

# Crowdsourcing Audio Annotations: Findings and Next Steps

---

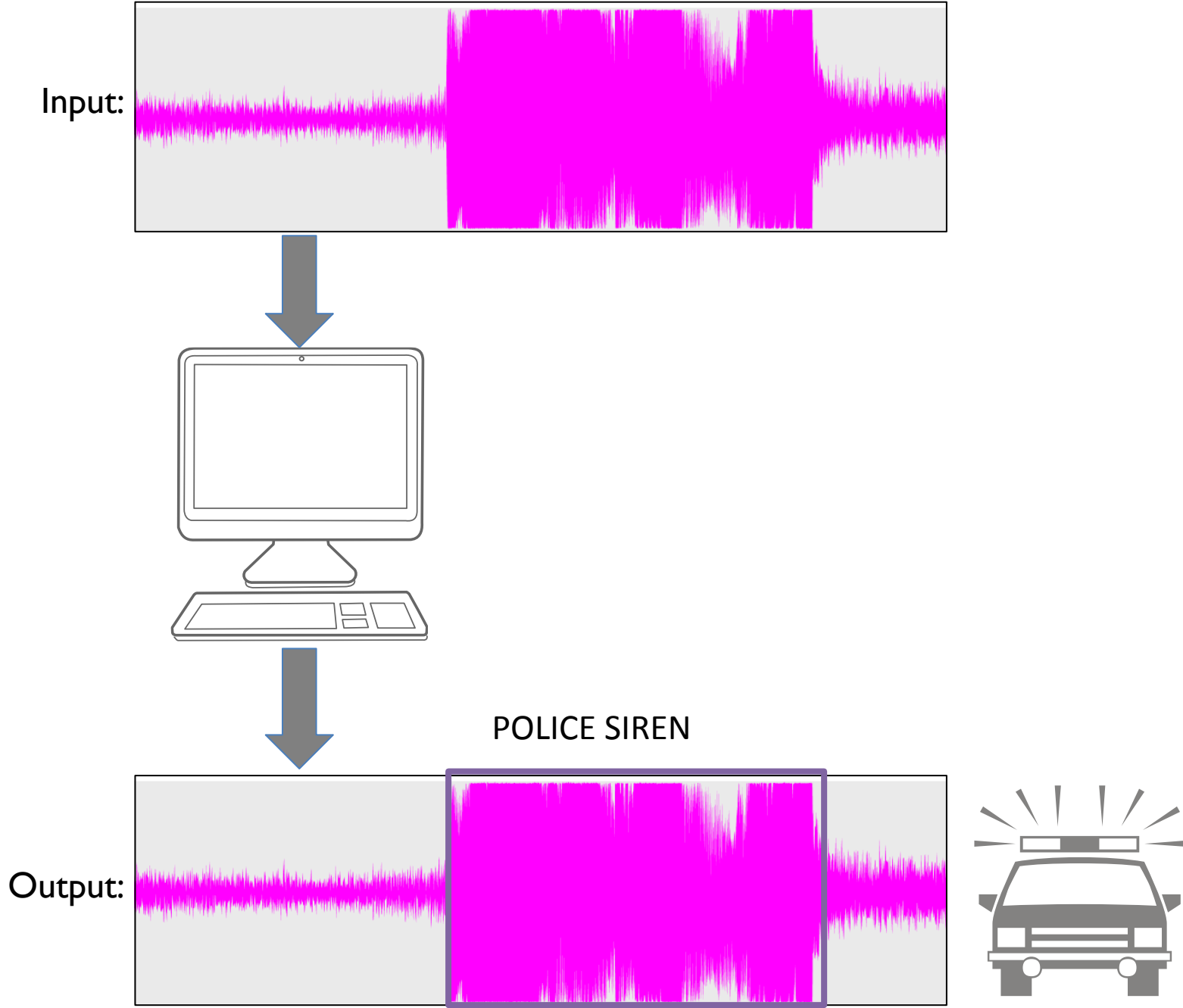
Mark Cartwright, Justin Salamon, Ayanna Seals, Ana Elisa Mendez Mendez,  
Graham Dove, Juan Bello, Oded Nov

Mark Cartwright  
Postdoctoral Researcher  
Music and Audio Research Laboratory



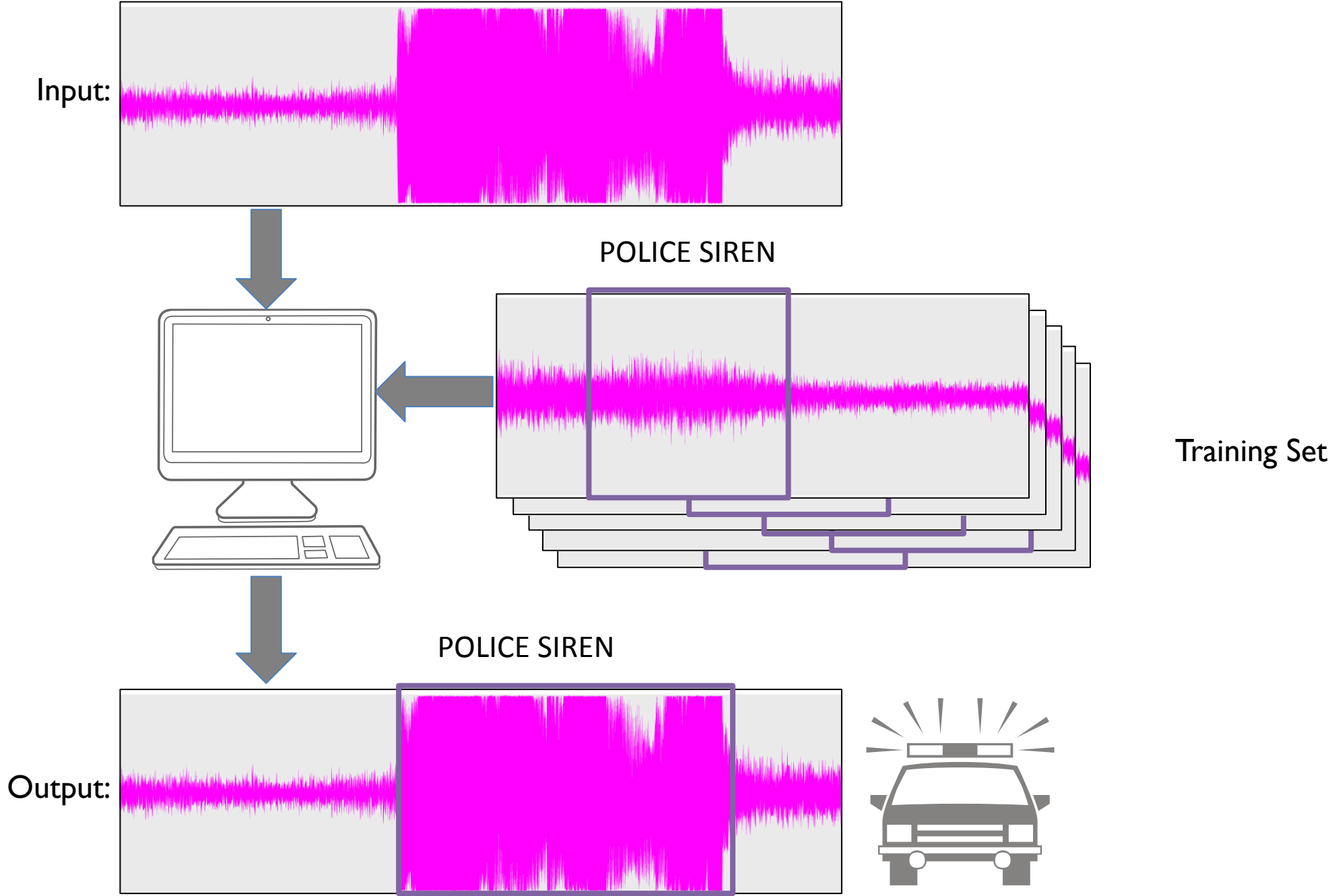
# Audio Annotation of Sound-Event Detection

---



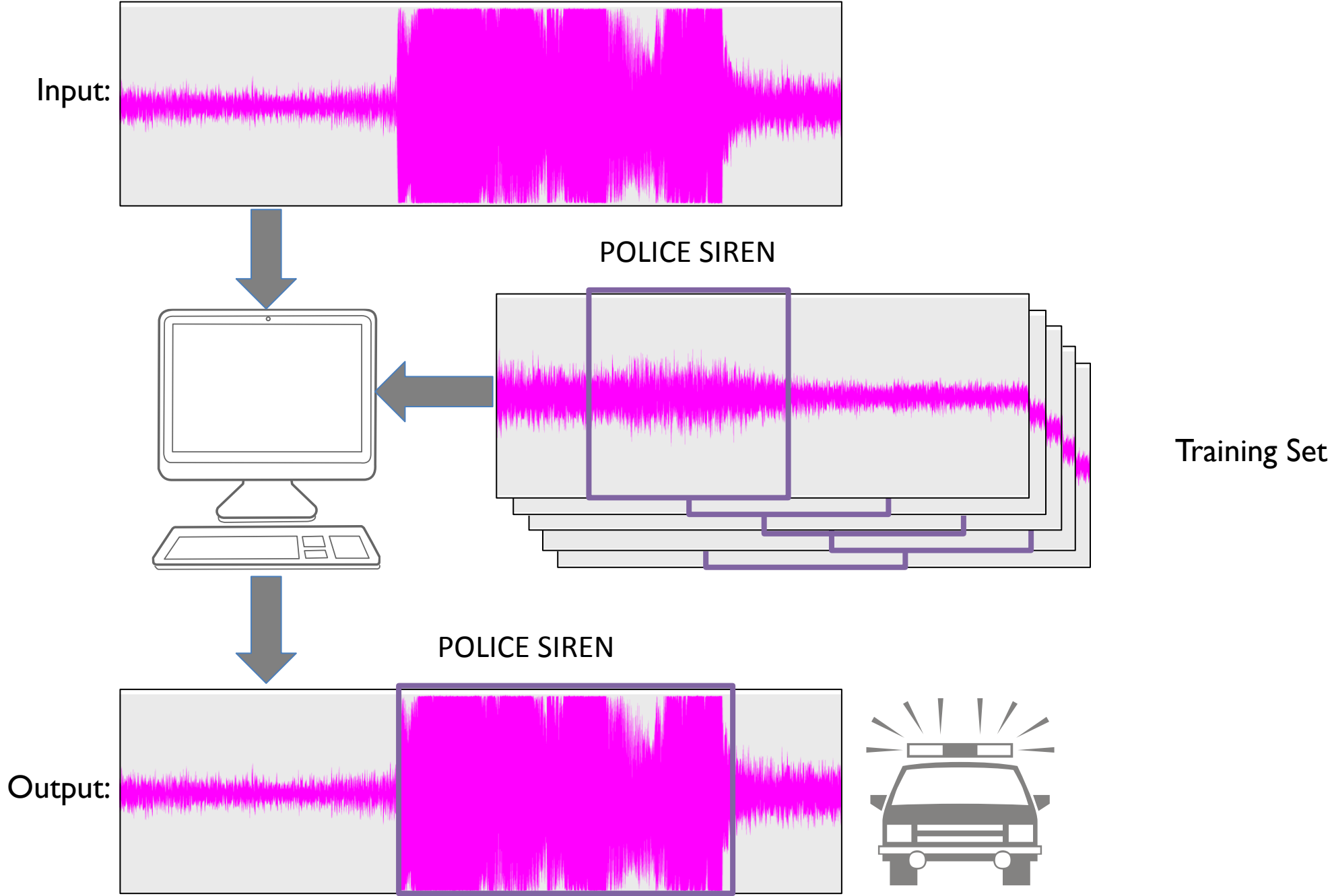
# Audio Annotation of Sound-Event Detection

---



# Audio Annotation of Sound-Event Detection

---

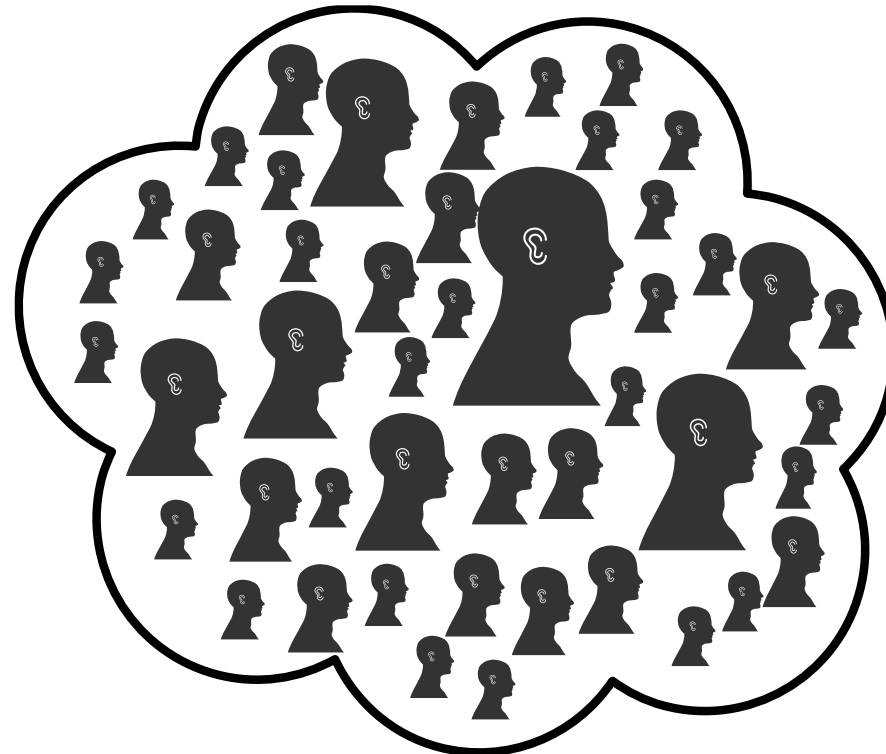




# Research Questions

---

- Which sound visualization aid yields the highest quality crowdsourced audio annotations?
- What limitations can we expect from crowdsourced audio annotations as a function of soundscape complexity?



# The Audio Annotator

---

Configured with the spectrogram visualization:

The screenshot displays the Audio Annotator interface. At the top, there are two tabs: 'SIREN WAILING' and 'ENGINE IDLING'. The 'ENGINE IDLING' tab is active, and its corresponding spectrogram segment is highlighted with a red border. The spectrogram shows frequency content over time, with a prominent horizontal band of energy in the lower-mid frequency range. Below the spectrogram, a play button is visible on the left, and the current time is shown as '00:10.796 / 00:10.796'. The segment's metadata is displayed as 'Start: 00:03.923 End: 00:08.198 Duration: 00:04.276'. Underneath, there are two rows of labels. The first row, labeled 'Label:', includes 'CAR HORN HONKING', 'DOG BARKING', 'ENGINE IDLING' (which is highlighted in red), 'GUN SHOOTING', 'JACK HAMMER DRILLING', and 'MUSIC PLAYING'. The second row, labeled 'The sound is:', includes 'NEAR', 'FAR', and 'NOT SURE'.

# The Audio Annotator

---

Configured with the waveform visualization:

The screenshot displays an audio annotation interface. At the top, a dark grey bar contains the label 'SIREN WAILING' with a play icon on the left and a close icon on the right. Below this, a teal bar highlights the 'ENGINE IDLING' segment, also with play and close icons. The main area shows a pink waveform on a grey background. A blue play button is located at the bottom left. The bottom right corner shows the time '00:10.796 / 00:10.796'. Below the waveform, the following metadata is displayed: 'Start: 00:03.923', 'End: 00:08.198', and 'Duration: 00:04.276'. Underneath, there are two rows of label buttons. The first row includes 'CAR HORN HONKING', 'DOG BARKING', 'ENGINE IDLING' (highlighted in teal), 'GUN SHOOTING', 'JACK HAMMER DRILLING', and 'MUSIC PLAYING'. The second row includes 'PEOPLE SHOUTING', 'PEOPLE TALKING', and 'SIREN WAILING'. At the bottom left, the text 'The sound is:' is followed by three buttons: 'NEAR', 'FAR', and 'NOT SURE'.

