Multivariate Classification of Children with Speech Delay of Unknown Origin

Jennell Vick, Christopher A. Moore, Thomas Campbell, Lawrence Shriberg, Jordan Green & Klaus Truemper
jennell@utdallas.edu

November 19, 2009
American Speech-Language Hearing Association
New Orleans, LA

Speech Delay

- Prevalence
  - 15.6% among 3-year-olds (Campbell et al., 2003)
  - 3.8% among 6-year-olds (Shriberg et al., 1999)
- Sequelae
  - Residual speech sound errors (cf., Shriberg et al., 1997; Flipsen, 2003)
  - Lack of increase in number of syllables per word (Flipsen, 2006)
  - Later difficulites with reading, spelling and literacy throughout life (Felsenfeld et al., 1992; Larrivee & Catts, 1999; Raitano et al., 2004; Shriberg et al., 1999)

Current Speech Disorders Classification System (SDCS) Categories (Shriberg et al., 2009)

No.	Type	Subtype	Abbreviation	Risk Factors	Processes Affected
1	Speech Delay	Speech Delay–Genetic	SD-GEN	Polygenic/ Environmental	Cognitive-Linguistic
2		Speech Delay– Otitis Media with Effusion	SD-OME	Polygenic/ Environmental	Auditory-Perceptual
3		Speech Delay– Developmental Psychosocial Involvement	SD-DPI	Polygenic/ Environmental	Affective-Temperamental
4	Motor Speech Disorder	Motor Speech Disorder– Apraxia of Speech	MSD-AOS	Monogenic? Oligogenic?	Speech-Motor Control
5		Motor Speech Disorder– Dysarthria	MSD-DYS	Monogenic? Oligogenic?	Speech-Motor Control
6		Motor Speech Disorder-Not Otherwise Specified	MSD-NOS	Monogenic? Polygenic? Oligogenic? Environmental?	Speech-Motor Control
7	Speech Errors	Speech Errors-Sibilants	SE-/s/	Environmental	Phonological Attunement
8		Speech Errors-Rhotics	SE-/r/	Environmental	Phonological Attunement

160 Participants

97 children with idiopathic speech delay (SD)

- 66 males; 34 females
- ID'd as SD by referring SLP or classified as such by SDCS
- Mean age: 46.02 months Range: 35 – 59 months

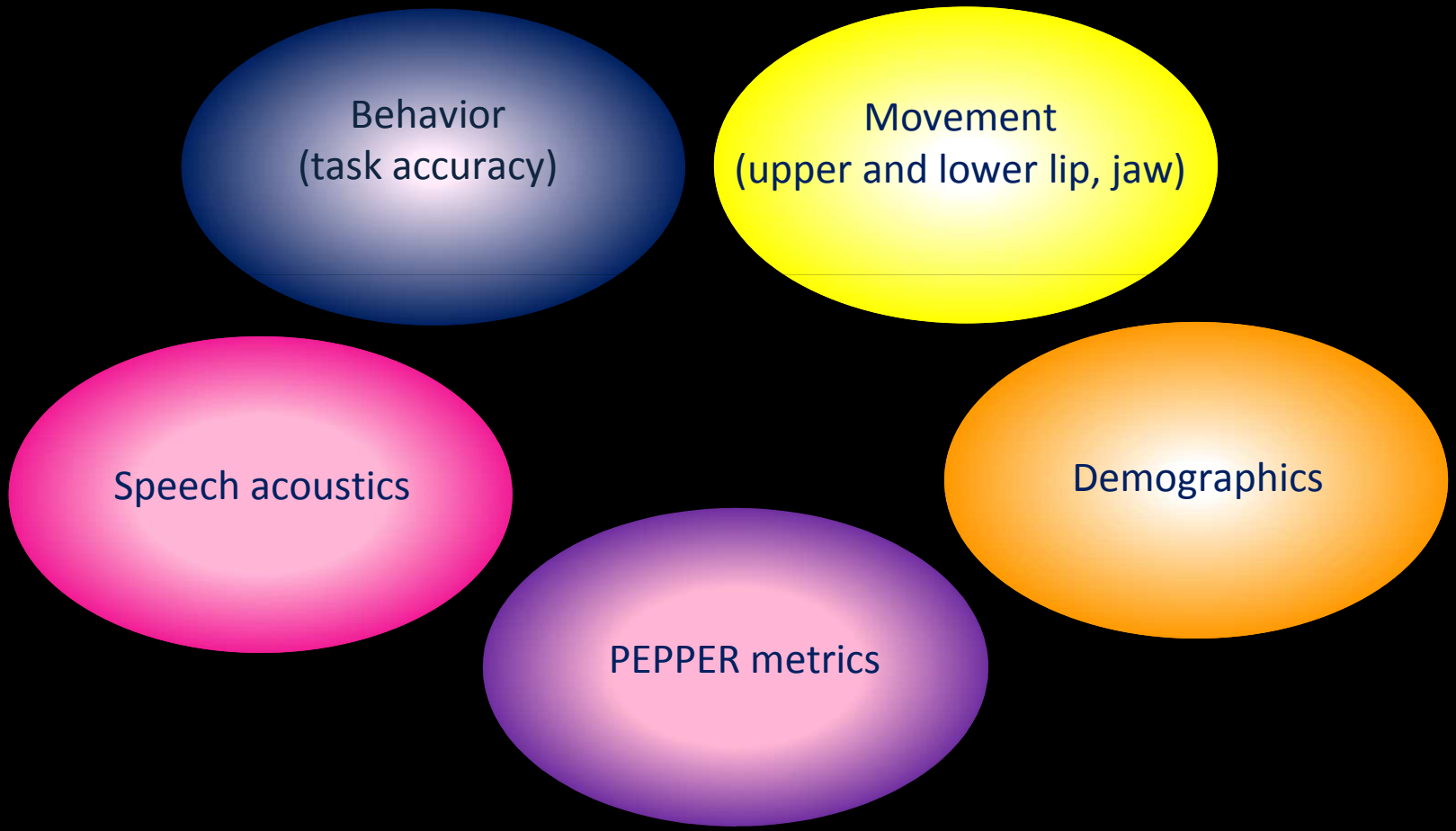
63 children with typical speech acquisition used as a reference while plotting discovered subgroups

