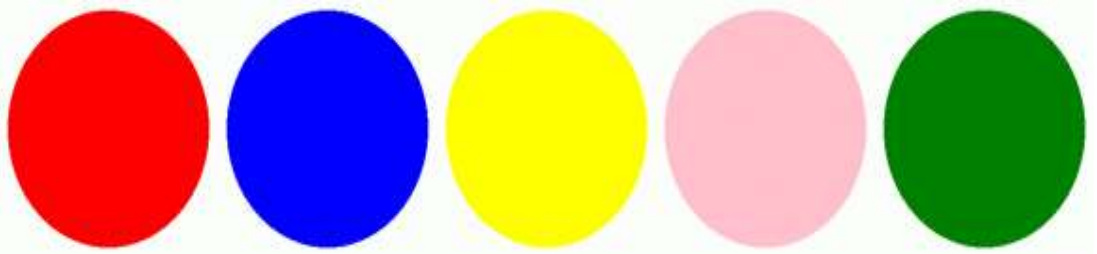
2



A



B



Construct Validity: Scale Fit

Length Increase	Perceived Length
+ 20 cm	20
+ 40 cm	40
+ 80 cm	80

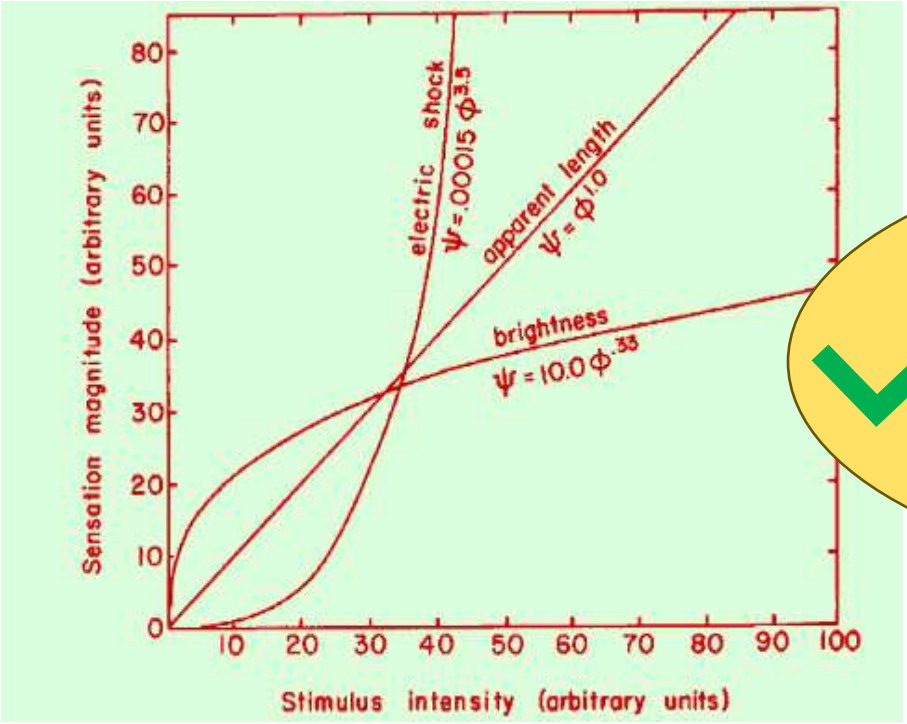
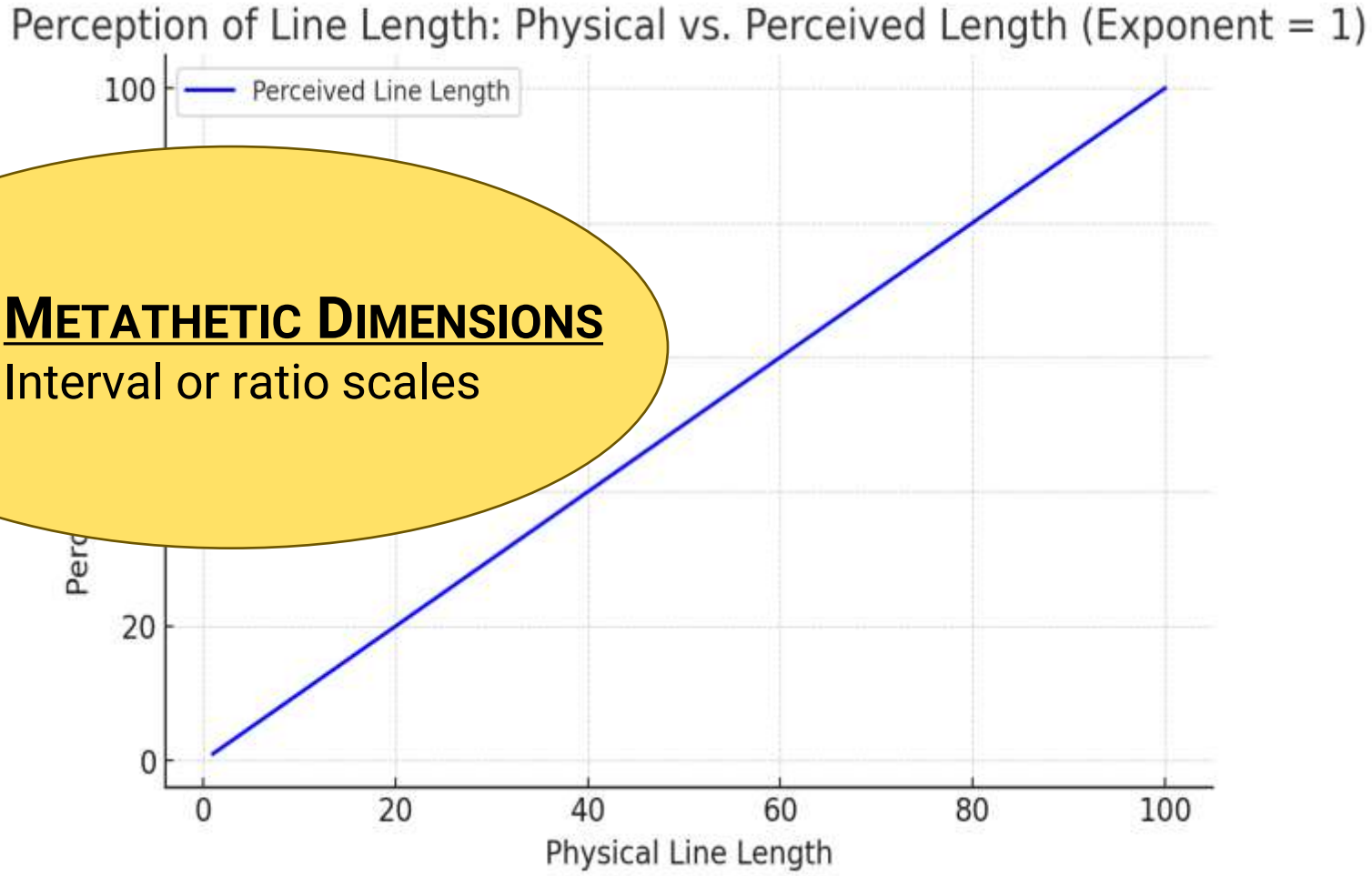