Start: 00:03.923    End: 00:08.198    Duration: 00:04.276

Label:

- CAR HORN HONKING
- DOG BARKING
- ENGINE IDLING**
- GUN SHOOTING
- JACK HAMMER DRILLING
- MUSIC PLAYING

The sound is:

- NEAR
- FAR
- NOT SURE

# The Audio Annotator

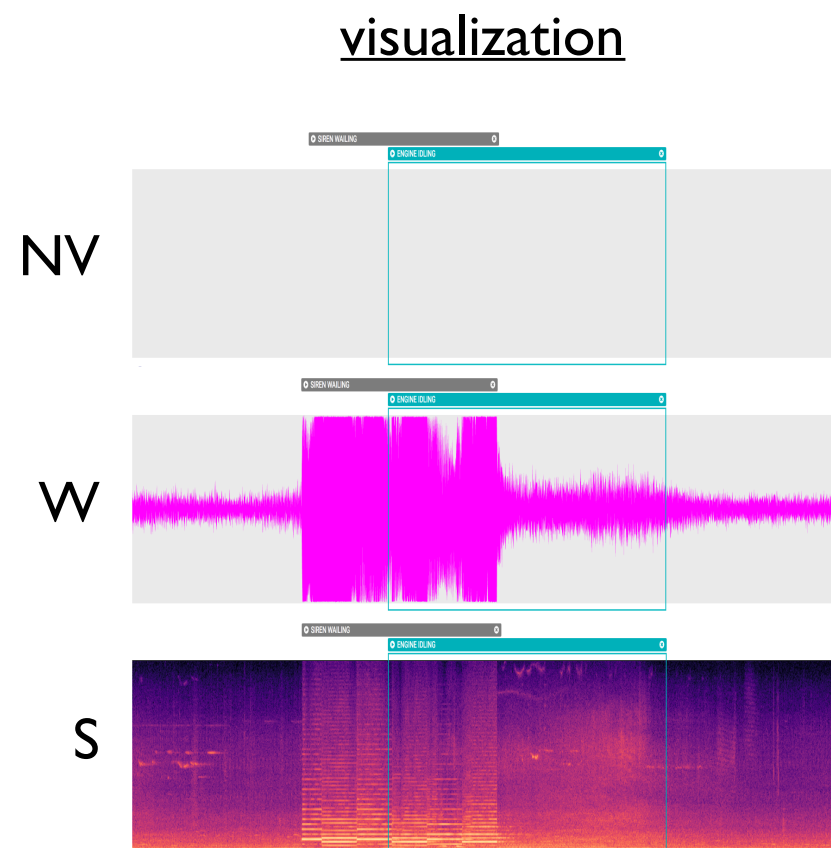
---

Configured without a visualization:

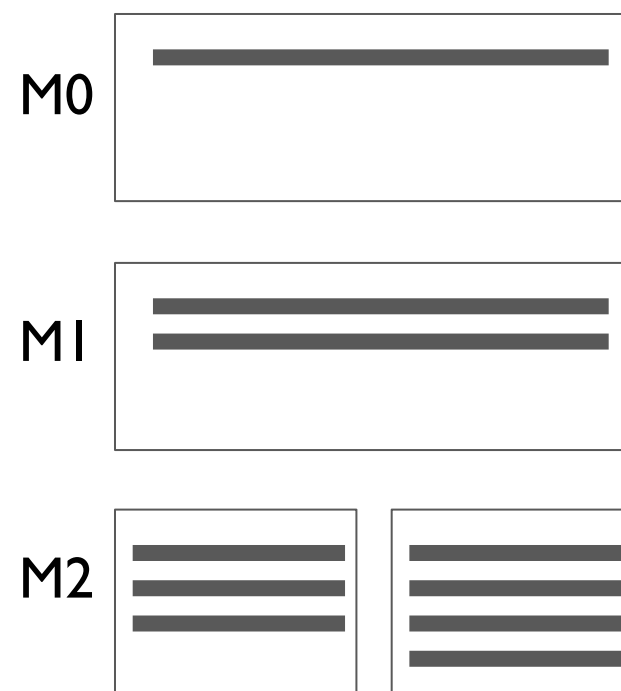
The screenshot displays the Audio Annotator interface. At the top, there are two floating labels: "SIREN WAILING" and "ENGINE IDLING". The "ENGINE IDLING" label is highlighted with a teal border. Below the labels is a large grey rectangular area representing the audio waveform, which is currently empty. A play button is located at the bottom left. The timeline at the bottom shows the current time as 00:10.796 / 00:10.796. The selected segment has a start time of 00:03.923, an end time of 00:08.198, and a duration of 00:04.276. The "Label:" section contains several buttons: "CAR HORN HONKING", "DOG BARKING", "ENGINE IDLING" (highlighted), "GUN SHOOTING", "JACK HAMMER DRILLING", and "MUSIC PLAYING". The "The sound is:" section contains buttons for "NEAR", "FAR", and "NOT SURE".

# Experiment

- 3 x 3 x 2 between-subjects factorial design:



max-polyphony



gini-polyphony



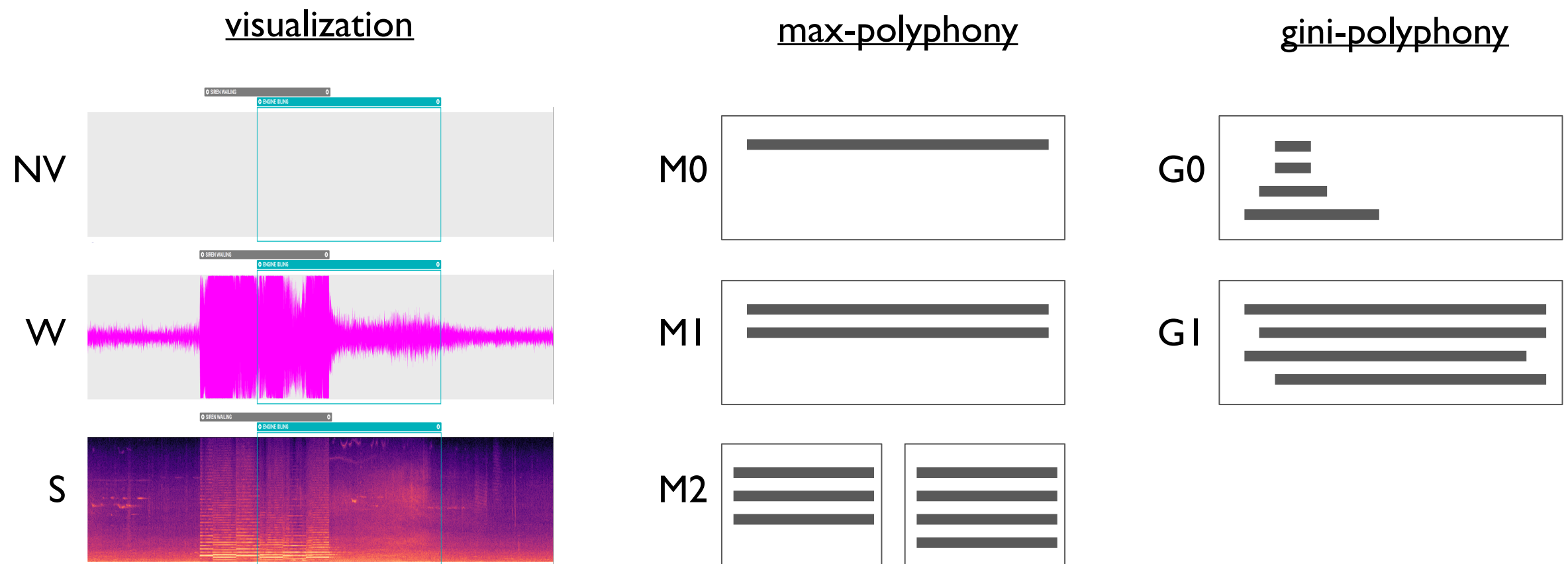
- Soundscape examples:  
M0G0                      M0G1

M2G0

M2G1

# Experiment

- 3 x 3 x 2 between-subjects factorial design:



- Soundscape examples:

M0G0

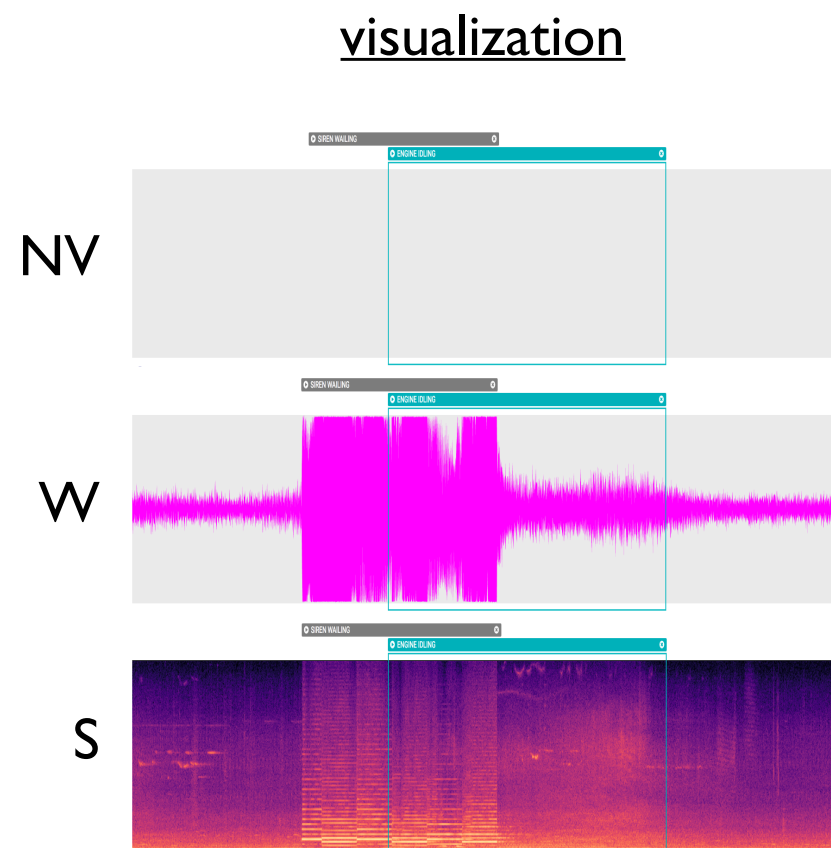
M0G1

M2G0

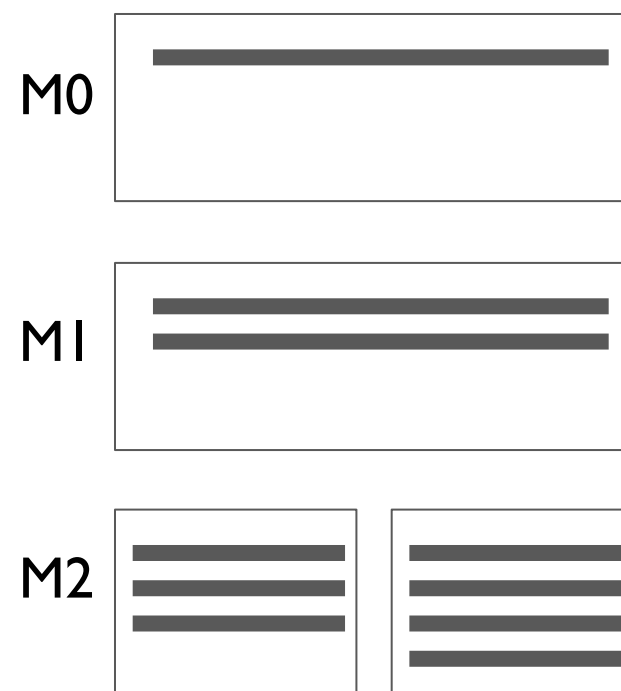
M2G1

# Experiment

- 3 x 3 x 2 between-subjects factorial design:



max-polyphony



gini-polyphony



- Soundscape examples:

M0G0

M0G1

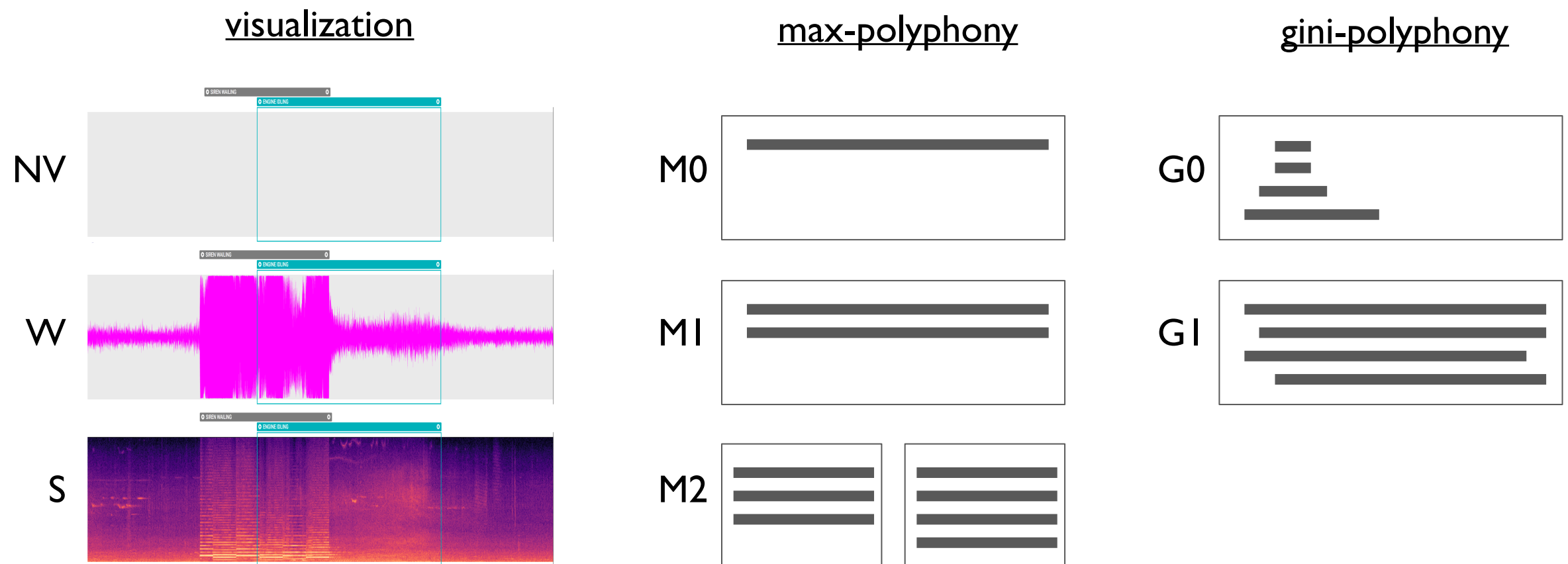
M2G0

M2G1



# Experiment

- 3 x 3 x 2 between-subjects factorial design:



- Soundscape examples:

M0G0

**M0G1**

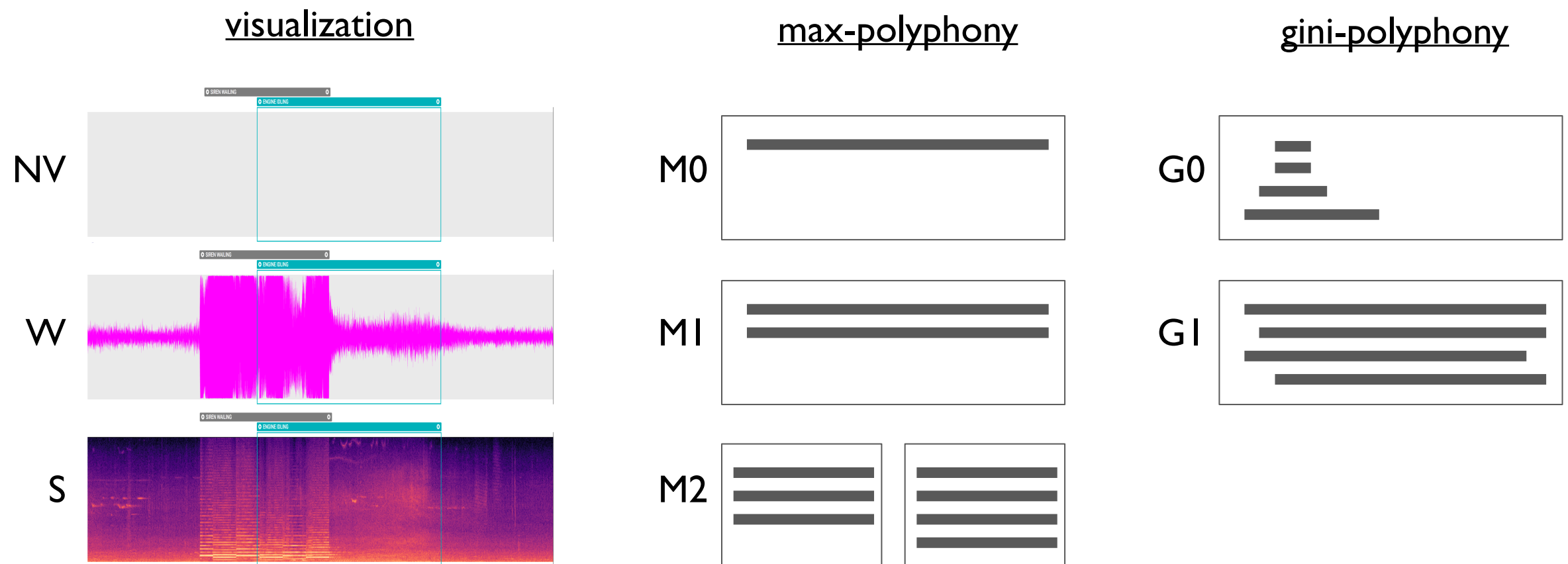
M2G0

M2G1



# Experiment

- 3 x 3 x 2 between-subjects factorial design:



- Soundscape examples:

M0G0

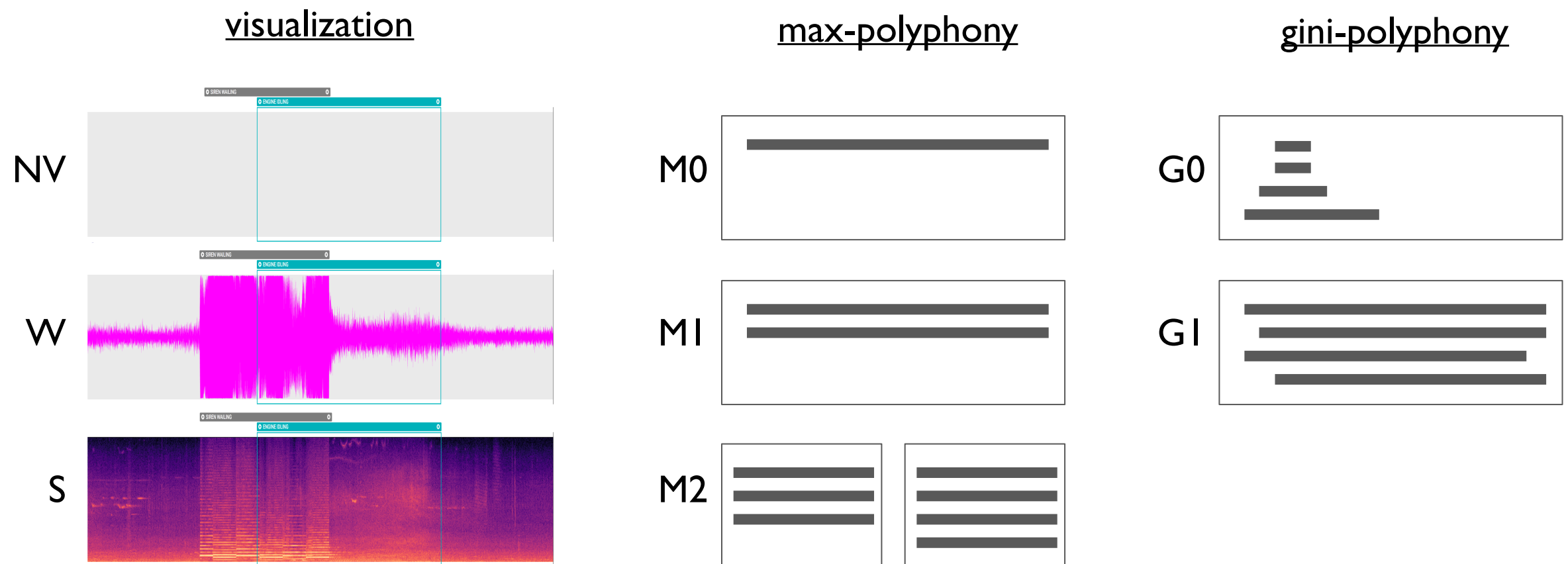
**M0G1**

M2G0

M2G1

# Experiment

- 3 x 3 x 2 between-subjects factorial design:



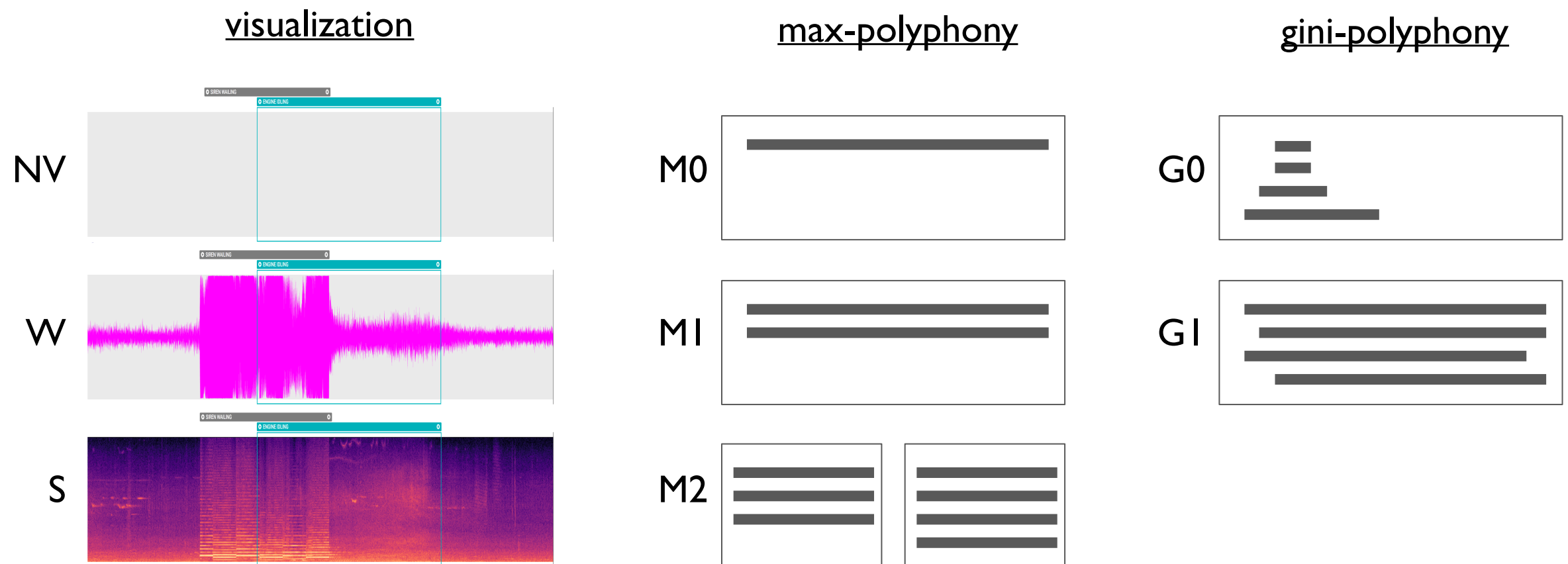
- Soundscape examples:  
M0G0                      M0G1

**M2G0**

M2G1

# Experiment

- 3 x 3 x 2 between-subjects factorial design:



- Soundscape examples:  
M0G0                      M0G1

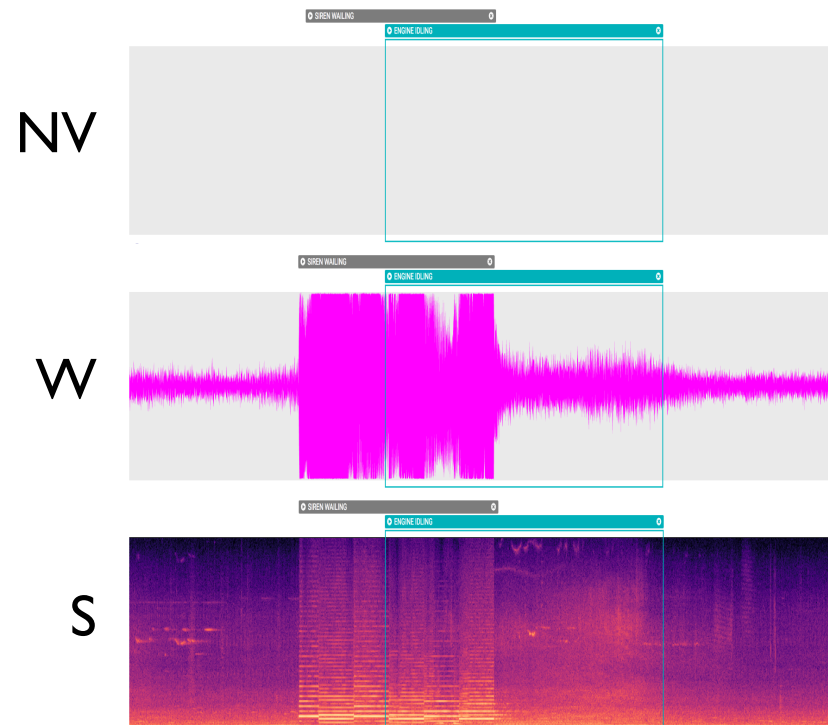
M2G0

M2G1

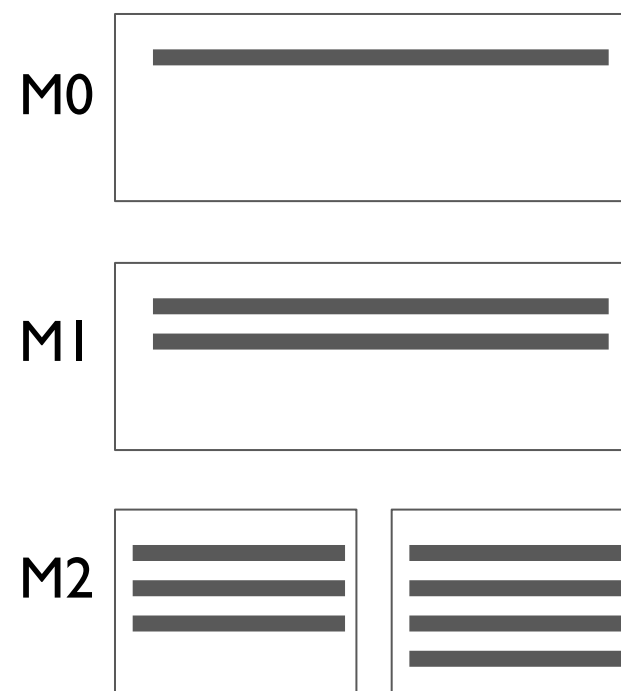
# Experiment

- 3 x 3 x 2 between-subjects factorial design:

## visualization



## max-polyphony



## gini-polyphony



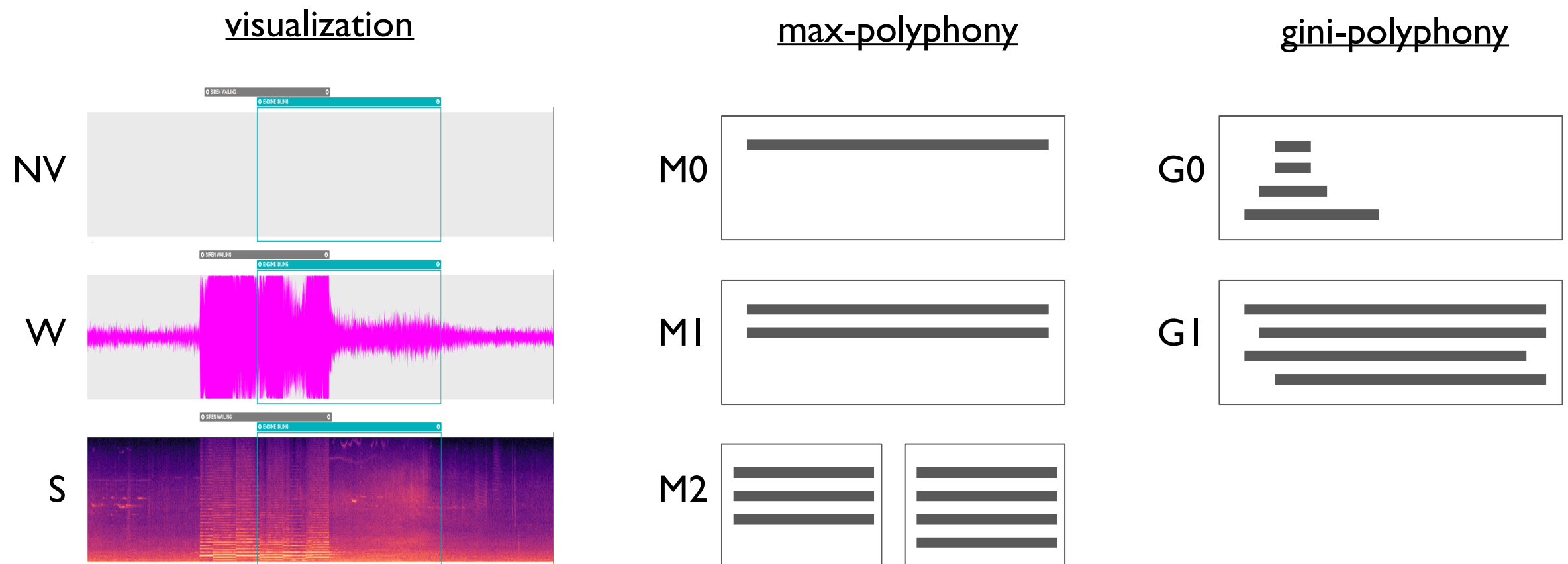
- Soundscape examples:  
M0G0                      M0G1

M2G0

**M2G1**

# Experiment

- 3 x 3 x 2 between-subjects factorial design:



- Soundscape examples:  
M0G0                      M0G1

M2G0

**M2G1**

# Experiment

---

- 10s synthesized urban soundscapes (i.e. audio stimuli)
- Classes: car horn honking, dog barking, engine idling, gun shooting, jack hammer drilling, music playing, people shouting, people talking, siren wailing
- 30 replications / 540 participants from Mechanical Turk
- 10 soundscapes per complexity condition (i.e. max- x gini-polyphony pair)
- Counterbalanced ordering of soundscapes