- 27 males; 36 females
- Mean age: 46.13 months Range: 36 – 57 months

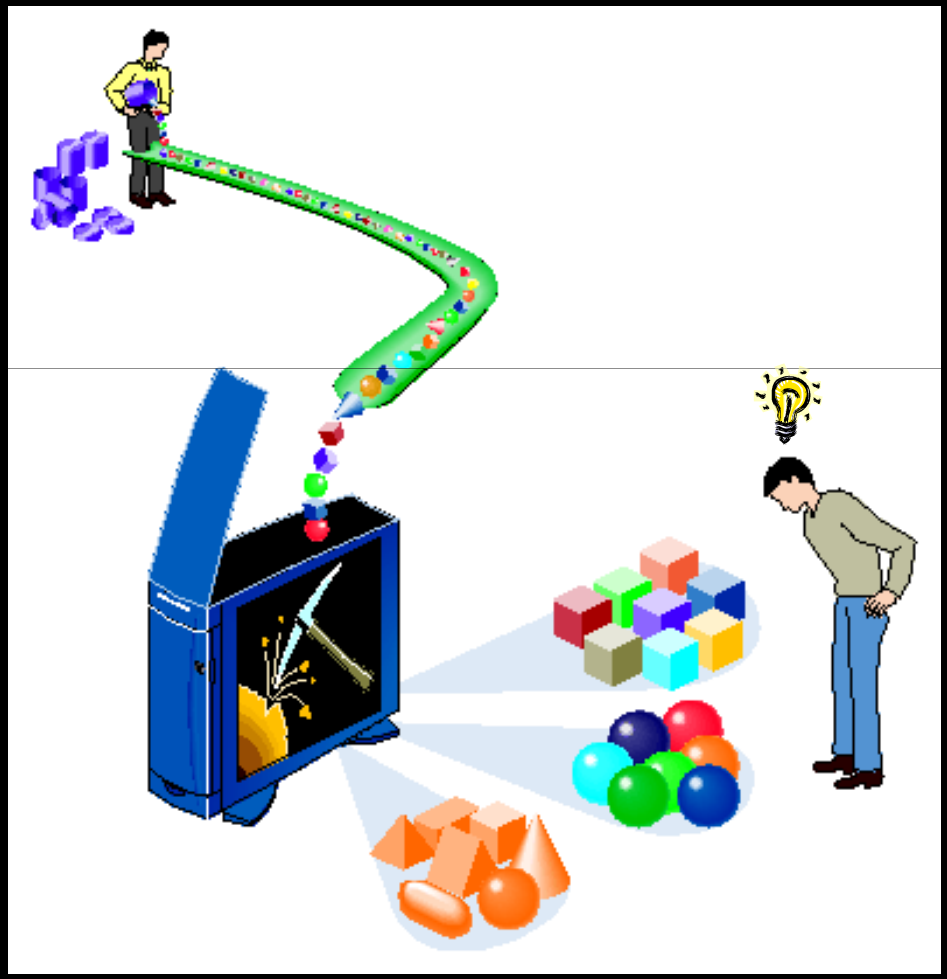
12 Tasks

- Contrastive lexical stress task
  - Trochaic and iambic stress
  - /baba/, /mama/, /papa/
- Syllable repetition task (SRT; Shriberg, Lohmeier, Campbell, Dollaghan, Green, and Moore, 2009; Shriberg, Lohmeier, Dollaghan, & Campbell, 2006)
  - /bama/ , /bada/ , /bamana/ , /manaba/
- Nonspeech tasks
    - Chewing, vertical jaw oscillation

75 Measures


Behavior
(task accuracy)
Movement
(upper and lower lip, jaw)
Speech acoustics
Demographics
PEPPER metrics

Subgroup Discovery



EXARP Algorithm

Truemper, 2009; Moreland & Truemper, 2009

- The subgroups are described by patterns or rules that relate an *explaining variable* to *target variables*
- Significance is determined by the exclusivity of the group
  - Training and testing sets