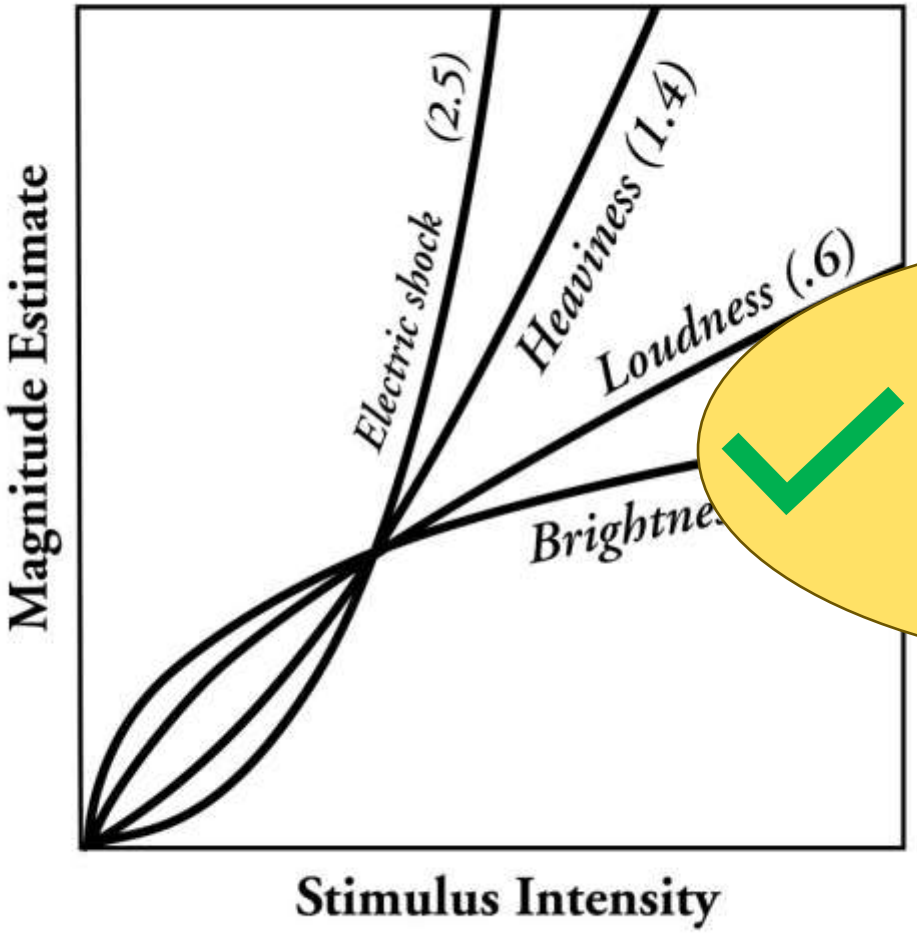
Figure 8.4 in: Stipancic et al. (2024). Tipping the scales: Indiscriminate use of interval scales to rate diverse dysarthric features. *Journal of Speech, Language, and Hearing Research*, *67*, 3673-3685.