# Participant Tasks

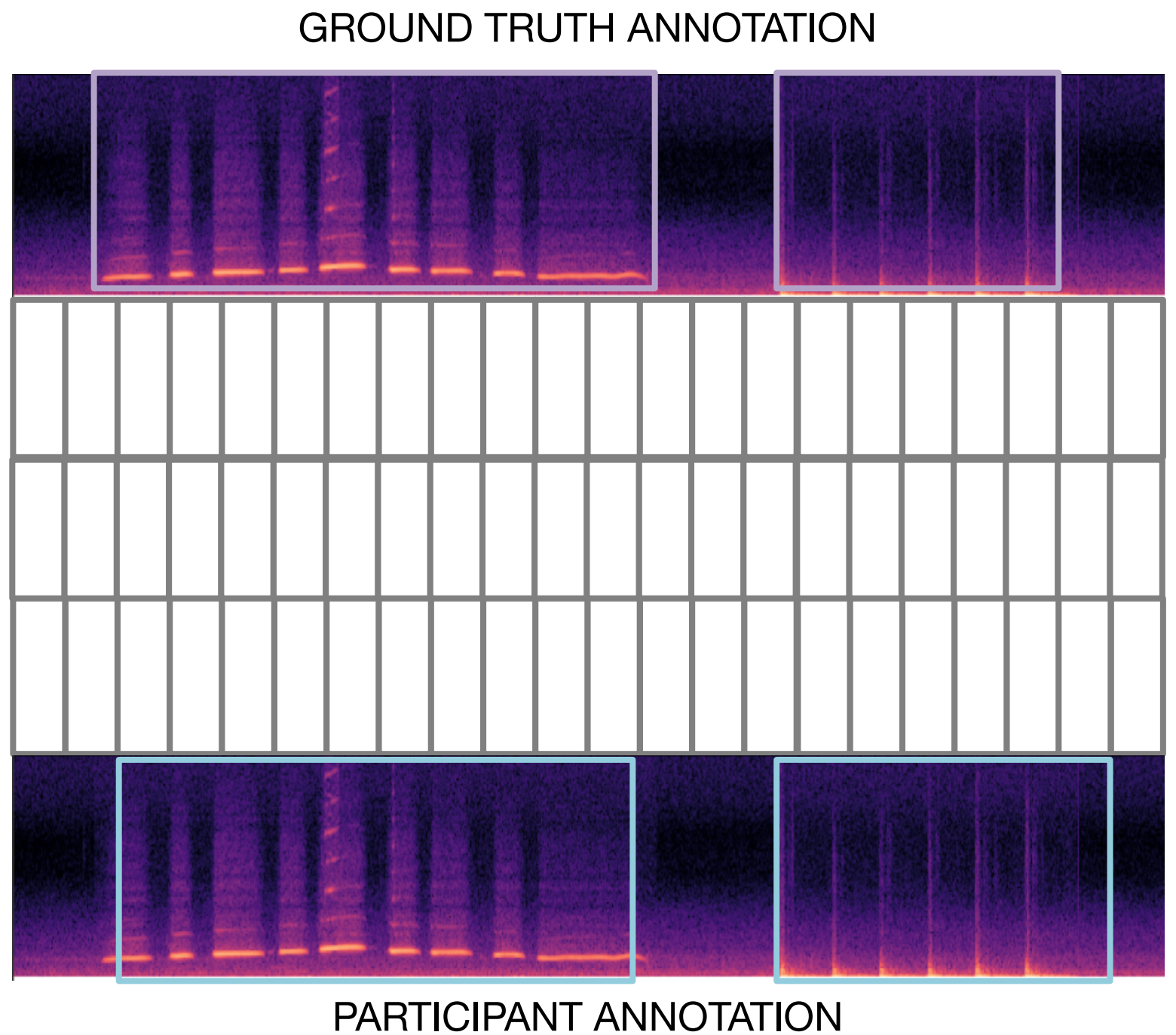
---

- Hearing screening
- Pre-task questionnaire
- Tutorial video
- Practice annotation task
- Series of 10 annotation tasks
- Post-task questionnaire

# Frame-based Evaluation

---

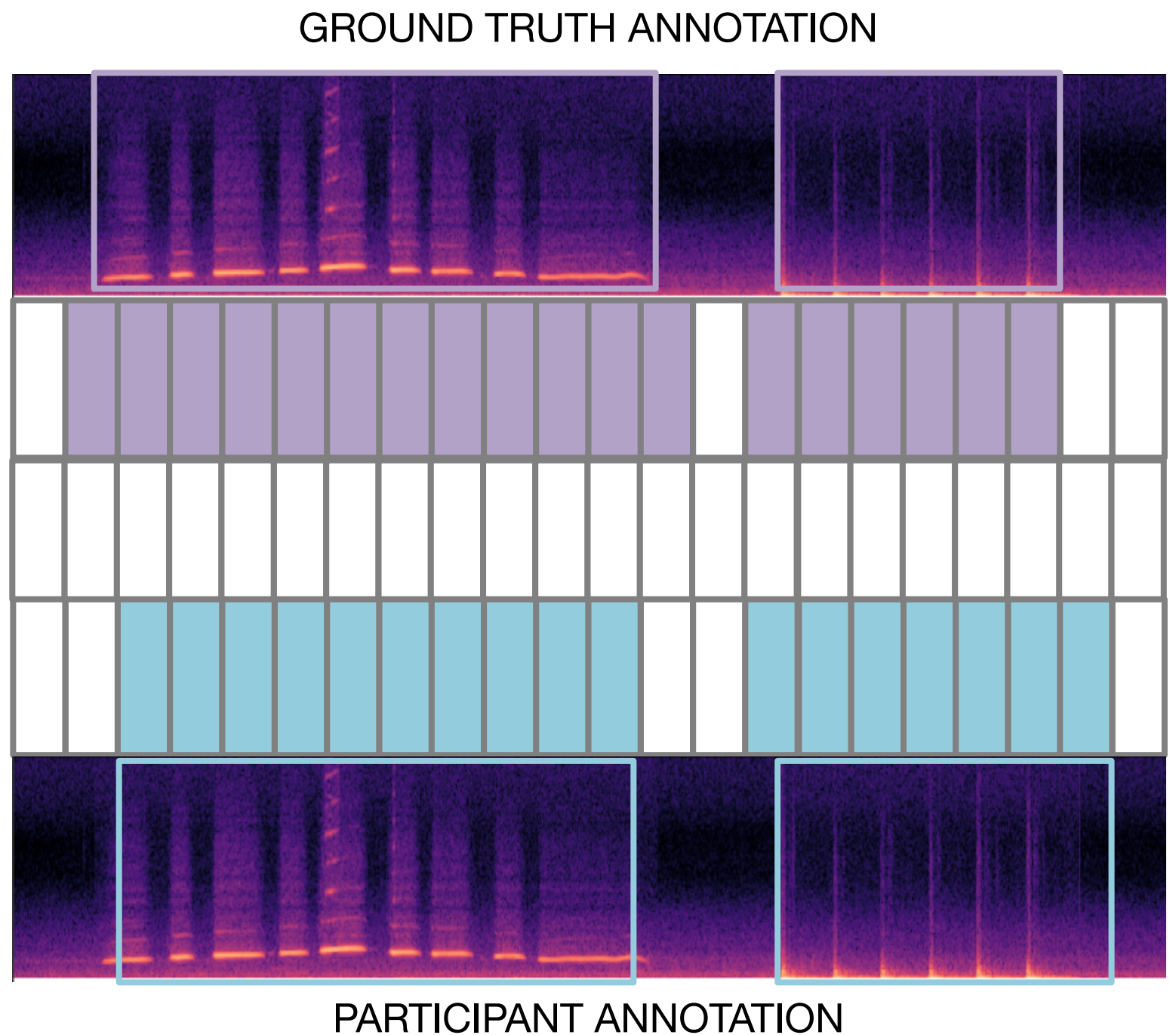
- Segment signal into 100ms frames.





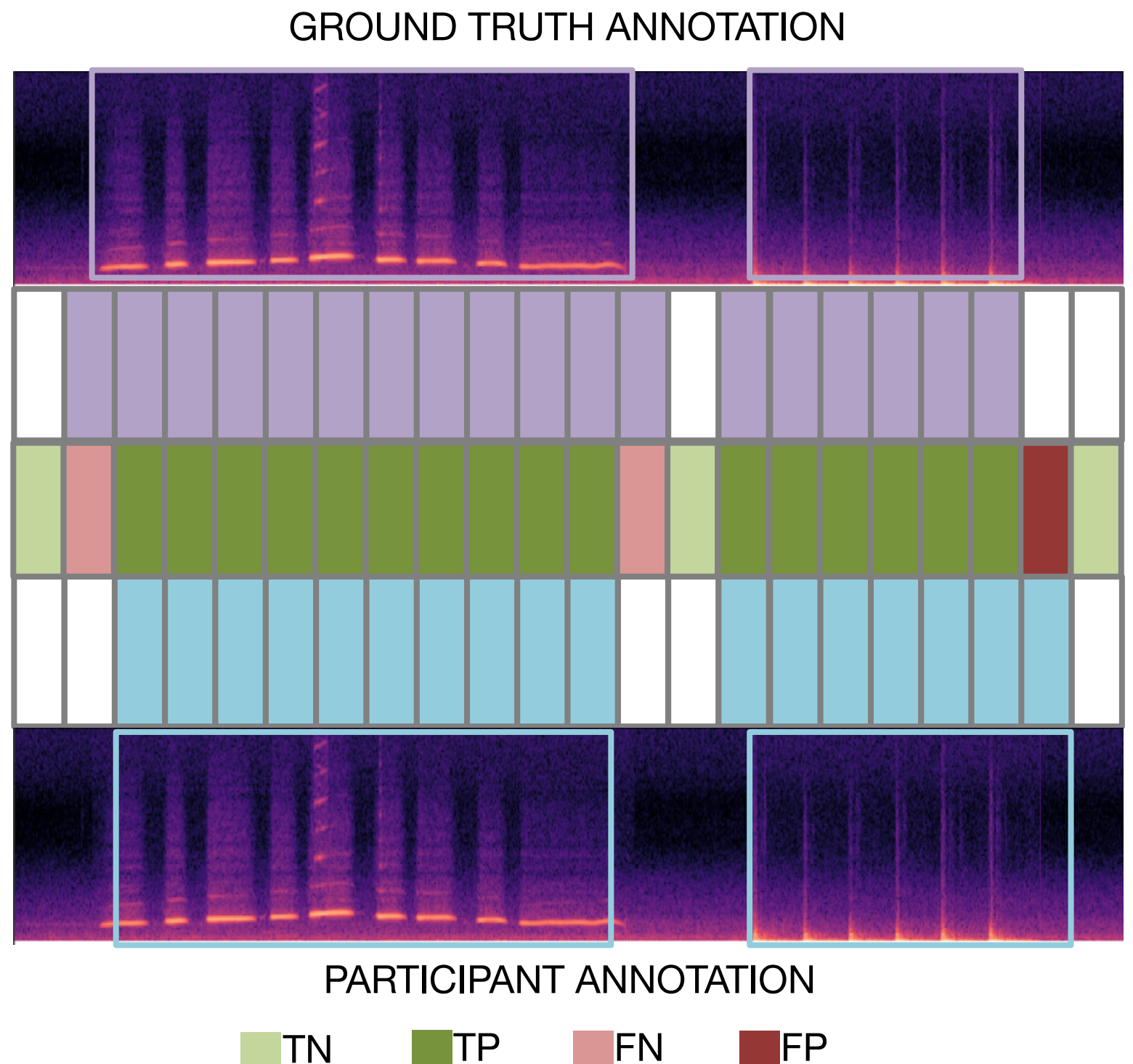
# Frame-based Evaluation

- Segment signal into 100ms frames.
- Round the annotations to the outer frame boundaries



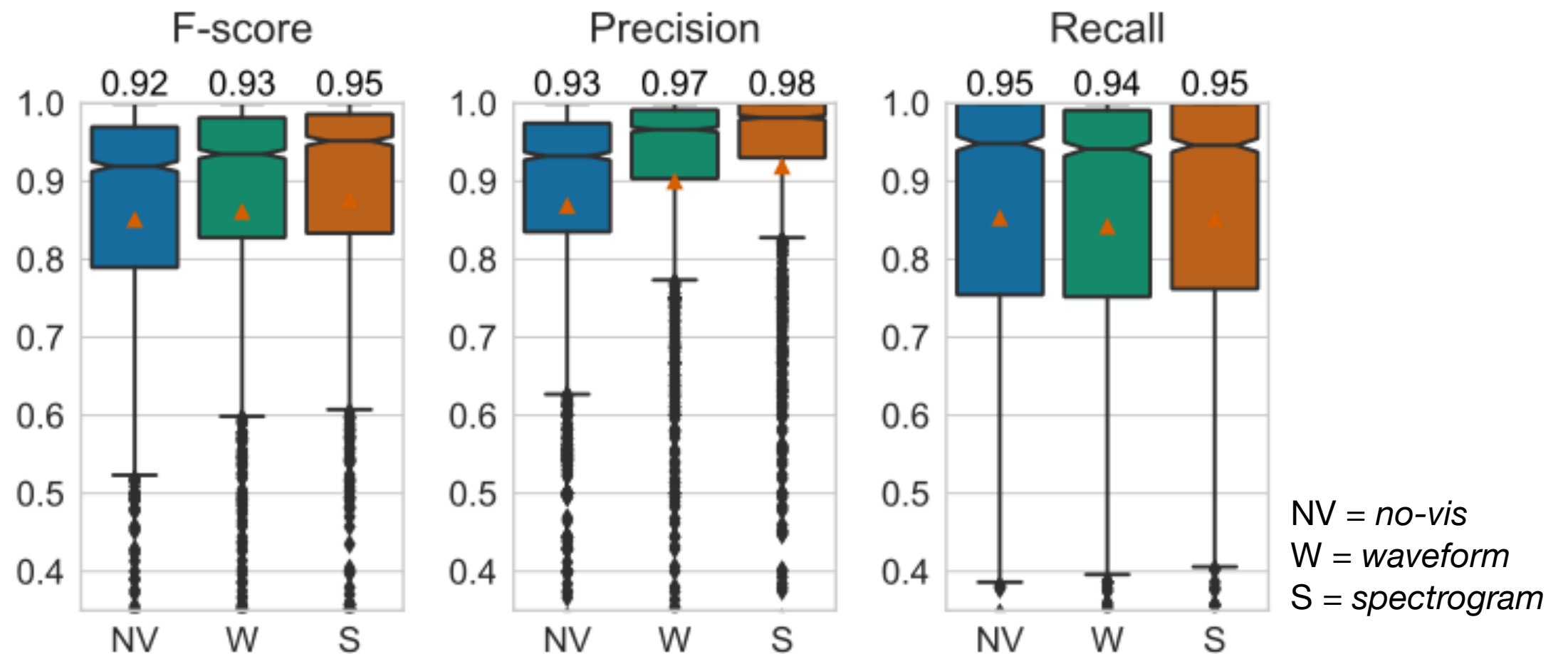
# Frame-based Evaluation

- Segment signal into 100ms frames.
- Round the annotations to the outer frame boundaries
- Count TP, FP, FN for each class and calculate precision, recall, F-score