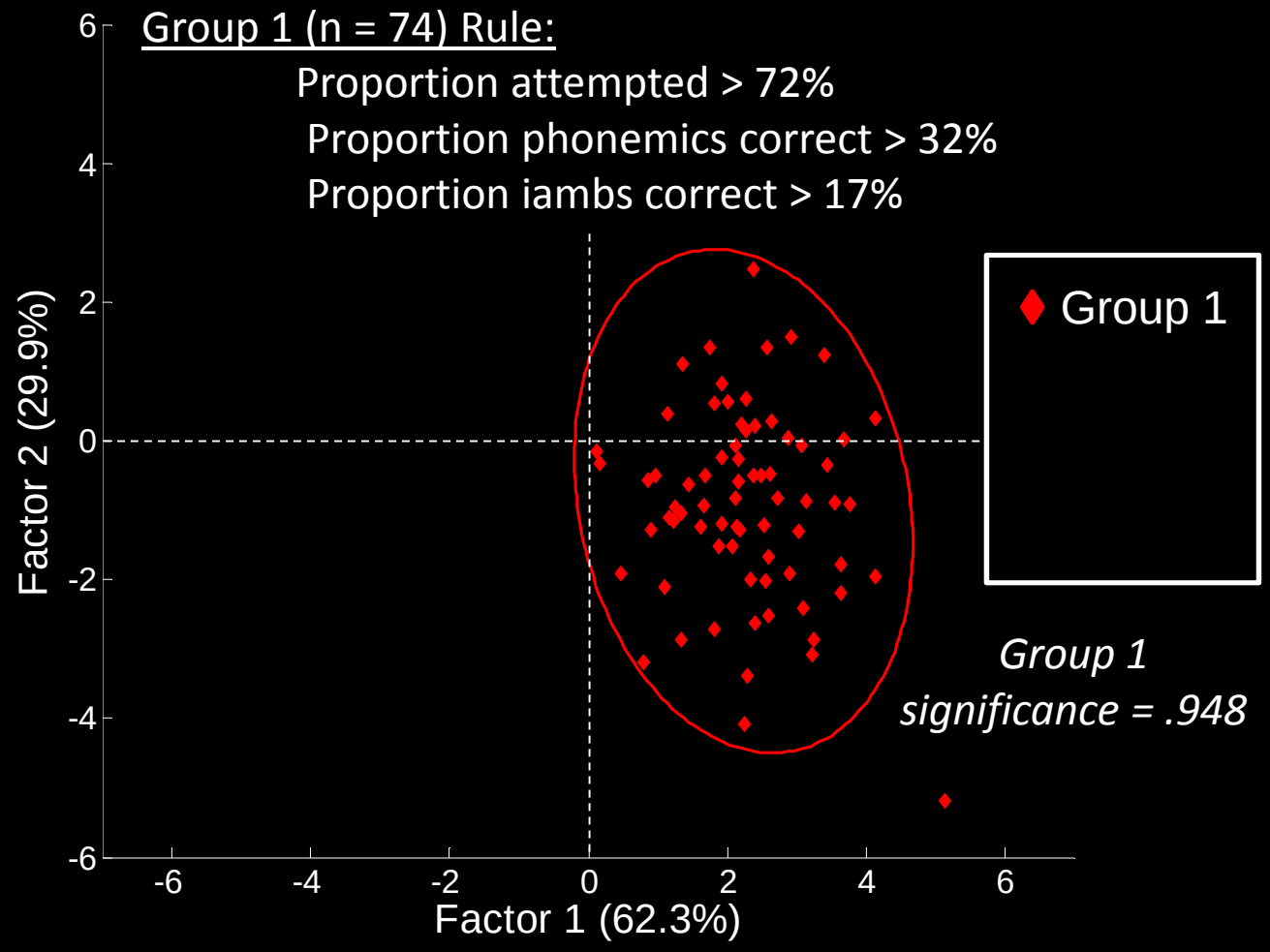
Goals of Subgroup Discovery

Our primary goal is to identify children who could be classified as *Motor Speech Disorder, Not Otherwise Specified (MSD-NOS)* on the SDCS

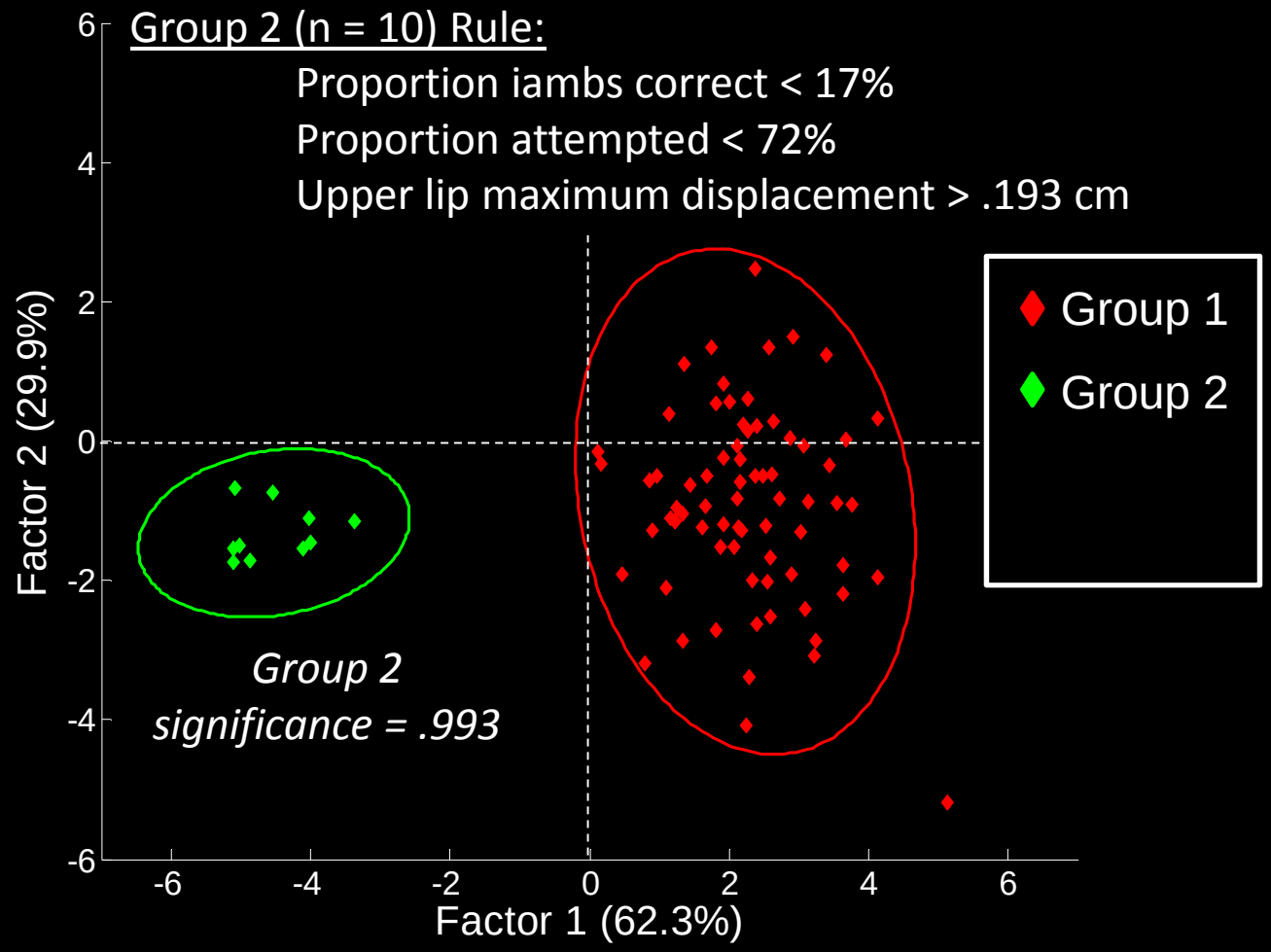
Additional benefits:

- Target intervention goals
- Predict prognosis
- Refine phenotypes for genetic studies