METATHETIC DIMENSIONS
Interval or ratio scales

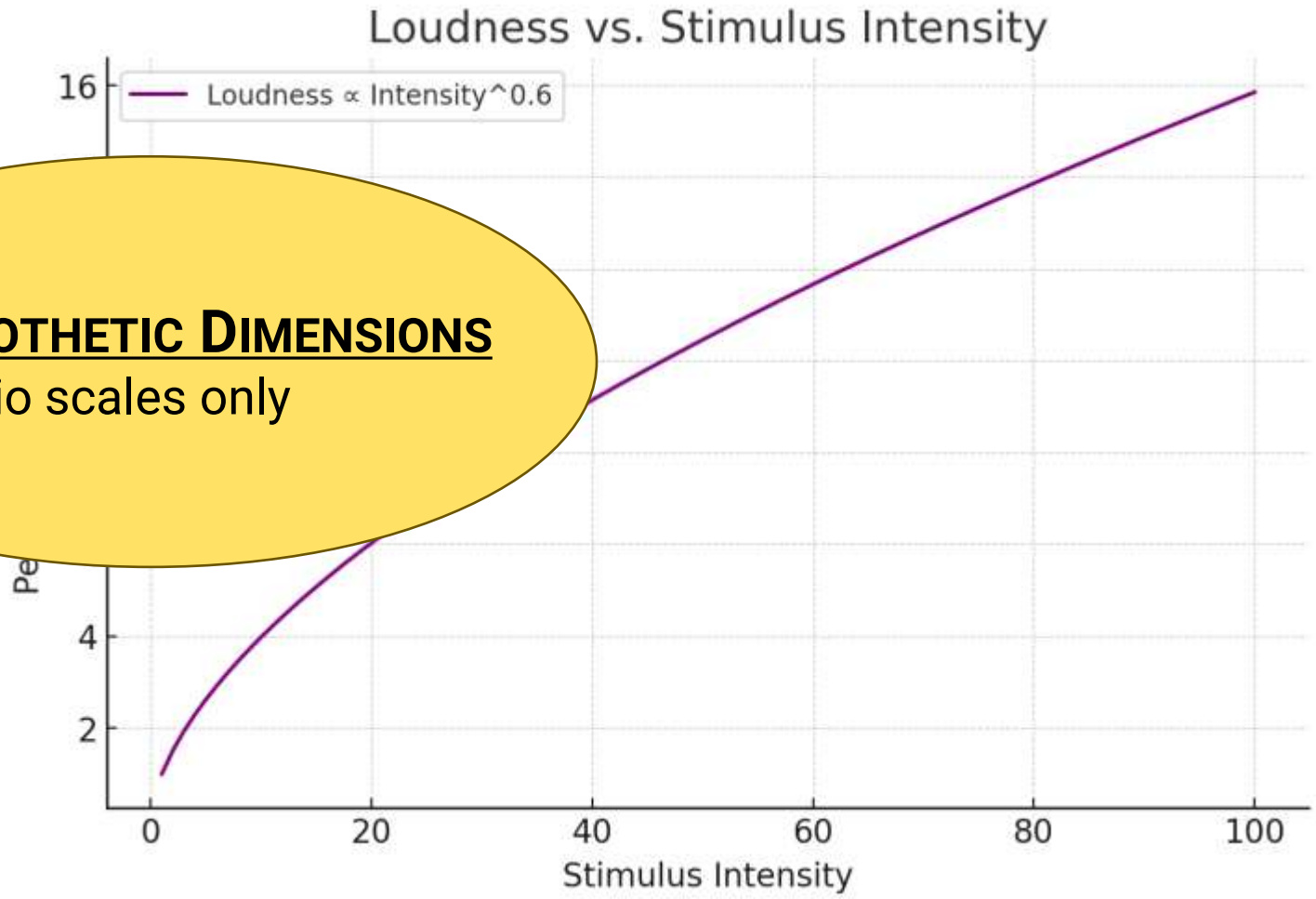


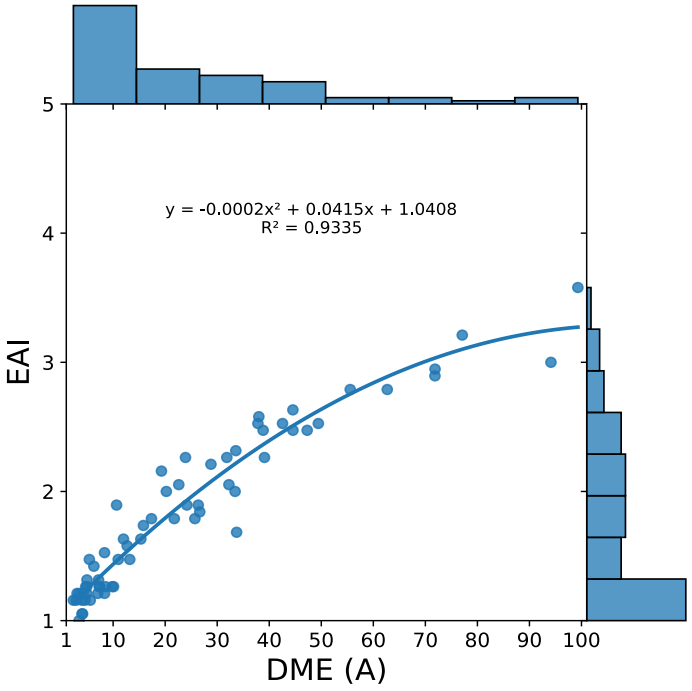