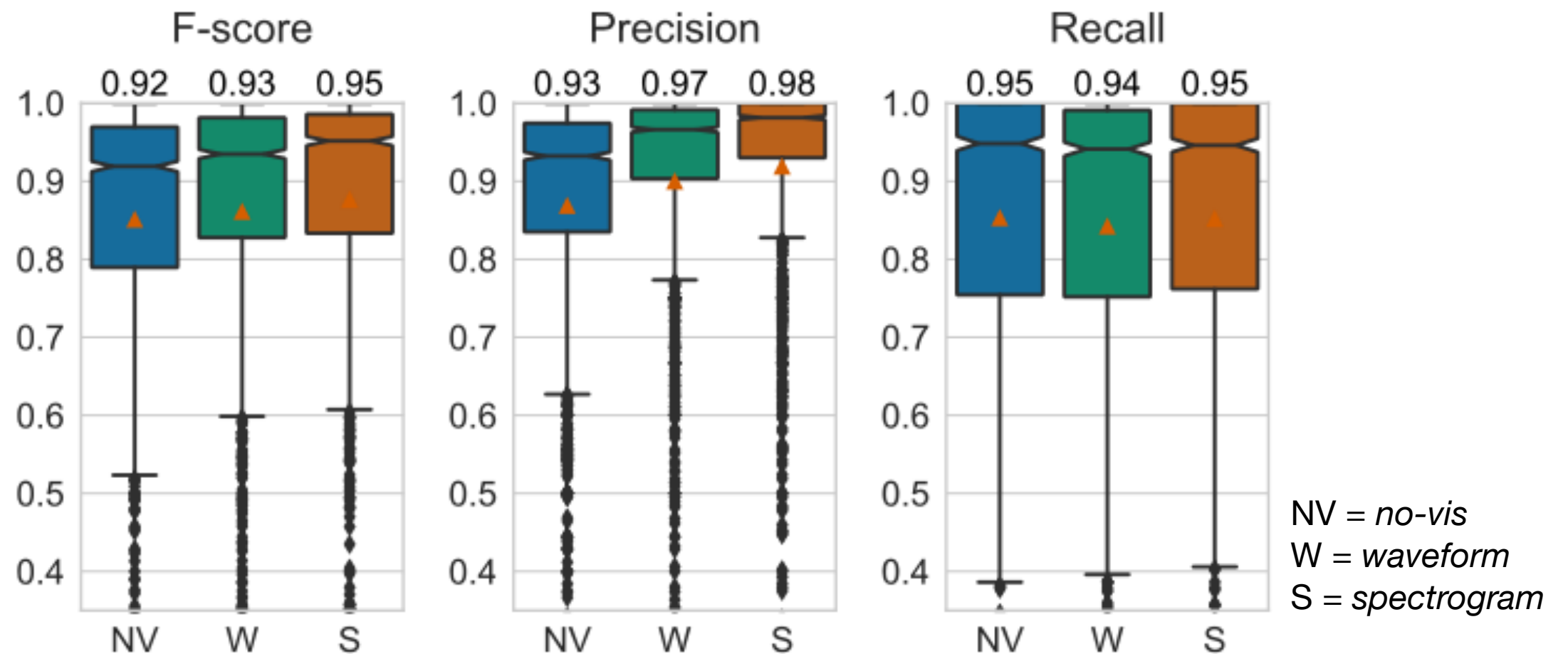


# Results

# Effect of Visualization on Quality of Annotations

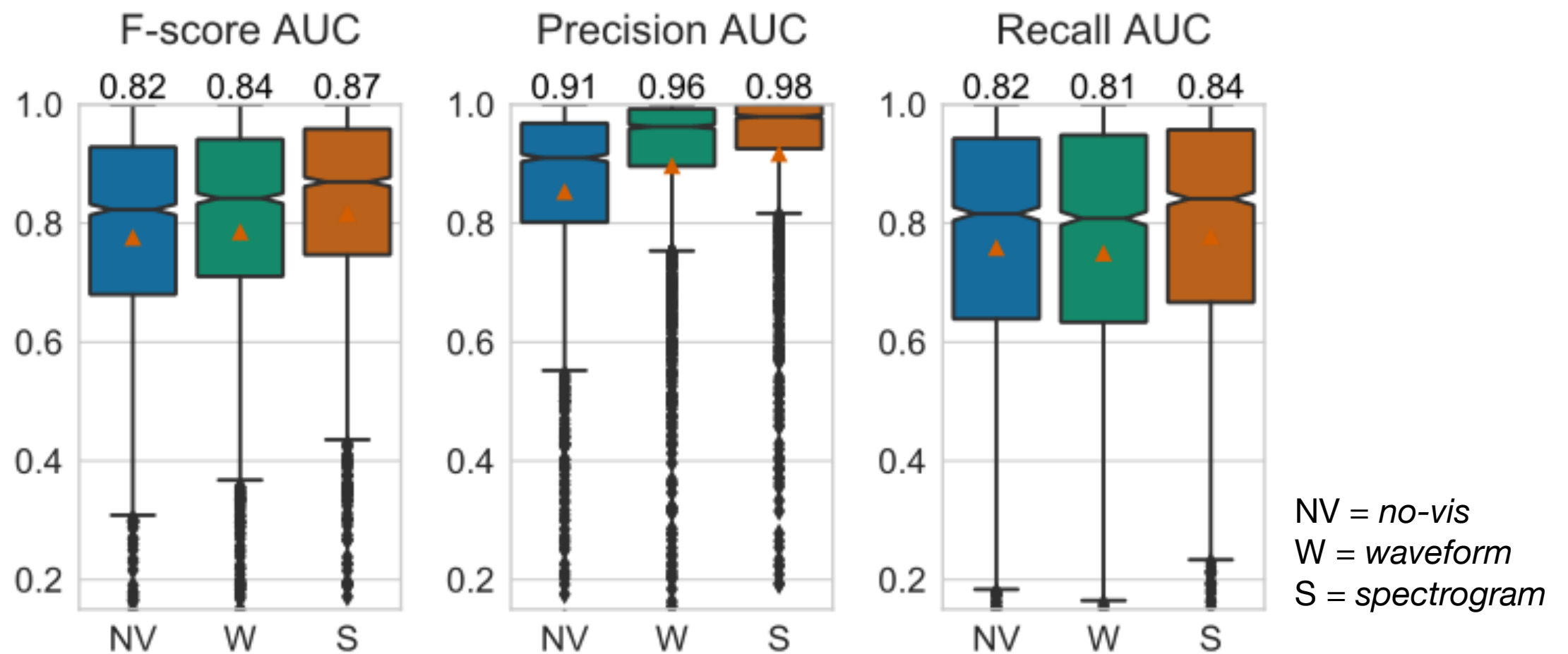


# Effect of Visualization on Quality of Annotations



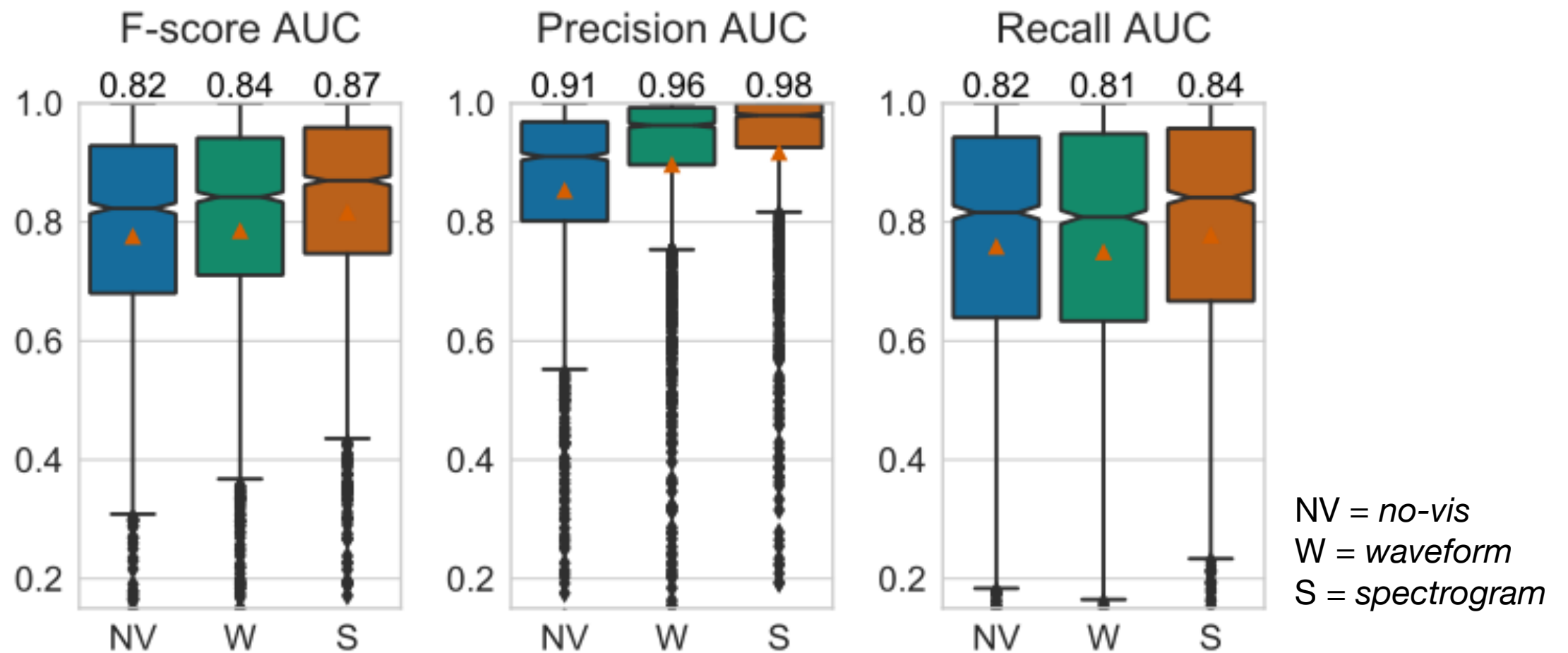
Spectrogram → higher-quality annotations

# Effect of Visualization on Quality and Speed of Annotations



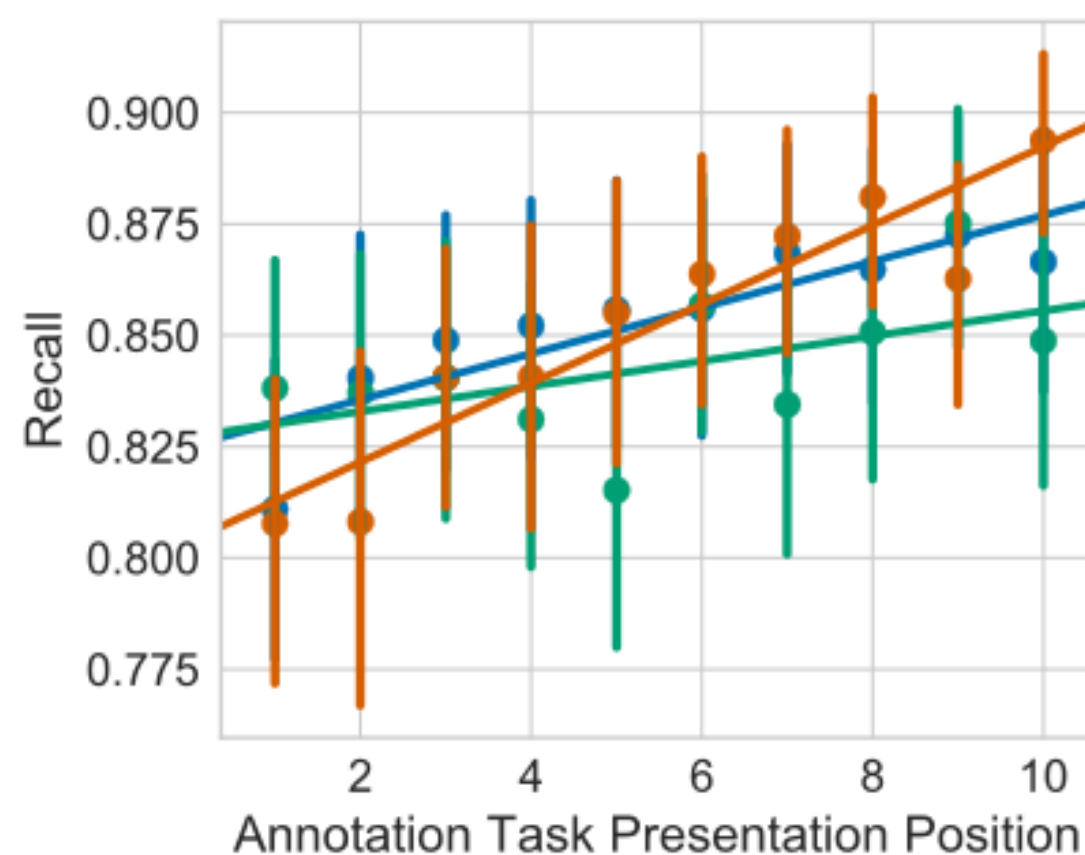
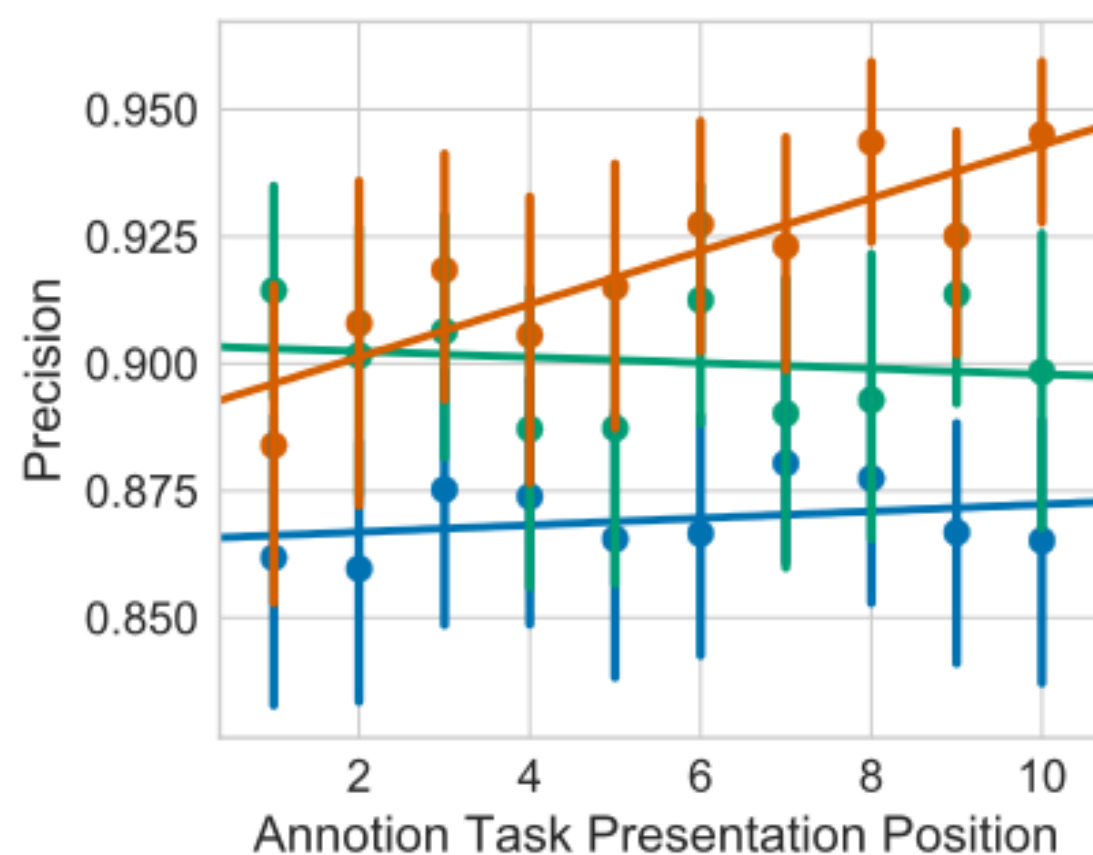