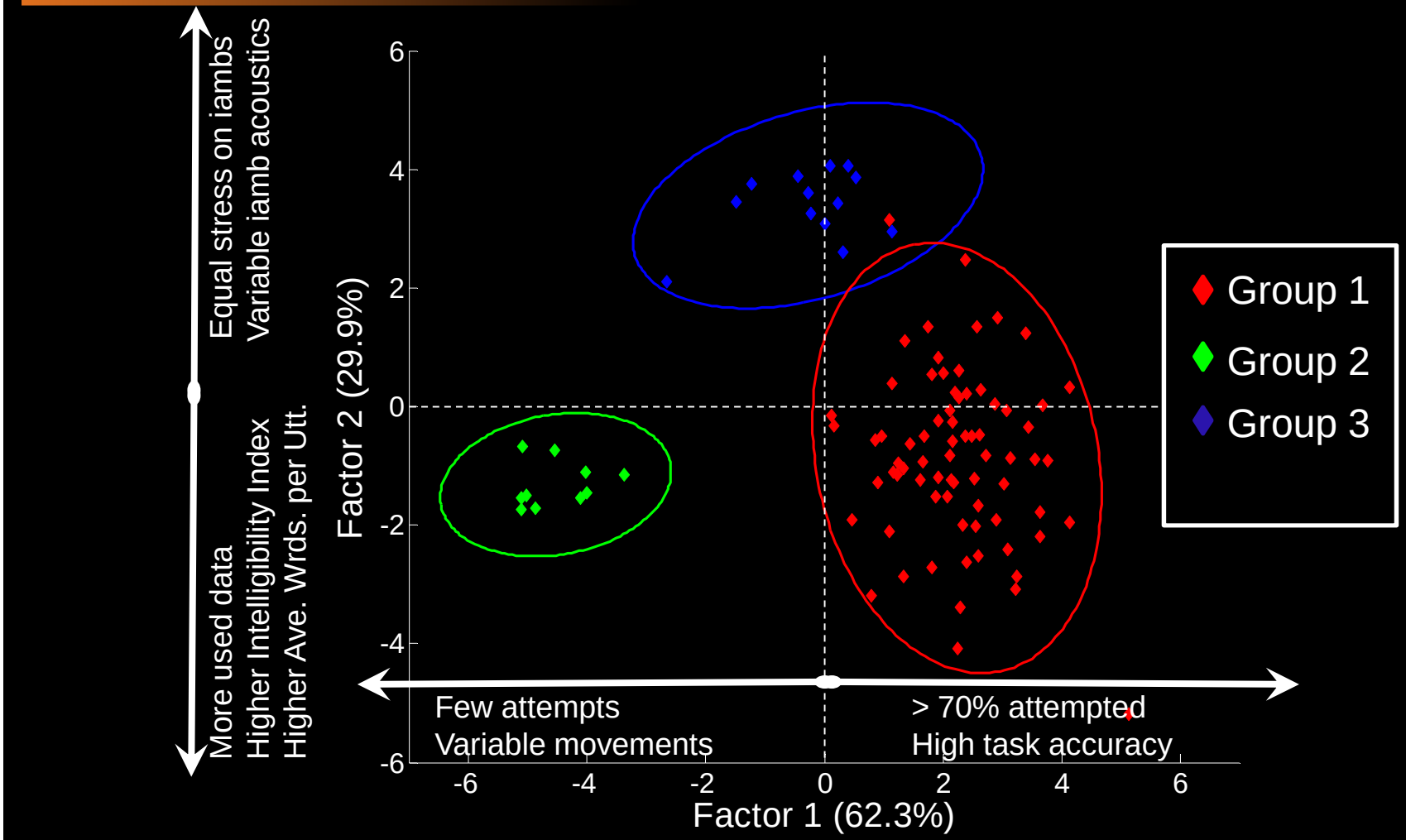
EXARP Results in 2D Discriminant Space



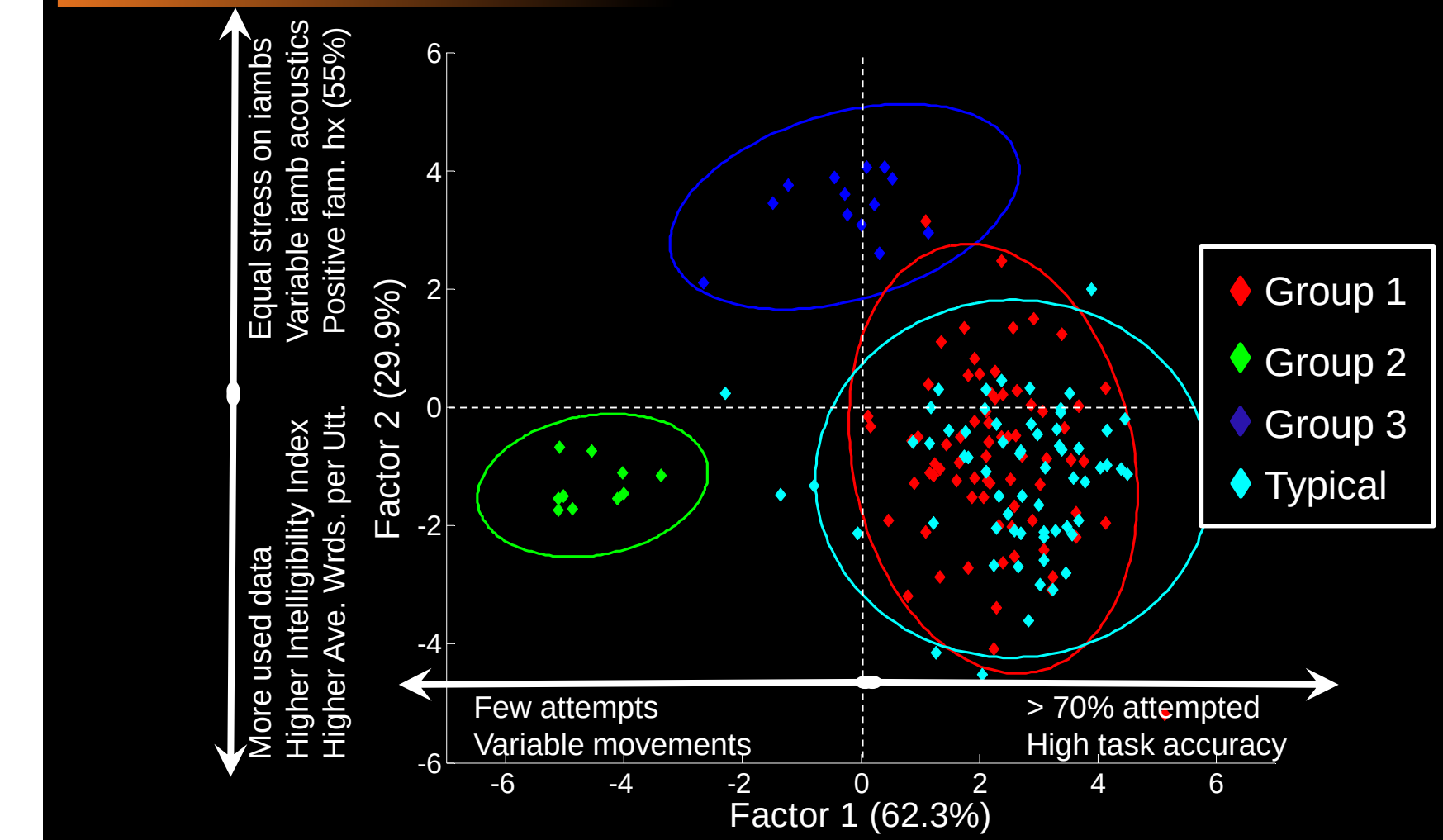
EXARP Results in 2D Discriminant Space



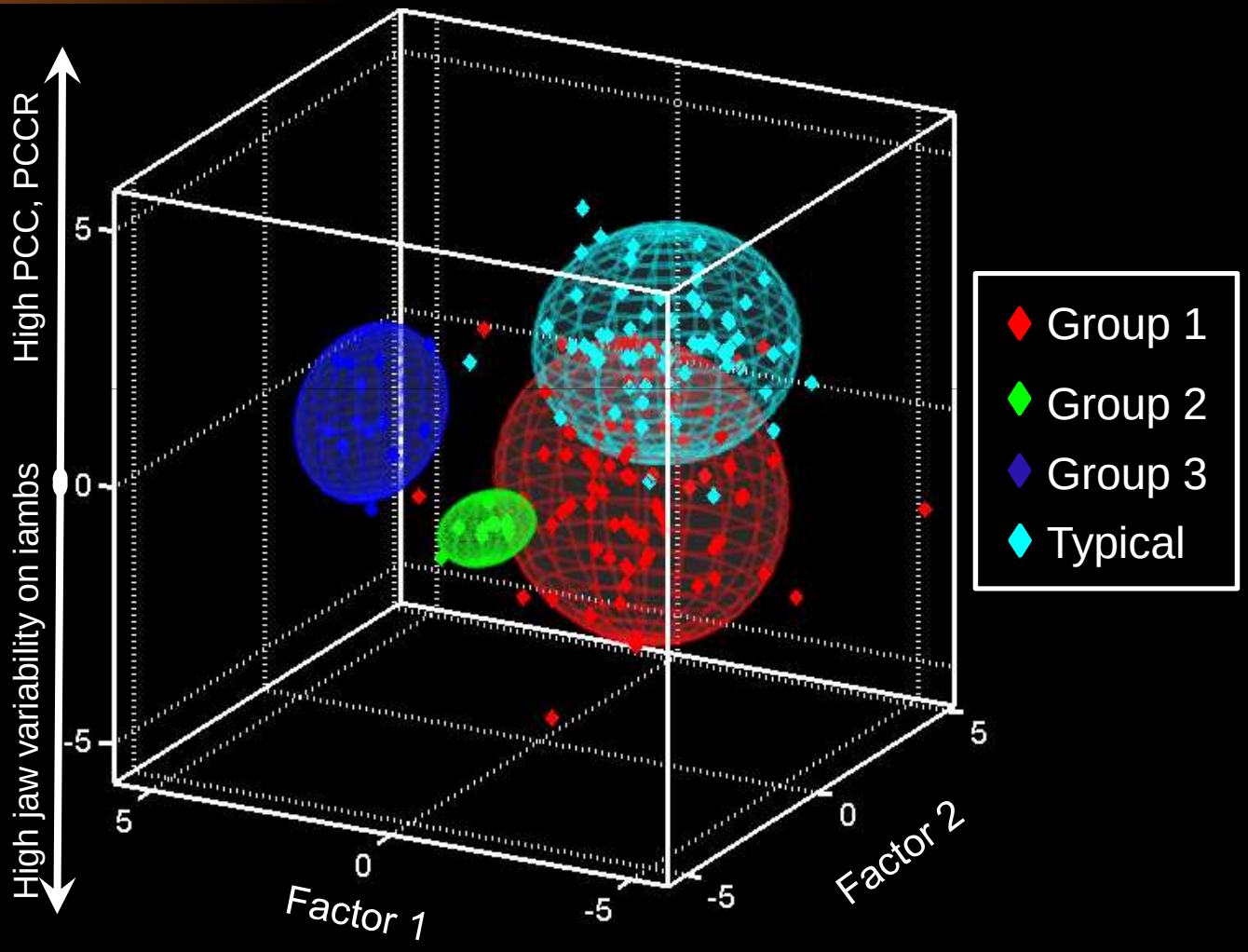
EXARP Results in 2D Discriminant Space