Construct Validity: Scale Fit

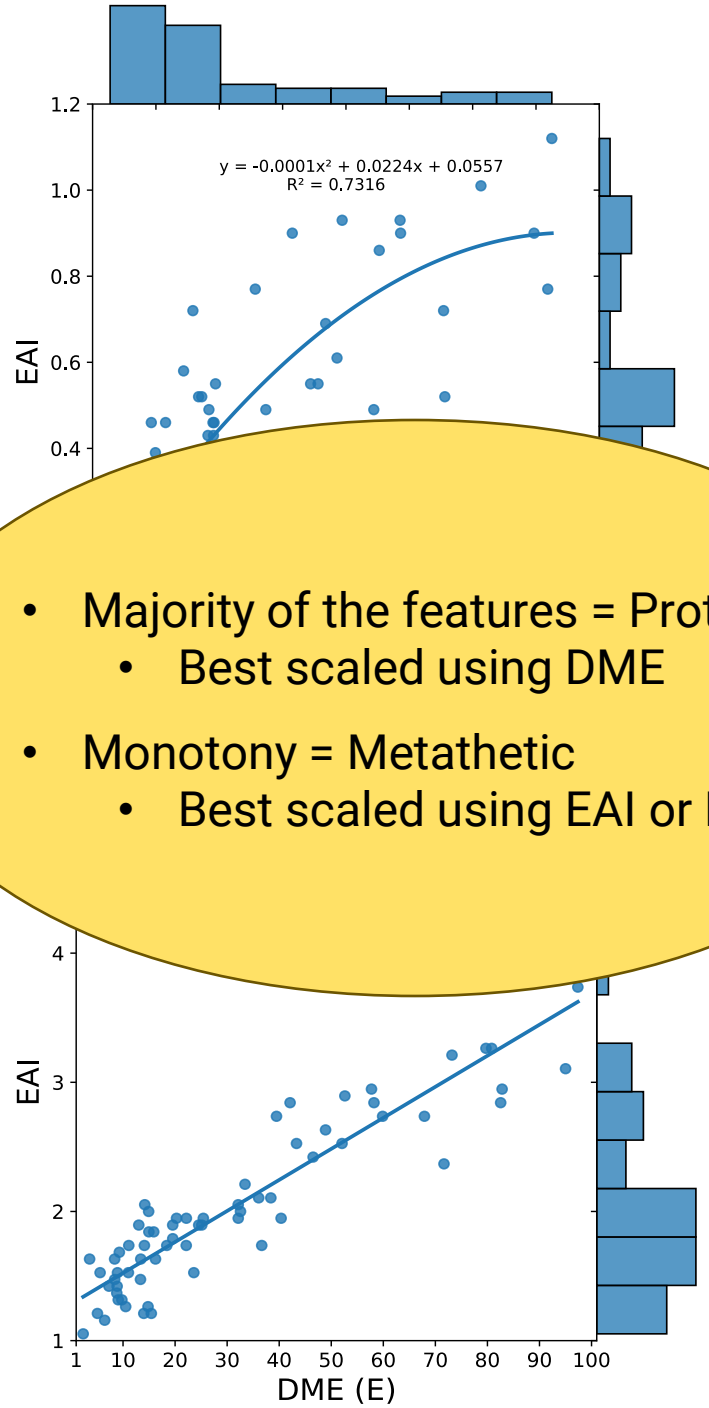
dB Increase	Perceived Loudness
+ 10 dB	3.98
+ 20 dB	6.03
+ 40 dB	9.15
+ 80 dB	13.86