# Effect of Visualization on Quality and Speed of Annotations



Spectrogram → higher-quality and faster annotations

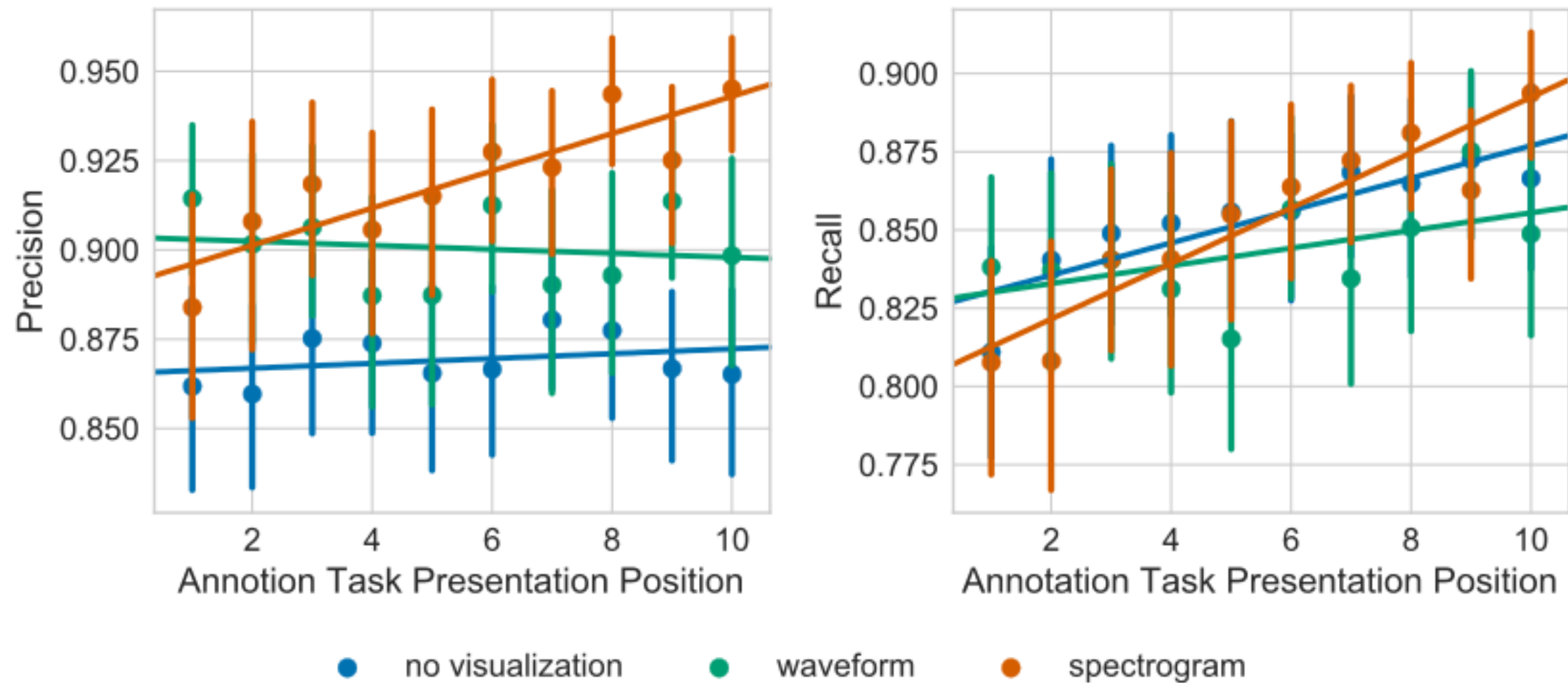
# Effect of Visualization on Task Learning



● no visualization    ● waveform    ● spectrogram

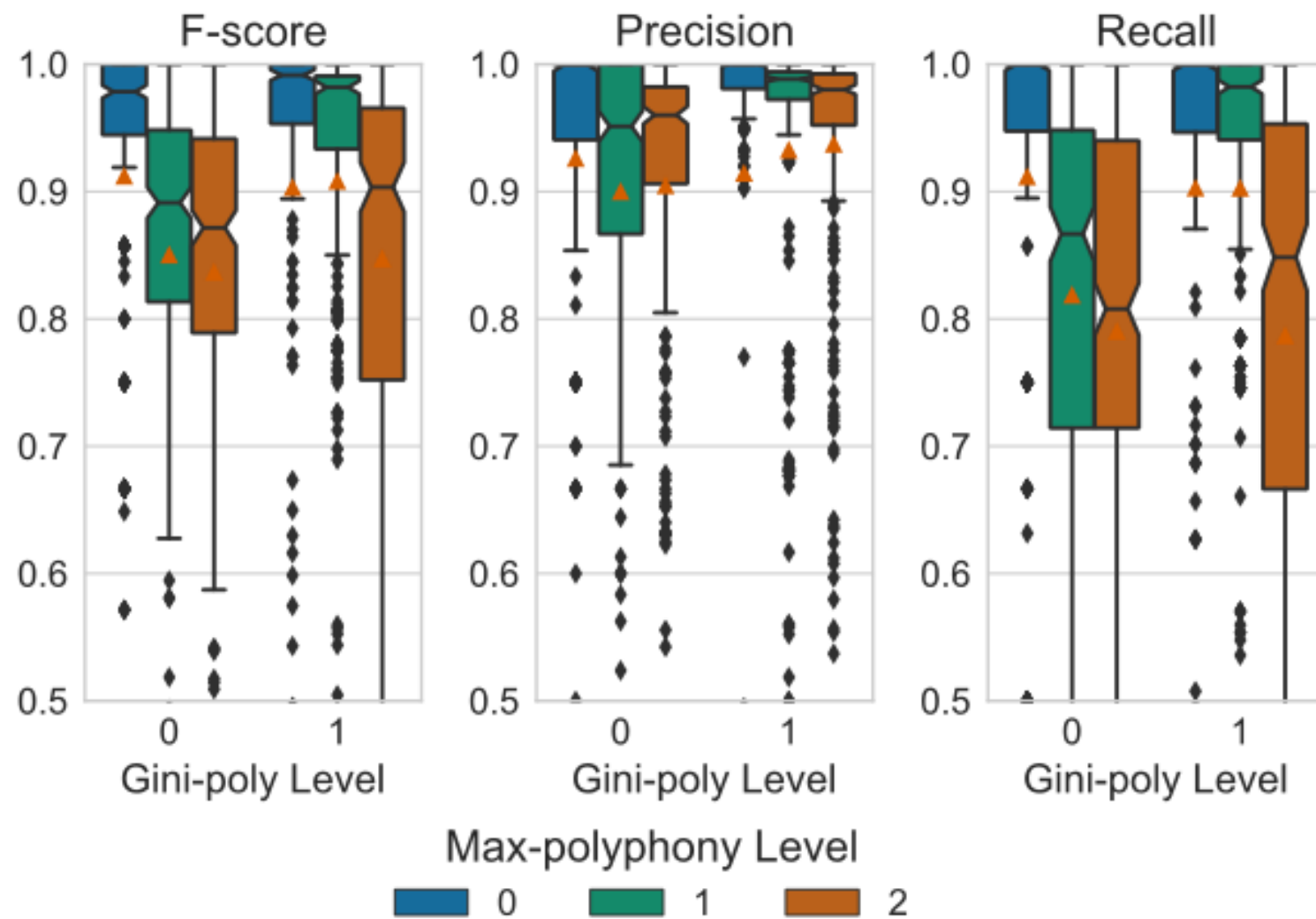


# Effect of Visualization on Task Learning

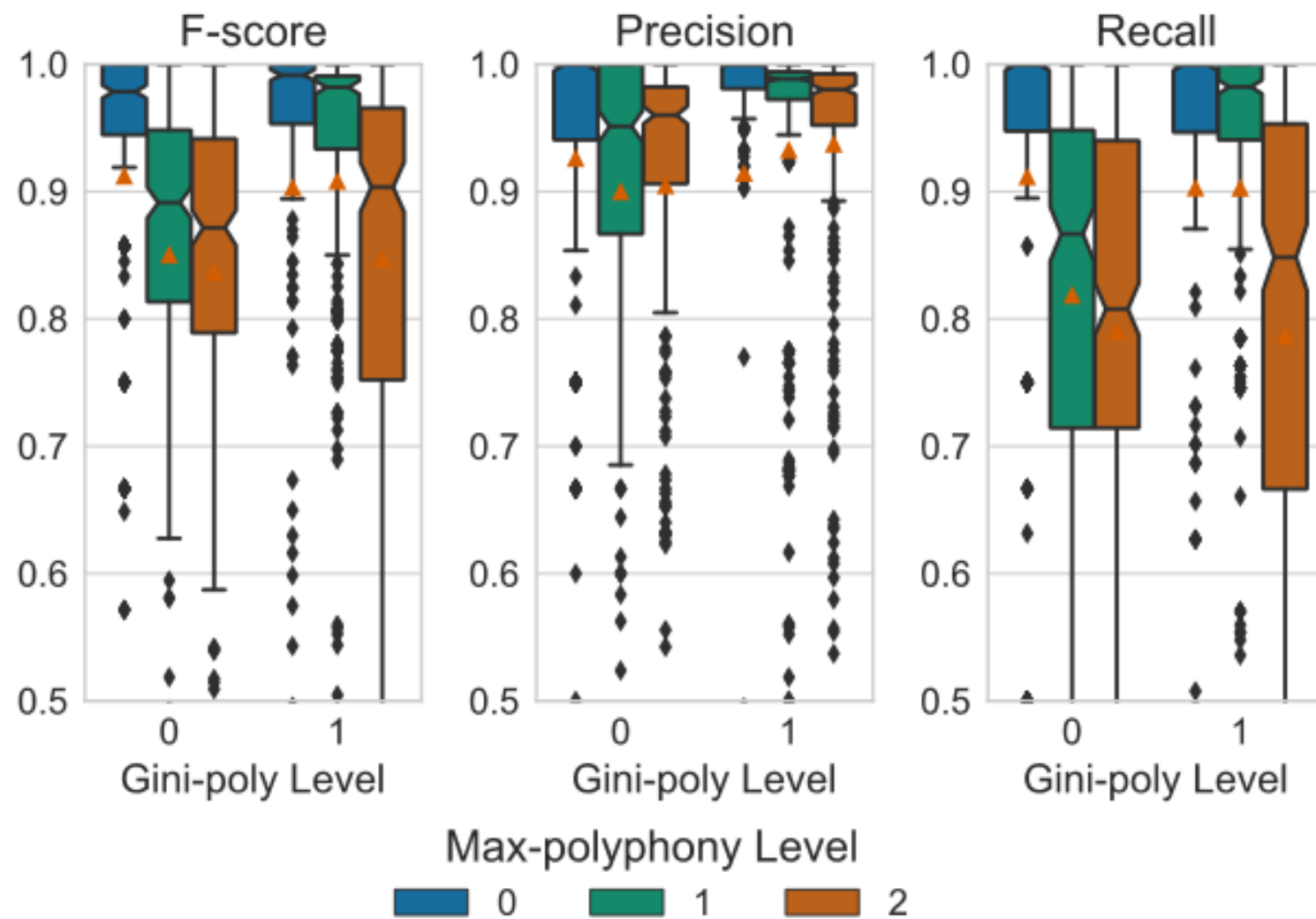


Expect even higher quality annotations after learning period

# Effect of Soundscape Complexity on Annotation Quality

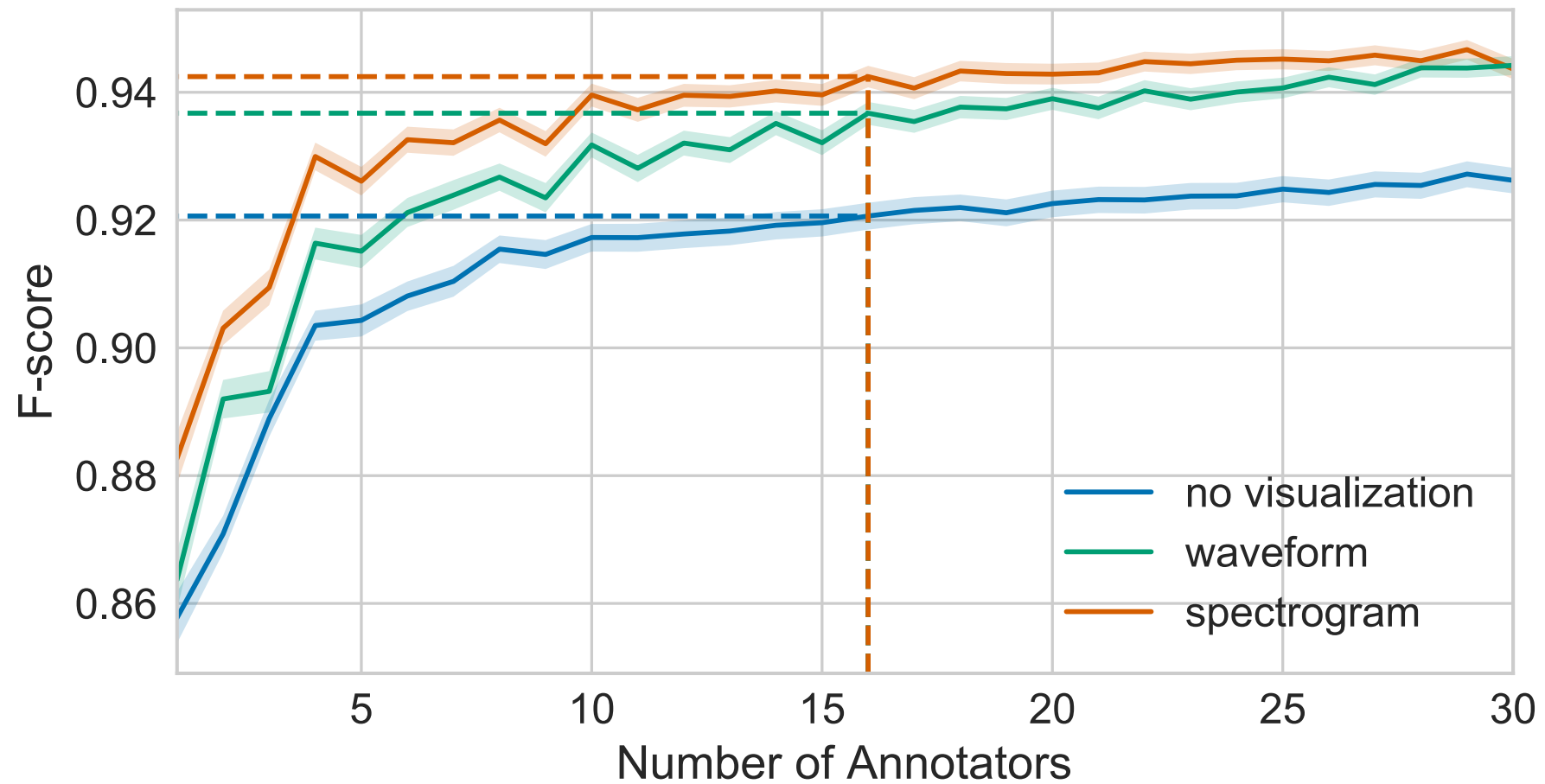


# Effect of Soundscape Complexity on Annotation Quality

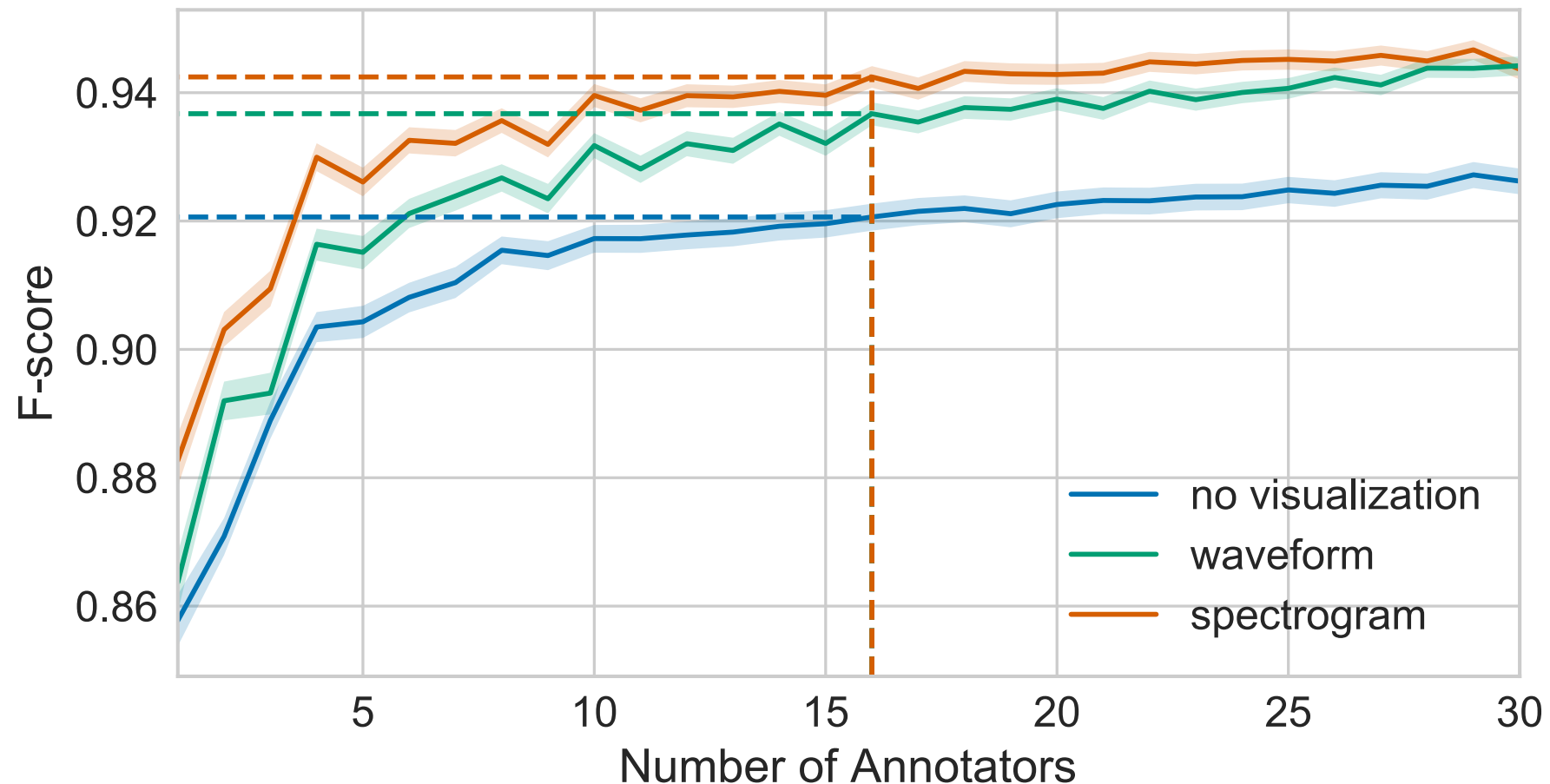


Complex soundscapes → expect precise but incomplete annotations

# Effect of Number of Annotators on Aggregate Annotation Quality



# Effect of Number of Annotators on Aggregate Annotation Quality



16 annotators captured 90% of gain in annotation quality, but 5 annotators is reasonable choice with respect to cost/quality trade-off