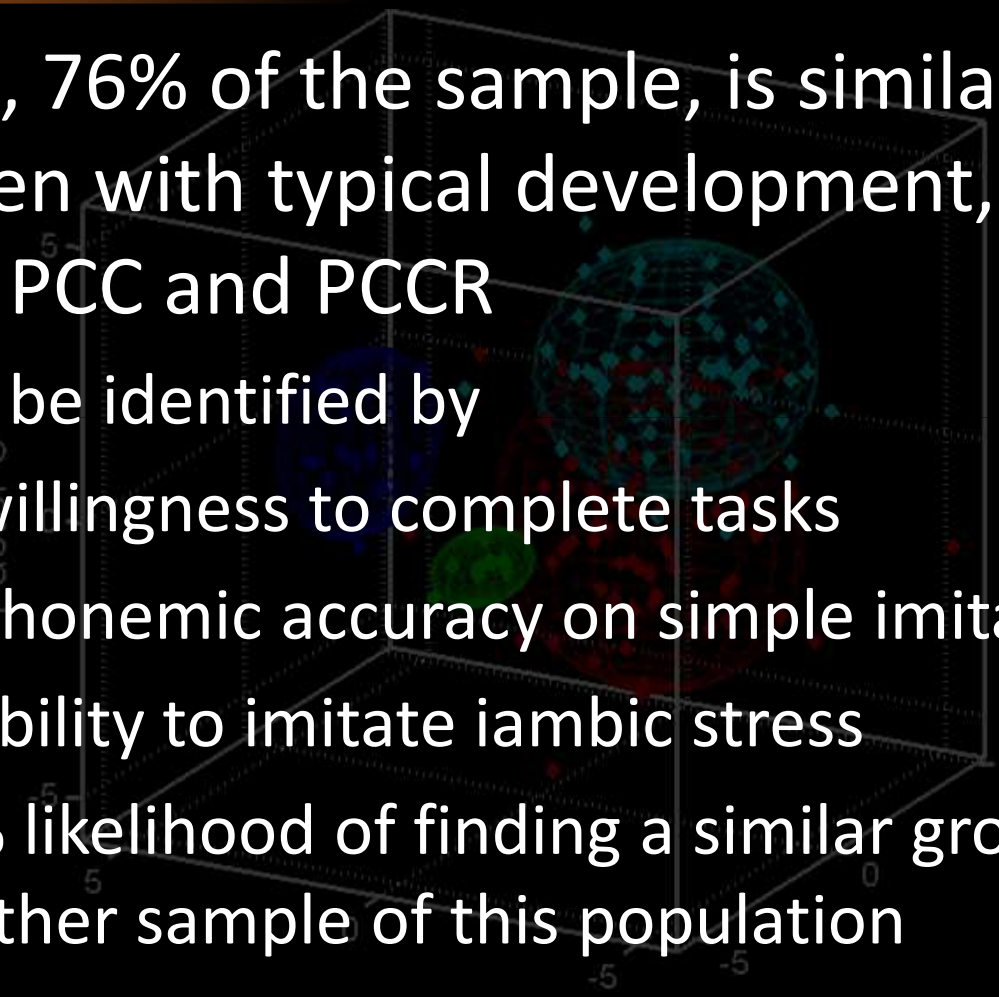
EXARP Results in 2D Discriminant Space



EXARP Results in 3D Discriminant Space



Conclusions



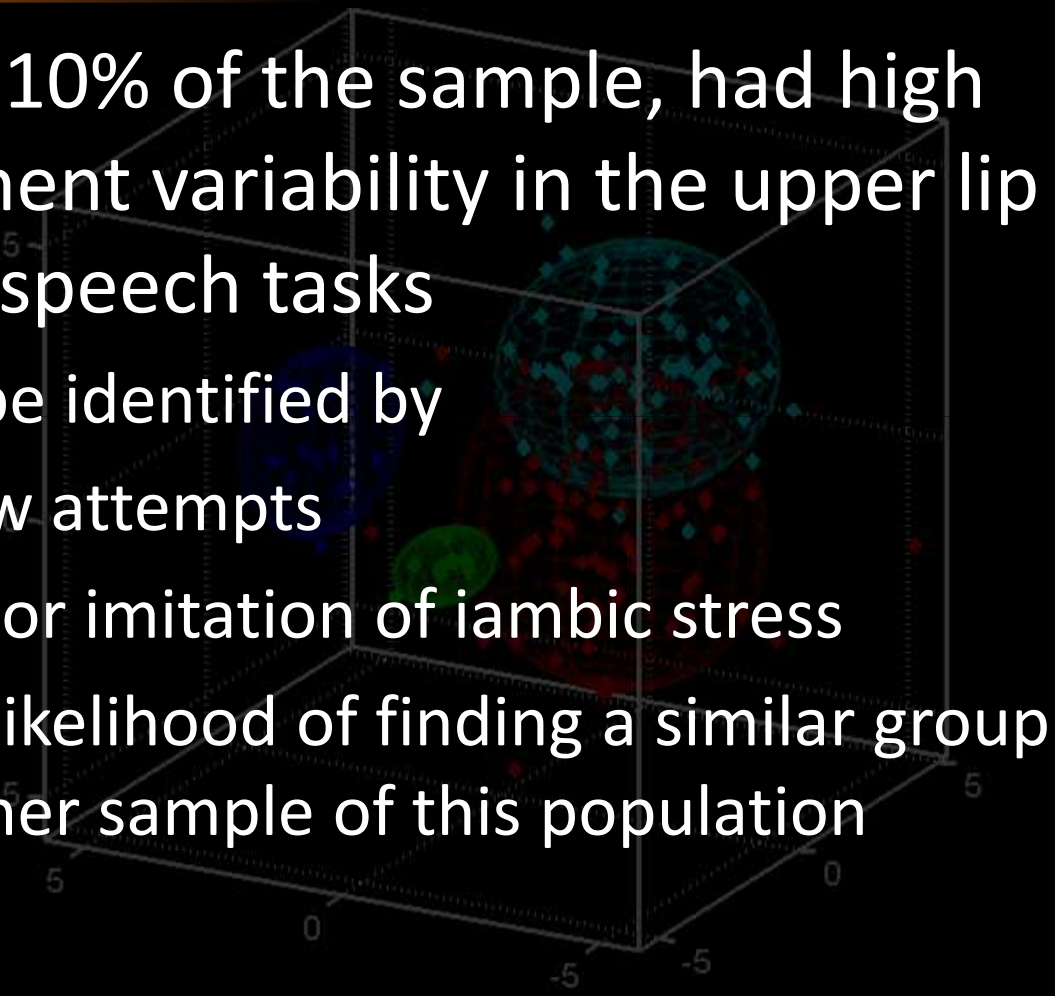
Group 1, 76% of the sample, is similar to children with typical development, except for lower PCC and PCCR

- Can be identified by
  - willingness to complete tasks
  - phonemic accuracy on simple imitation
  - ability to imitate iambic stress
- 95% likelihood of finding a similar group in another sample of this population