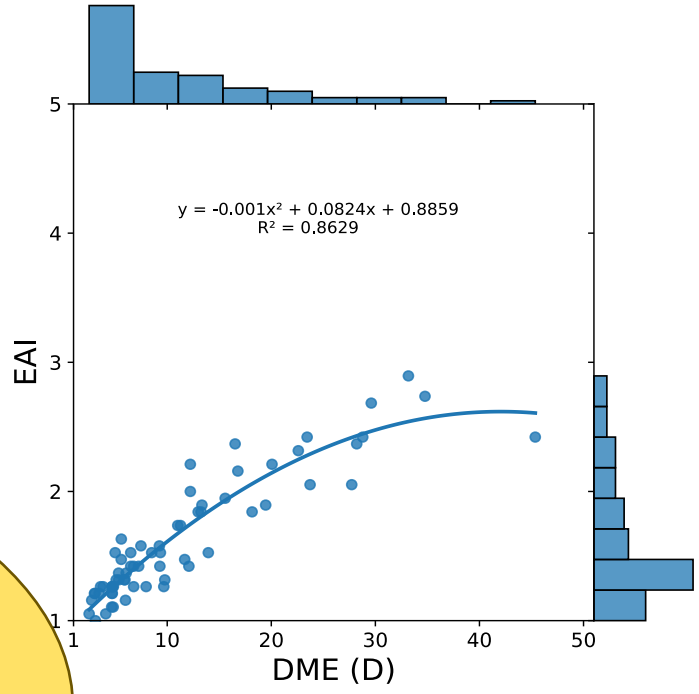


PROTHETIC DIMENSIONS
Ratio scales only





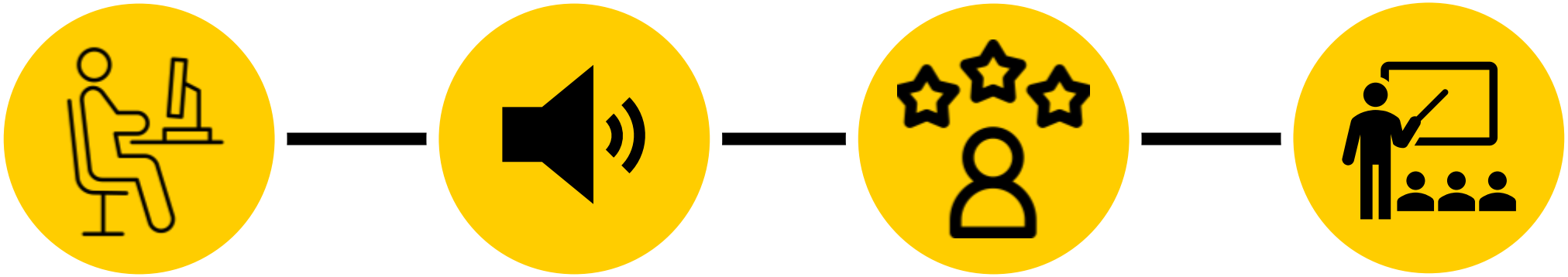




- Majority of the features = Prothetic
  - Best scaled using DME
- Monotony = Metathetic
  - Best scaled using EAI or DME

Figure 2 in: Stipancic et al. (2024). Tipping the scales: Indiscriminate use of interval scales to rate diverse dysarthric features. *Journal of Speech, Language, and Hearing Research*, *67*, 3673-3685.

Optimizing Auditory-Perceptual Scaling: Training



Explicit Definitions

External Anchors

Multidimensionality

Uniform Training Experiences

	PRE-TEST	TRAINING	POST-TEST
Speech Samples	SIT sentences (6-9 word) *N* ALS = 18, *N* PD = 18, *N* controls = 12 48 samples + 11 samples repeated for intra-rater reliability = 59 samples	Grandfather Passage segments *N* ALS = 12, *N* PD = 6, *N* controls = 18 36 training samples + 7 samples repeated for intra-rater reliability + 18 anchors for familiarization = 61 samples	SIT sentences (6-9 word) *N* ALS = 18, *N* PD = 18, *N* controls = 12 48 samples + 11 samples repeated for intra-rater reliability = 59 samples
Listeners	Training group (*N* = 22) Non-training group (*N* = 22)	Training group (*N* = 22)	Training group (*N* = 22) Non-training group (*N* = 22)
Procedures	Listeners rated overall severity, imprecision, monotony, and speech rate of each sample on individual visual analog scales	Orientation → Familiarization → Stimulus-Response → Feedback	Listeners rated overall severity, imprecision, monotony, and speech rate of each sample on individual visual analog scales
Outcome Measures	ICC and VAS scores for overall severity, imprecision, monotony, and slow rate	N/A	ICC and VAS scores for overall severity, imprecision, monotony, and slow rate