# Practical Implications

---

- Spectrogram → higher-quality and faster annotations
- Expect even higher quality annotations after learning period
- Complex soundscapes → expect precise but incomplete annotations
- 5 annotators is reasonable choice with respect to cost/quality trade-off

# Publications

---

- Cartwright, M., Seals, A., Salamon, J., Williams, A., Mikloska, S., MacConnell, D., Law, E., Bello, J.P., Nov, O. **Seeing Sound: Investigating the Effects of Visualizations and Complexity on Crowdsourced Audio Annotations.** In *Proceedings of the ACM on Human-Computer Interaction*, vol. 1(2), 2017.
- Cartwright, M., Salamon, J., Seals, A., Nov, O., Bello, J.P. **Investigating the Effect of Sound-Event Loudness on Crowdsourced Audio Annotations.** In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2018

Next Steps



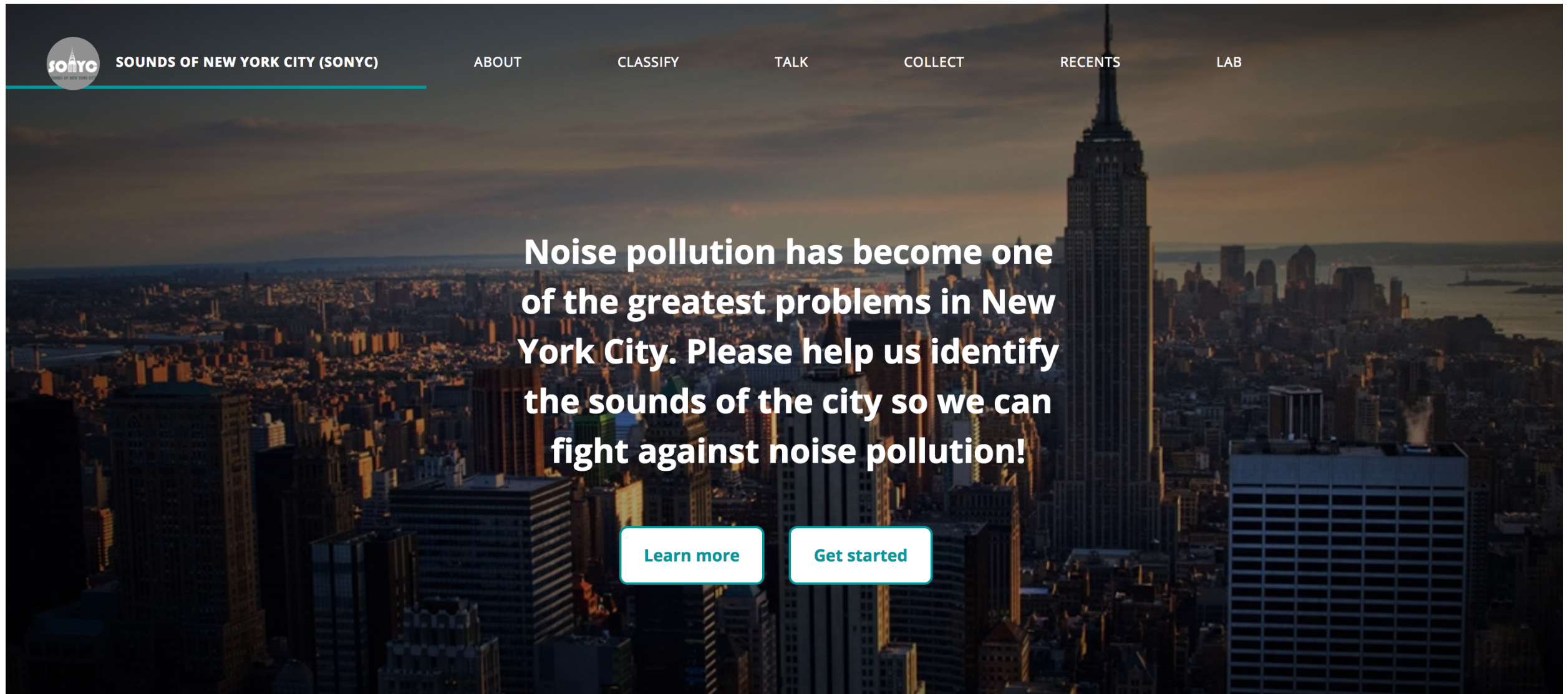
# Annotation Goals for Machine Listening

---

- Annotate SONYC sensor recordings on 22 urban sound classes
- Identify at least 1000 positive examples for each sound class
- “Positive example” -> “weak label” (binary) on 10s sensor recordings
- All recordings assessed for all sound classes

# Citizen Science Annotation Campaign

---



**SONYC** SOUNDS OF NEW YORK CITY (SONYC) ABOUT CLASSIFY TALK COLLECT RECENTS LAB

**Noise pollution has become one of the greatest problems in New York City. Please help us identify the sounds of the city so we can fight against noise pollution!**

[Learn more](#) [Get started](#)

# Deciding Which Recordings to Annotate

---

- Uniform sampling -> skew towards everyday street sound (would need much more annotation to reach targets)
- Bias sampling probabilities by similarity to known sound class examples from YouTube
- Compute similarity by distance in VGGish features (a deep audio embedding), extracted for 1/10th of data, ~5 million audio recording
- Release to Zooniverse in groups that are spaced in location and time to protect privacy
- Eventually move to an active sampling approach...

# Urban Sound Classes

---

## ▶ **Engines:**

- ▶ Small-sounding engine
- ▶ Medium-sounding engine
- ▶ Large-sounding engine
- ▶ Other/unknown engine

## ▶ **Powered sawing tools:**

- ▶ Chain saw
- ▶ Small->medium rotating saw
- ▶ Large rotating saw
- ▶ Other/unknown saw

## ▶ **Music:**

- ▶ Stationary music
- ▶ Mobile music
- ▶ Ice-cream truck
- ▶ Other/unknown music

## ▶ **Impact sounds:**

- ▶ Rock drill
- ▶ Jackhammer
- ▶ Hoe ram
- ▶ Pile driver
- ▶ Other/unknown impact sound

## ▶ **Alert signals:**

- ▶ Car horn
- ▶ Car alarm
- ▶ Siren
- ▶ Reverse beeper
- ▶ Other/unknown alert signal

## ▶ **Human and animal vocalization sounds:**

- ▶ Person or small group talking
- ▶ Person shouting
- ▶ Crowd
- ▶ Amplified speech
- ▶ Dog barking/whining
- ▶ Other/unknown human or animal vocalization sound

## ▶ **Other/unknown sound:**

- ▶ Other/unknown sound
- ▶ Other/unknown construction sound

# Experiment on Weak Multi-Label Audio Annotations

---

How do **label set specificity** and **task assignment strategy** affect

- annotation quality?
- annotation speed/throughput?
- annotator satisfaction?
- annotator retention?

Label set specificity:

- General (1-pass annotation)
- 2-stage
- Specific (multi-pass annotation)

Task assignment strategy:

- Random
- Annotator-choice

# Varying Label Set Specificity

---

## General (Single-pass) Annotation Task

- ▶ Small-sounding engine
- ▶ Medium-sounding engine
- ▶ Large-sounding engine
- ▶ Other/unknown engine
- ▶ Chain saw
- ▶ Small->medium rotating saw
- ▶ Large rotating saw
- ▶ Other/unknown saw
- ▶ Stationary music
- ▶ Mobile music
- ▶ Ice-cream truck
- ▶ Other/unknown music
- ▶ Rock drill
- ▶ Jackhammer
- ▶ Hoe ram
- ▶ Pile driver
- ▶ Other/unknown impact sound
- ▶ Car horn
- ▶ Car alarm
- ▶ Siren
- ▶ Reverse beeper
- ▶ Other/unknown alert signal
- ▶ Person or small group talking
- ▶ Person shouting
- ▶ Crowd
- ▶ Amplified speech
- ▶ Dog barking/whining
- ▶ Other/unknown human or animal vocalization sound
- ▶ Other/unknown sound
- ▶ Other/unknown construction sound

# Varying Label Set Specificity

---

## General (Single-pass) Annotation Task

- ▶ Small-sounding engine
- ▶ Medium-sounding engine
- ▶ Large-sounding engine
- ▶ Other/unknown engine
- ▶ Chain saw
- ▶ Small->medium rotating saw
- ▶ Large rotating saw
- ▶ Other/unknown saw
- ▶ Stationary music
- ▶ Mobile music
- ▶ Ice-cream truck
- ▶ Other/unknown music
- ▶ Rock drill
- ▶ Jackhammer
- ▶ Hoe ram
- ▶ Pile driver
- ▶ Other/unknown impact sound
- ▶ Car horn
- ▶ Car alarm
- ▶ Siren
- ▶ Reverse beeper
- ▶ Other/unknown alert signal
- ▶ Person or small group talking
- ▶ Person shouting
- ▶ Crowd
- ▶ Amplified speech
- ▶ Dog barking/whining
- ▶ Other/unknown human or animal vocalization sound
- ▶ Other/unknown sound
- ▶ Other/unknown construction sound

# Varying Label Set Specificity

---

## General (Single-pass) Annotation Task

- ▶ Small-sounding engine
- ▶ Medium-sounding engine
- ▶ Large-sounding engine
- ▶ Other/unknown engine
- ▶ Chain saw
- ▶ Small-medium rotating saw
- ▶ Large rotating saw
- ▶ Other/unknown saw
- ▶ Stationary music
- ▶ Mobile music
- ▶ Ice-cream truck
- ▶ Other/unknown music
- ▶ Rock drill
- ▶ Jackhammer
- ▶ Hoe ram
- ▶ Pile driver
- ▶ Other/unknown impact sound
- ▶ Car horn
- ▶ Car alarm
- ▶ Siren
- ▶ Reverse beeper
- ▶ Other/unknown alert signal
- ▶ Person or small group talking
- ▶ Person shouting
- ▶ Crowd
- ▶ Amplified speech
- ▶ Dog barking/whining
- ▶ Other/unknown human or animal vocalization sound
- ▶ Other/unknown sound
- ▶ Other/unknown construction sound



# Varying Label Set Specificity

---

## General (Single-pass) Annotation Task

- ▶ Small-sounding engine
- ▶ Medium-sounding engine
- ▶ Large-sounding engine
- ▶ Other/unknown engine
- ▶ Chain saw
- ▶ Small- & medium rotating saw
- ▶ Large rotating saw
- ▶ Other/unknown saw
- ▶ Stationary music
- ▶ Mobile music
- ▶ Ice-cream truck
- ▶ Other/unknown music
- ▶ Rock drill
- ▶ Jackhammer
- ▶ Pile ram
- ▶ Pile driver
- ▶ Other/unknown impact sound
- ▶ Car horn
- ▶ Car alarm
- ▶ Siren
- ▶ Reverse beeper
- ▶ Other/unknown alert signal
- ▶ Person or small group talking
- ▶ Person shouting
- ▶ Crowd
- ▶ Amplified speech
- ▶ Dog barking/whining
- ▶ Other/unknown human or animal vocalization sound
- ▶ Other/unknown sound
- ▶ Other/unknown construction sound

# Varying Label Set Specificity

---

## General (Single-pass) Annotation Task

- ▶ Small-sounding engine
- ▶ Medium-sounding engine
- ▶ Large-sounding engine
- ▶ Other/unknown engine
- ▶ Chain saw
- ▶ Small-medium rotating saw
- ▶ Large rotating saw
- ▶ Other/unknown saw
- ▶ Stationary music
- ▶ Mobile music
- ▶ Ice-cream truck
- ▶ Other/unknown music
- ▶ Rock drill
- ▶ Jackhammer
- ▶ Hoop ram
- ▶ Pile driver
- ▶ Other/unknown impact sound
- ▶ Car horn
- ▶ Car alarm
- ▶ Siren
- ▶ Reverse beeper
- ▶ Other/unknown alert signal
- ▶ Person or small group talking
- ▶ Person shouting
- ▶ Crowd
- ▶ Amplified speech
- ▶ Dog barking/whining
- ▶ Other/unknown human or animal vocalization sound
- ▶ Other/unknown sound
- ▶ Other/unknown construction sound

## Hybrid (2-stage) Annotation Task

### Stage 1

- ▶ Engines
- ▶ Powered sawing tools
- ▶ Music
- ▶ Impact sounds
- ▶ Alert signals
- ▶ Human and animal vocalization sounds
- ▶ Other/unknown sound
- ▶ Other/unknown construction sound

# Varying Label Set Specificity

---

## General (Single-pass) Annotation Task

- ▶ Small-sounding engine
- ▶ Medium-sounding engine
- ▶ Large-sounding engine
- ▶ Other/unknown engine
- ▶ Chain saw
- ▶ Small- & medium rotating saw
- ▶ Large rotating saw
- ▶ Other/unknown saw
- ▶ Stationary music
- ▶ Mobile music
- ▶ Ice-cream truck
- ▶ Other/unknown music
- ▶ Rock drill
- ▶ Jackhammer
- ▶ Hoop ram
- ▶ Pile driver
- ▶ Other/unknown impact sound
- ▶ Car horn
- ▶ Car alarm
- ▶ Siren
- ▶ Reverse beeper
- ▶ Other/unknown alert signal
- ▶ Person or small group talking
- ▶ Person shouting
- ▶ Crowd
- ▶ Amplified speech
- ▶ Dog barking/whining
- ▶ Other/unknown human or animal vocalization sound
- ▶ Other/unknown sound
- ▶ Other/unknown construction sound

## Hybrid (2-stage) Annotation Task

### Stage 1

- ▶ Engines
- ▶ Powered sawing tools
- ▶ Music
- ▶ Impact sounds
- ▶ Alert signals
- ▶ Human and animal vocalization sounds
- ▶ Other/unknown sound
- ▶ Other/unknown construction sound