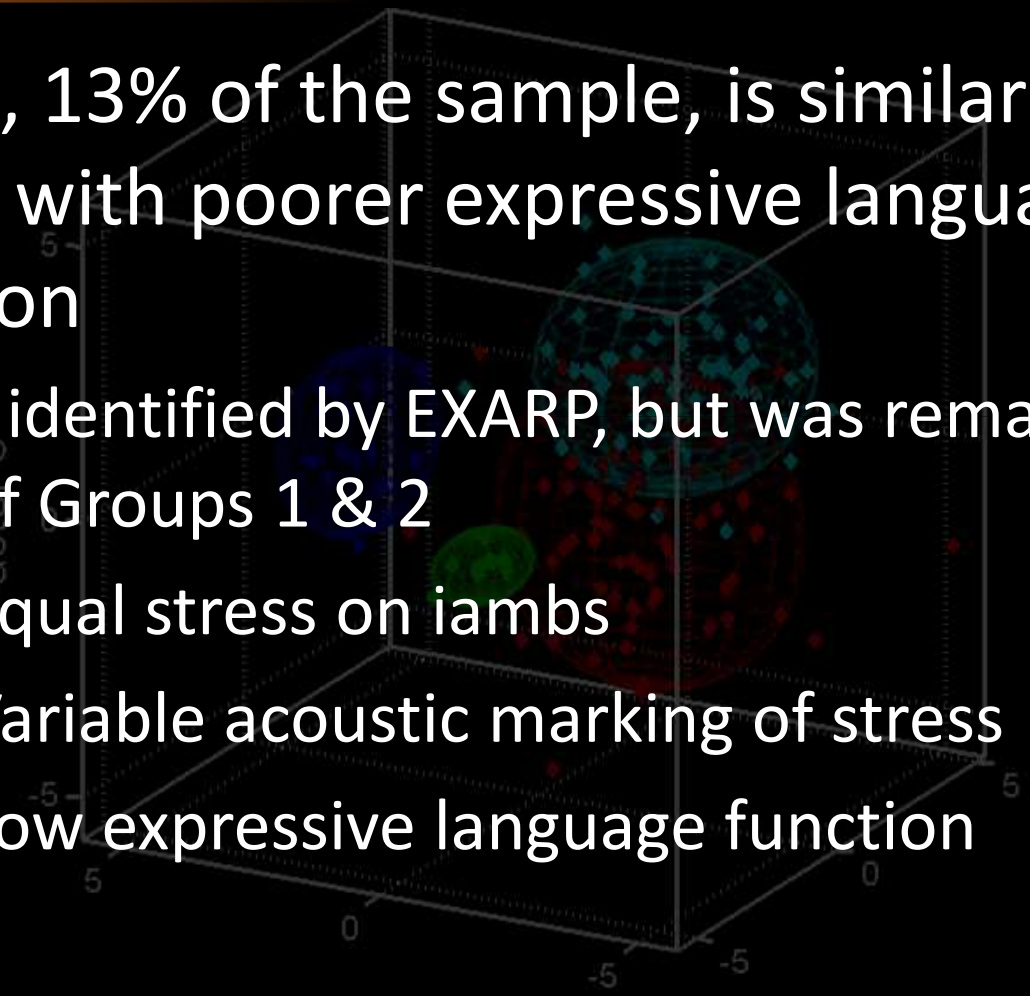
Conclusions

Group 2, 10% of the sample, had high movement variability in the upper lip and jaw during speech tasks

- Can be identified by
  - few attempts
  - poor imitation of iambic stress
- 99% likelihood of finding a similar group in another sample of this population



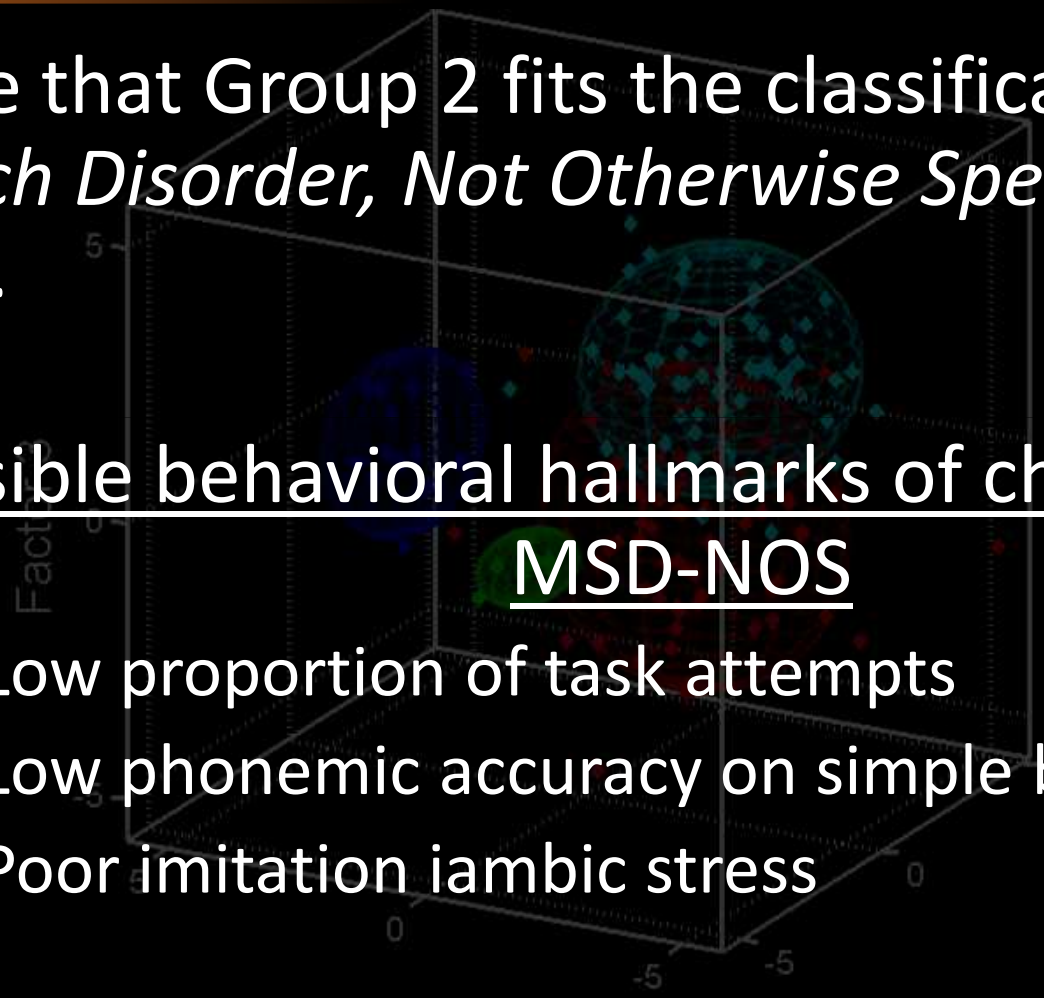
Conclusions



- Group 3, 13% of the sample, is similar to Group 2, but with poorer expressive language function
  - Not identified by EXARP, but was remaining after ID of Groups 1 & 2
    - Equal stress on iambs
    - Variable acoustic marking of stress
    - Low expressive language function

Conclusions

Propose that Group 2 fits the classification *Motor Speech Disorder, Not Otherwise Specified* on the SDCS.



<u>Possible behavioral hallmarks of children with MSD-NOS</u>

- Low proportion of task attempts
- Low phonemic accuracy on simple bisyllables
- Poor imitation iambic stress

Acknowledgments

Funding: NIDCD R01 DC00822

Hervé Abdi

Greg Lee

Lakshmi Venkatesh

Michelle Yee

Kristie Spencer

Data Acquisition Laboratory in the Center for Research on Acquired and Neurologic Communication Disorders in Children at Children's Hospital of Pittsburgh

Heather Leavy Rusiewicz, Tammy Nash, Jill Brady, Dayna Pitcairn

Developmental Speech Physiology Lab at the University of Washington

Yumi Sumida, Alyssa Mosely, Jeanette Wu, Jenny Morus, Kelsey Moore, Mary Reeves, Nicholas Moon, Dennis Tang, Adam Politis, Andrea Kettler, Laura Worthen