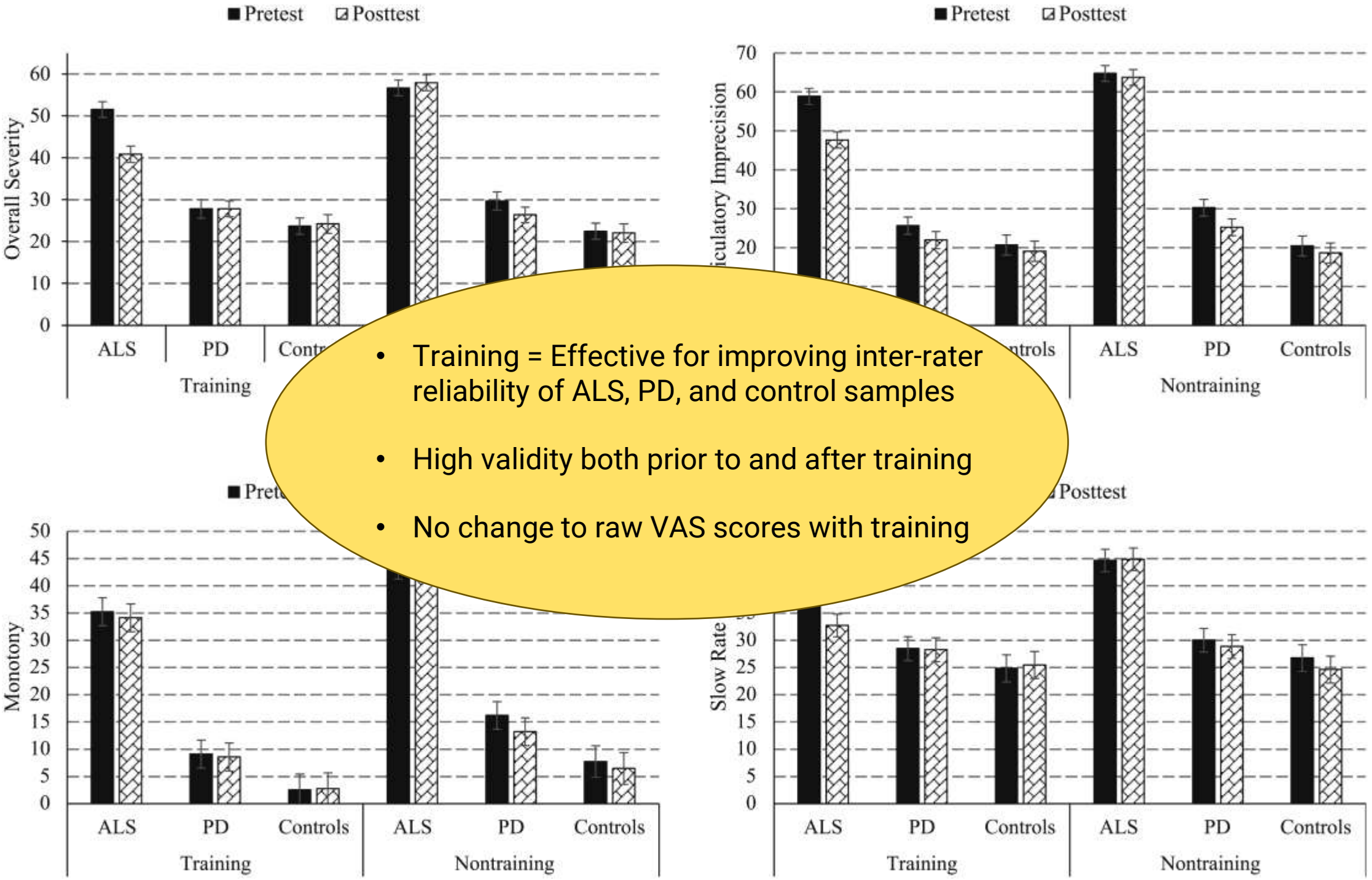
Same as pre-training testing

Scores compared

Study Aim 1:
Inter-rater and intra-rater reliability (ICCs)
– within the training and non-training groups across pre-and post-test timepoints

Study Aim 2:
Raw VAS scores
– between the non-expert listeners (i.e., training and non-training) and expert listeners across pre- and post-test timepoints

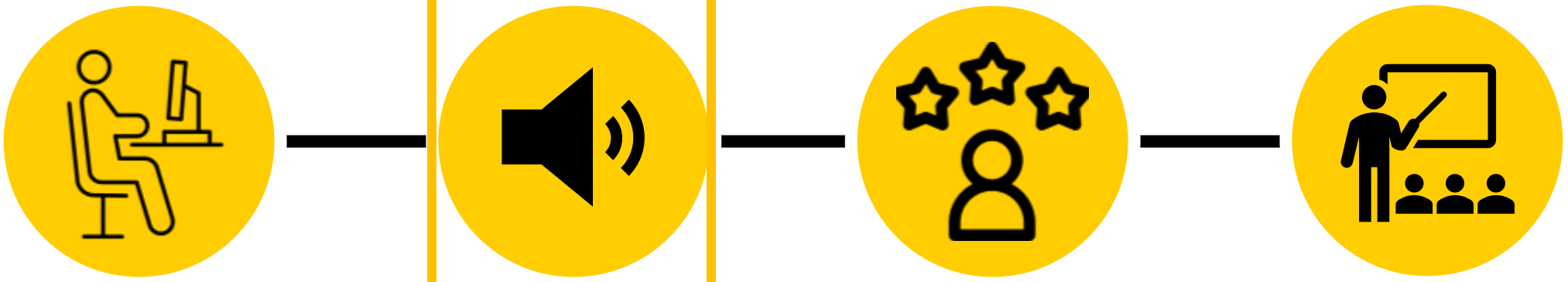
Study Aim 3:
Raw VAS scores
– within and between the training and non-training groups across pre- and post-test timepoints



- Training = Effective for improving inter-rater reliability of ALS, PD, and control samples
- High validity both prior to and after training
- No change to raw VAS scores with training

Figure 2 in: Stipancic et al. (2023). Improving perceptual speech ratings: The effects of auditory training on judgments of dysarthric speech. *Journal of Speech, Language, and Hearing Research*, 66(11), 4236-4258..