Katie Moreland

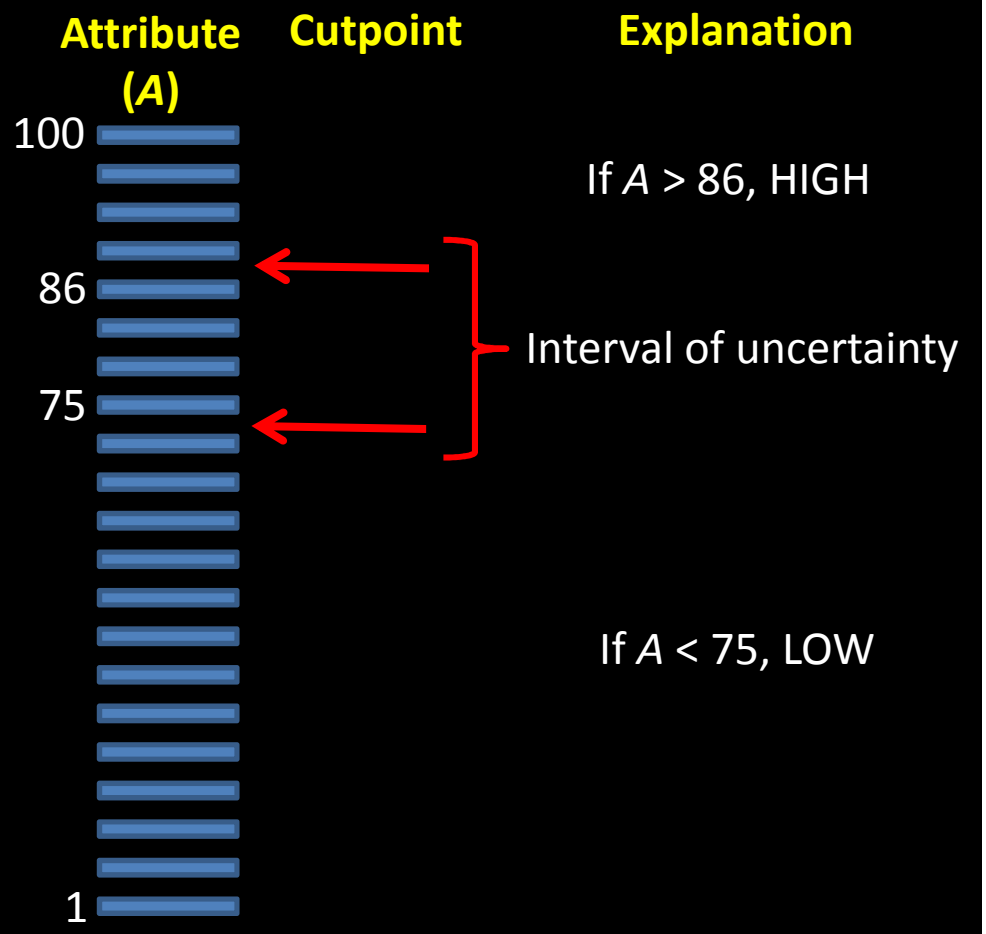
References

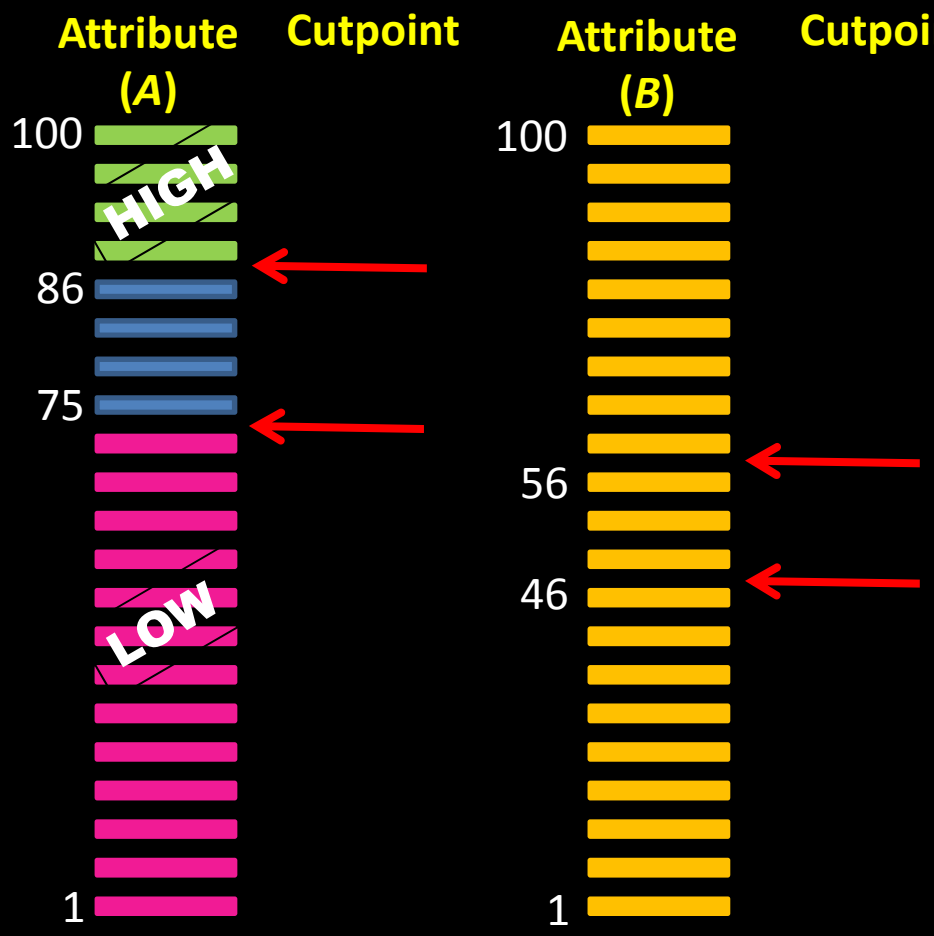
K. Truemper, "Improved Comprehensibility and Reliabity of Explanations via Restricted Halfspace Discretization,'‘ Proceedings of International Conference on Machine Learning and Data Mining (MLDM 2009) (2009) 1--15.

K. Moreland and K. Truemper, "Discretization of Target Attributes for Subgroup Discovery,'‘ Proceedings of International Conference on Machine Learning and Data Mining (MLDM 2009) (2009) 44-52.

Discretization: For each variable (*attribute*) a cutpoint is identified; The precision of this point is established by comparing it to alternate random processes. For instance, values above this point form a cohesive group, but not all values below this point can necessarily be excluded– there are middle points that supply some uncertainty or that cannot be definitively classified after comparison to random processes.

Alternate Random Processes: May make unimportant contrasts seem important; Use these as benchmarks against which to compare discretization results




Attribute (A)
Cutpoint
Attribute (B)
Cutpoint
Explanation
100
HIGH
86
75
LOW
1
100
56
46
1
If B > 56, HIGH
If B < 46, LOW
• 8 of 13 A-LOW fit this
• 0 of 4 A HIGH fit this