Optimizing Auditory-Perceptual Scaling: External Anchors



Explicit Definitions

External Anchors

Multidimensionality

Uniform Training Experiences

External Anchors

Natural Anchors

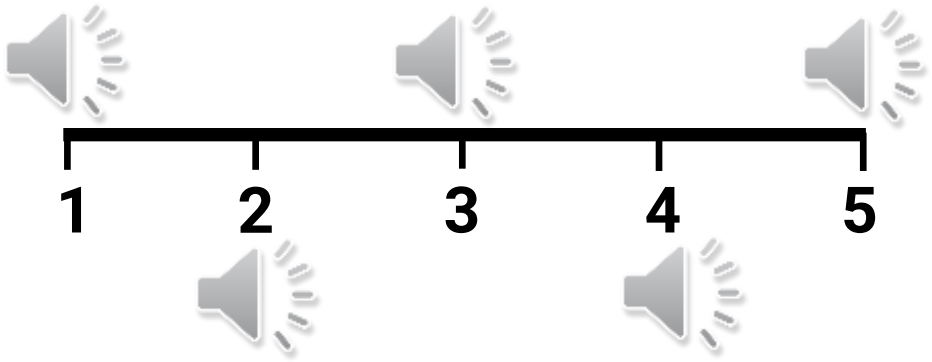
- True representation of dysarthria
- Real-world representation of how different dysarthria features manifest
- Must minimize extraneous perceptual features that are distracting
- May be achievable to various extents

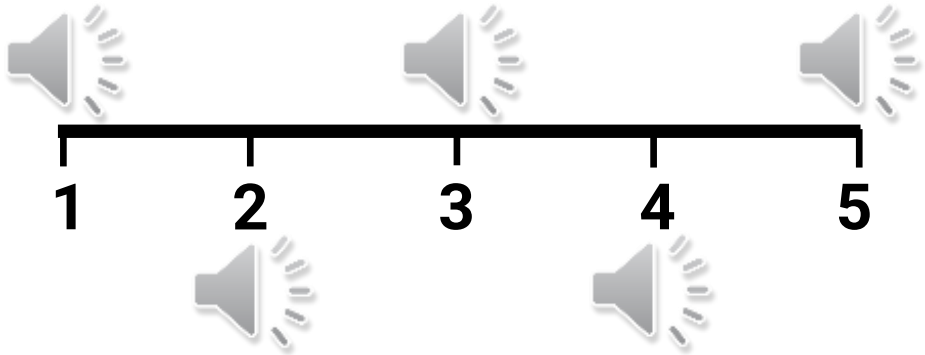
Synthetic Anchors

- Synthesized speech samples = Precise control over speech features
- Do not capture naturally occurring dysarthria
- Easy to illustrate isolated features

Anchored Scales

EAI with Anchors





IOWA

Use of Natural Anchors

- Natural anchors ≠ improve reliability
- High validity in the non-anchor and anchor conditions
- PD scores > Control scores for both anchor conditions

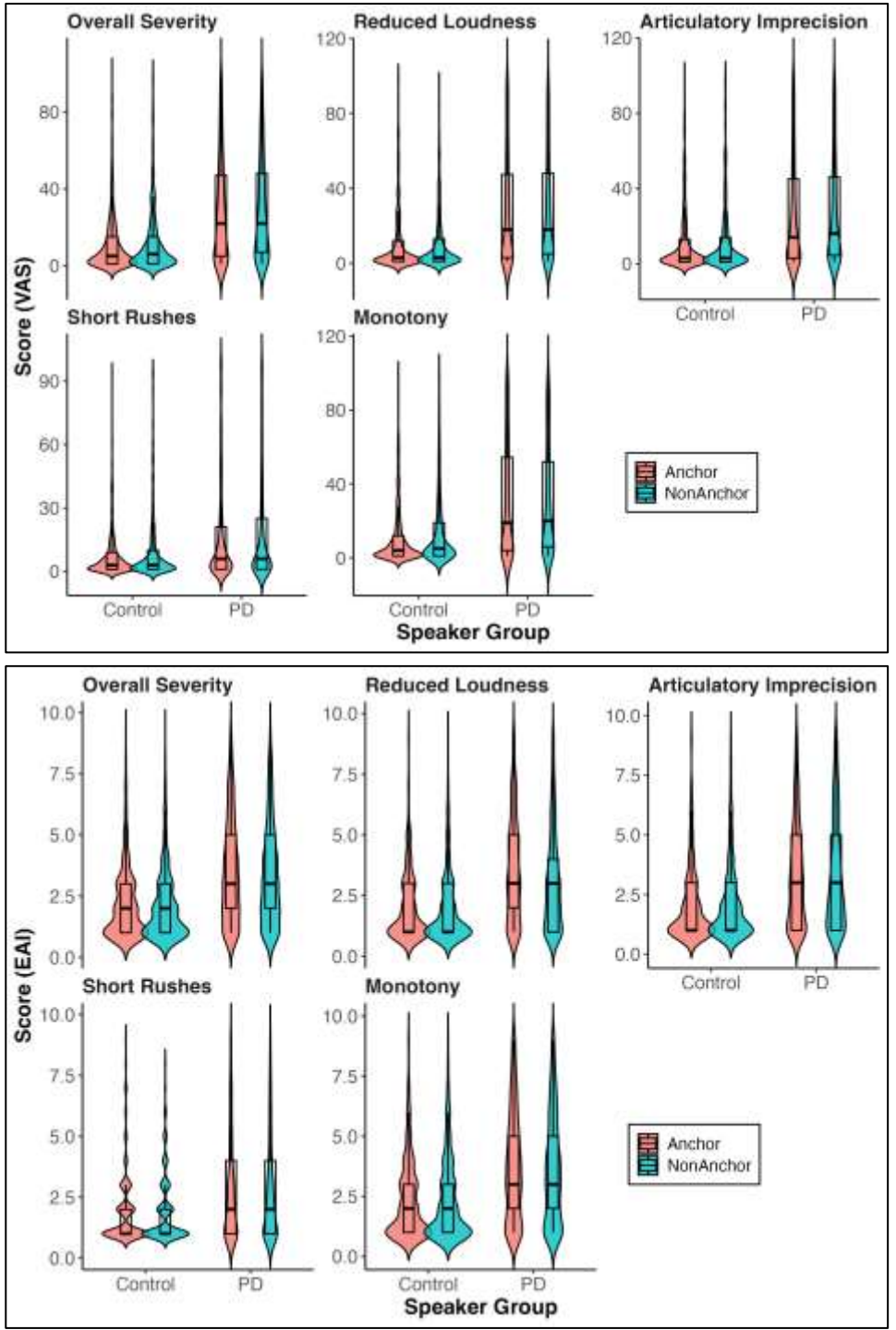
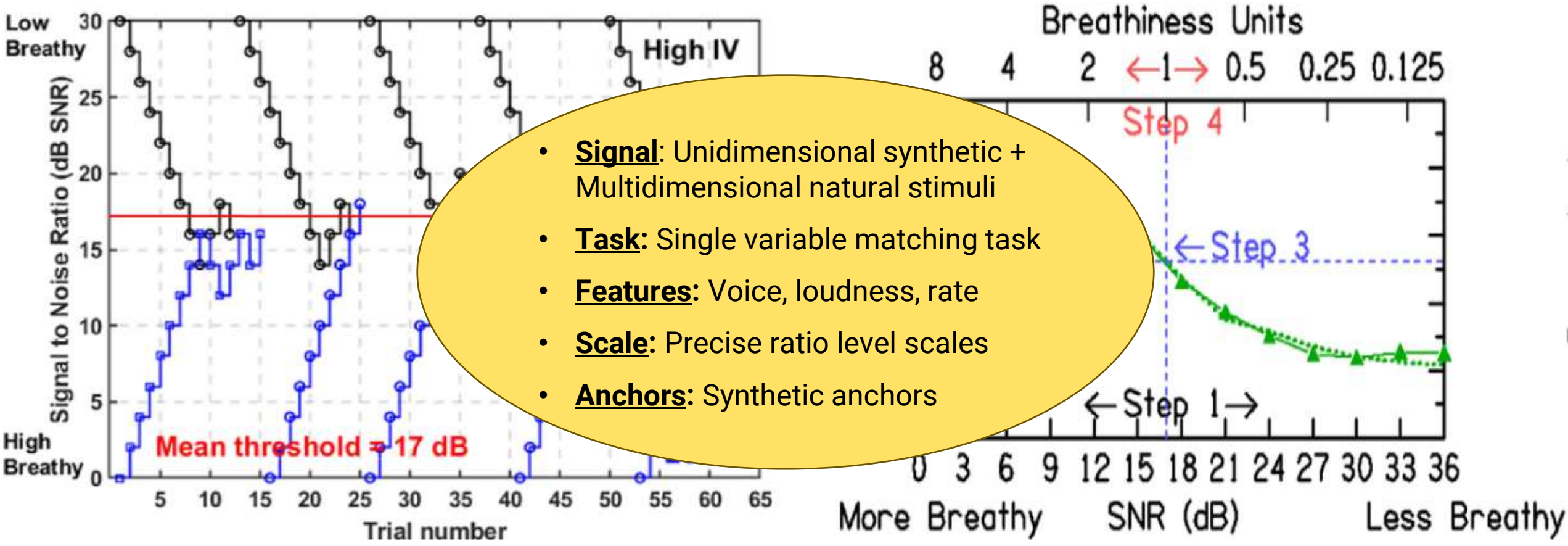


Figure 3 in: Crasta et al. (2025). *Anchored Interval and Visual Analog Scales: Impact on Reliability, Validity, and Scaled Scores for Dysarthria Assessment*. Manuscript in preparation.

Next Steps: Psychophysical Approach



- **Signal**: Unidimensional synthetic + Multidimensional natural stimuli
- **Task:** Single variable matching task
- **Features:** Voice, loudness, rate
- **Scale:** Precise ratio level scales
- **Anchors:** Synthetic anchors

Figures 1 & 6 in: Eddins et al. (2019). Developing clinically relevant scales of breathy and rough voice quality. *Journal of Voice*, *35*, 663.e9–663.e16.

Next Steps

- Identify optimal physical units for several dysarthria features
- Develop precise ratio-level scales for dysarthria features
  - Identify physical units for different features
  - Mapped the unit to scale intervals
- Determine the relationship between stimulus and perceptual magnitude
  - Help establish the input-output functions of the auditory-perceptual system for dysarthric speech