# Varying Label Set Specificity

---

## General (Single-pass) Annotation Task

- ▶ Small-sounding engine
- ▶ Medium-sounding engine
- ▶ Large-sounding engine
- ▶ Other/unknown engine
- ▶ Chain saw
- ▶ Small- & medium rotating saw
- ▶ Large rotating saw
- ▶ Other/unknown saw
- ▶ Stationary music
- ▶ Mobile music
- ▶ Ice-cream truck
- ▶ Other/unknown music
- ▶ Rock drill
- ▶ Jackhammer
- ▶ Hoop ram
- ▶ Pile driver
- ▶ Other/unknown impact sound
- ▶ Car horn
- ▶ Car alarm
- ▶ Siren
- ▶ Reverse beeper
- ▶ Other/unknown alert signal
- ▶ Person or small group talking
- ▶ Person shouting
- ▶ Crowd
- ▶ Amplified speech
- ▶ Dog barking/whining
- ▶ Other/unknown human or animal vocalization sound
- ▶ Other/unknown sound
- ▶ Other/unknown construction sound

## Hybrid (2-stage) Annotation Task

### Stage 1

- ▶ Engines
- ▶ Powered sawing tools
- ▶ Music
- ▶ Impact sounds
- ▶ Alert signals
- ▶ Human and animal vocalization sounds
- ▶ Other/unknown sound
- ▶ Other/unknown construction sound

# Varying Label Set Specificity

---

## General (Single-pass) Annotation Task

- ▶ Small-sounding engine
- ▶ Medium-sounding engine
- ▶ Large-sounding engine
- ▶ Other/unknown engine
- ▶ Chain saw
- ▶ Small->medium rotating saw
- ▶ Large rotating saw
- ▶ Other/unknown saw
- ▶ Stationary music
- ▶ Mobile music
- ▶ Ice-cream truck
- ▶ Other/unknown music
- ▶ Rock drill
- ▶ Jackhammer
- ▶ Hoop ram
- ▶ Pile driver
- ▶ Other/unknown impact sound
- ▶ Car horn
- ▶ Car alarm
- ▶ Siren
- ▶ Reverse beeper
- ▶ Other/unknown alert signal
- ▶ Person or small group talking
- ▶ Person shouting
- ▶ Crowd
- ▶ Amplified speech
- ▶ Dog barking/whining
- ▶ Other/unknown human or animal vocalization sound
- ▶ Other/unknown sound
- ▶ Other/unknown construction sound

## Hybrid (2-stage) Annotation Task

### Stage 1

- ▶ Engines
- ▶ Powered sawing tools
- ▶ Music
- ▶ Impact sounds
- ▶ Alert signals
- ▶ Human and animal vocalization sounds
- ▶ Other/unknown sound
- ▶ Other/unknown construction sound

### Stage 2

- ▶ Powered sawing tools
  - ▶ Chain saw
  - ▶ Small->medium rotating saw
  - ▶ Large rotating saw
  - ▶ Other/unknown saw

# Varying Label Set Specificity

---

## General (Single-pass) Annotation Task

- ▶ Small-sounding engine
- ▶ Medium-sounding engine
- ▶ Large-sounding engine
- ▶ Other/unknown engine
- ▶ Chain saw
- ▶ Small->medium rotating saw
- ▶ Large rotating saw
- ▶ Other/unknown saw
- ▶ Stationary music
- ▶ Mobile music
- ▶ Ice-cream truck
- ▶ Other/unknown music
- ▶ Rock drill
- ▶ Jackhammer
- ▶ Hoop ram
- ▶ Pile driver
- ▶ Other/unknown impact sound
- ▶ Car horn
- ▶ Car alarm
- ▶ Siren
- ▶ Reverse beeper
- ▶ Other/unknown alert signal
- ▶ Person or small group talking
- ▶ Person shouting
- ▶ Crowd
- ▶ Amplified speech
- ▶ Dog barking/whining
- ▶ Other/unknown human or animal vocalization sound
- ▶ Other/unknown sound
- ▶ Other/unknown construction sound

## Hybrid (2-stage) Annotation Task

### Stage 1

- ▶ Engines
- ▶ Powered sawing tools
- ▶ Music
- ▶ Impact sounds
- ▶ Alert signals
- ▶ Human and animal vocalization sounds
- ▶ Other/unknown sound
- ▶ Other/unknown construction sound

### Stage 2

- ▶ Powered sawing tools
  - ▶ Chain saw
  - ▶ Small->medium rotating saw
  - ▶ Large rotating saw
  - ▶ Other/unknown saw

# Varying Label Set Specificity

---

## General (Single-pass) Annotation Task

- ▶ Small-sounding engine
- ▶ Medium-sounding engine
- ▶ Large-sounding engine
- ▶ Other/unknown engine
- ▶ Chain saw
- ▶ Small- & medium rotating saw
- ▶ Large rotating saw
- ▶ Other/unknown saw
- ▶ Stationary music
- ▶ Mobile music
- ▶ Ice-cream truck
- ▶ Other/unknown music
- ▶ Rock drill
- ▶ Jackhammer
- ▶ Hoop ram
- ▶ Pile driver
- ▶ Other/unknown impact sound
- ▶ Car horn
- ▶ Car alarm
- ▶ Siren
- ▶ Reverse beeper
- ▶ Other/unknown alert signal
- ▶ Person or small group talking
- ▶ Person shouting
- ▶ Crowd
- ▶ Amplified speech
- ▶ Dog barking/whining
- ▶ Other/unknown human or animal vocalization sound
- ▶ Other/unknown sound
- ▶ Other/unknown construction sound

## Hybrid (2-stage) Annotation Task

### Stage 1

- ▶ Engines
- ▶ Powered sawing tools
- ▶ Music
- ▶ Impact sounds
- ▶ Alert signals
- ▶ Human and animal vocalization sounds
- ▶ Other/unknown sound
- ▶ Other/unknown construction sound

# Varying Label Set Specificity

---

## General (Single-pass) Annotation Task

- ▶ Small-sounding engine
- ▶ Medium-sounding engine
- ▶ Large-sounding engine
- ▶ Other/unknown engine
- ▶ Chain saw
- ▶ Small- & medium rotating saw
- ▶ Large rotating saw
- ▶ Other/unknown saw
- ▶ Stationary music
- ▶ Mobile music
- ▶ Ice-cream truck
- ▶ Other/unknown music
- ▶ Rock drill
- ▶ Jackhammer
- ▶ Hoe ram
- ▶ Pile driver
- ▶ Other/unknown impact sound
- ▶ Car horn
- ▶ Car alarm
- ▶ Siren
- ▶ Reverse beeper
- ▶ Other/unknown alert signal
- ▶ Person or small group talking
- ▶ Person shouting
- ▶ Crowd
- ▶ Amplified speech
- ▶ Dog barking/whining
- ▶ Other/unknown human or animal vocalization sound
- ▶ Other/unknown sound
- ▶ Other/unknown construction sound

## Hybrid (2-stage) Annotation Task

### Stage 1

- ▶ Engines
- ▶ Powered sawing tools
- ▶ Music
- ▶ Impact sounds
- ▶ Alert signals
- ▶ Human and animal vocalization sounds
- ▶ Other/unknown sound
- ▶ Other/unknown construction sound

### Stage 2

- ▶ **Impact sounds:**
  - ▶ Rock drill
  - ▶ Jackhammer
  - ▶ Hoe ram
  - ▶ Pile driver
  - ▶ Other/unknown impact sound



# Varying Label Set Specificity

---

## General (Single-pass) Annotation Task

- ▶ Small-sounding engine
- ▶ Medium-sounding engine
- ▶ Large-sounding engine
- ▶ Other/unknown engine
- ▶ Chain saw
- ▶ Small- & medium rotating saw
- ▶ Large rotating saw
- ▶ Other/unknown saw
- ▶ Stationary music
- ▶ Mobile music
- ▶ Ice-cream truck
- ▶ Other/unknown music
- ▶ Rock drill
- ▶ Jackhammer
- ▶ Hoe ram
- ▶ Pile driver
- ▶ Other/unknown impact sound
- ▶ Car horn
- ▶ Car alarm
- ▶ Siren
- ▶ Reverse beeper
- ▶ Other/unknown alert signal
- ▶ Person or small group talking
- ▶ Person shouting
- ▶ Crowd
- ▶ Amplified speech
- ▶ Dog barking/whining
- ▶ Other/unknown human or animal vocalization sound
- ▶ Other/unknown sound
- ▶ Other/unknown construction sound

## Hybrid (2-stage) Annotation Task

### Stage 1

- ▶ Engines
- ▶ Powered sawing tools
- ▶ Music
- ▶ Impact sounds
- ▶ Alert signals
- ▶ Human and animal vocalization sounds
- ▶ Other/unknown sound
- ▶ Other/unknown construction sound

### Stage 2

- ▶ Impact sounds:
  - ▶ Rock drill
  - ▶ Jackhammer
  - ▶ Hoe ram
  - ▶ Pile driver
  - ▶ Other/unknown impact sound

# Varying Label Set Specificity

---

## General (Single-pass) Annotation Task

- ▶ Small-sounding engine
- ▶ Medium-sounding engine
- ▶ Large-sounding engine
- ▶ Other/unknown engine
- ▶ Chain saw
- ▶ Small- & medium rotating saw
- ▶ Large rotating saw
- ▶ Other/unknown saw
- ▶ Stationary music
- ▶ Mobile music
- ▶ Ice-cream truck
- ▶ Other/unknown music
- ▶ Rock drill
- ▶ Jackhammer
- ▶ Hoe ram
- ▶ Pile driver
- ▶ Other/unknown impact sound
- ▶ Car horn
- ▶ Car alarm
- ▶ Siren
- ▶ Reverse beeper
- ▶ Other/unknown alert signal
- ▶ Person or small group talking
- ▶ Person shouting
- ▶ Crowd
- ▶ Amplified speech
- ▶ Dog barking/whining
- ▶ Other/unknown human or animal vocalization sound
- ▶ Other/unknown sound
- ▶ Other/unknown construction sound

## Hybrid (2-stage) Annotation Task

### Stage 1

- ▶ Engines
- ▶ Powered sawing tools
- ▶ Music
- ▶ Impact sounds
- ▶ Alert signals
- ▶ Human and animal vocalization sounds
- ▶ Other/unknown sound
- ▶ Other/unknown construction sound

### Stage 2

- ▶ Impact sounds:
  - ▶ Rock drill
  - ▶ Jackhammer
  - ▶ Hoe ram
  - ▶ Pile driver
  - ▶ Other/unknown impact sound

## Specific (Multi-pass) Annotation Task

- ▶ Car horn

# Varying Label Set Specificity

---

## General (Single-pass) Annotation Task

- ▶ Small-sounding engine
- ▶ Medium-sounding engine
- ▶ Large-sounding engine
- ▶ Other/unknown engine
- ▶ Chain saw
- ▶ Small- & medium rotating saw
- ▶ Large rotating saw
- ▶ Other/unknown saw
- ▶ Stationary music
- ▶ Mobile music
- ▶ Ice-cream truck
- ▶ Other/unknown music
- ▶ Rock drill
- ▶ Jackhammer
- ▶ Hoe ram
- ▶ Pile driver
- ▶ Other/unknown impact sound
- ▶ Car horn
- ▶ Car alarm
- ▶ Siren
- ▶ Reverse beeper
- ▶ Other/unknown alert signal
- ▶ Person or small group talking
- ▶ Person shouting
- ▶ Crowd
- ▶ Amplified speech
- ▶ Dog barking/whining
- ▶ Other/unknown human or animal vocalization sound
- ▶ Other/unknown sound
- ▶ Other/unknown construction sound

## Hybrid (2-stage) Annotation Task

### Stage 1

- ▶ Engines
- ▶ Powered sawing tools
- ▶ Music
- ▶ Impact sounds
- ▶ Alert signals
- ▶ Human and animal vocalization sounds
- ▶ Other/unknown sound
- ▶ Other/unknown construction sound

### Stage 2

- ▶ Impact sounds:
  - ▶ Rock drill
  - ▶ Jackhammer
  - ▶ Hoe ram
  - ▶ Pile driver
  - ▶ Other/unknown impact sound

## Specific (Multi-pass) Annotation Task

- ▶ Jackhammer

# Varying Label Set Specificity

---

## General (Single-pass) Annotation Task

- ▶ Small-sounding engine
- ▶ Medium-sounding engine
- ▶ Large-sounding engine
- ▶ Other/unknown engine
- ▶ Chain saw
- ▶ Small- & medium rotating saw
- ▶ Large rotating saw
- ▶ Other/unknown saw
- ▶ Stationary music
- ▶ Mobile music
- ▶ Ice-cream truck
- ▶ Other/unknown music
- ▶ Rock drill
- ▶ Jackhammer
- ▶ Hoe ram
- ▶ Pile driver
- ▶ Other/unknown impact sound
- ▶ Car horn
- ▶ Car alarm
- ▶ Siren
- ▶ Reverse beeper
- ▶ Other/unknown alert signal
- ▶ Person or small group talking
- ▶ Person shouting
- ▶ Crowd
- ▶ Amplified speech
- ▶ Dog barking/whining
- ▶ Other/unknown human or animal vocalization sound
- ▶ Other/unknown sound
- ▶ Other/unknown construction sound

## Hybrid (2-stage) Annotation Task

### Stage 1

- ▶ Engines
- ▶ Powered sawing tools
- ▶ Music
- ▶ Impact sounds
- ▶ Alert signals
- ▶ Human and animal vocalization sounds
- ▶ Other/unknown sound
- ▶ Other/unknown construction sound

### Stage 2

- ▶ Impact sounds:
  - ▶ Rock drill
  - ▶ Jackhammer
  - ▶ Hoe ram
  - ▶ Pile driver
  - ▶ Other/unknown impact sound

## Specific (Multi-pass) Annotation Task

- ▶ Jackhammer

# Varying Label Set Specificity

---

## General (Single-pass) Annotation Task

- ▶ Small-sounding engine
- ▶ Medium-sounding engine
- ▶ Large-sounding engine
- ▶ Other/unknown engine
- ▶ Chain saw
- ▶ Small- & medium rotating saw
- ▶ Large rotating saw
- ▶ Other/unknown saw
- ▶ Stationary music
- ▶ Mobile music
- ▶ Ice-cream truck
- ▶ Other/unknown music
- ▶ Rock drill
- ▶ Jackhammer
- ▶ Hoe ram
- ▶ Pile driver
- ▶ Other/unknown impact sound
- ▶ Car horn
- ▶ Car alarm
- ▶ Siren
- ▶ Reverse beeper
- ▶ Other/unknown alert signal
- ▶ Person or small group talking
- ▶ Person shouting
- ▶ Crowd
- ▶ Amplified speech
- ▶ Dog barking/whining
- ▶ Other/unknown human or animal vocalization sound
- ▶ Other/unknown sound
- ▶ Other/unknown construction sound

## Hybrid (2-stage) Annotation Task

### Stage 1

- ▶ Engines
- ▶ Powered sawing tools
- ▶ Music
- ▶ Impact sounds
- ▶ Alert signals
- ▶ Human and animal vocalization sounds
- ▶ Other/unknown sound
- ▶ Other/unknown construction sound

### Stage 2

- ▶ Impact sounds:
  - ▶ Rock drill
  - ▶ Jackhammer
  - ▶ Hoe ram
  - ▶ Pile driver
  - ▶ Other/unknown impact sound

## Specific (Multi-pass) Annotation Task

▶ etc....



# Example of “General” Task on Zooniverse

The screenshot displays the 'CLASSIFY' page of the 'SOUNDS OF NEW YORK CITY (SONYC)' project. The top navigation bar includes 'ABOUT', 'CLASSIFY' (highlighted), 'TALK', 'COLLECT', 'RECENTS', and 'LAB'. The main content area is split into two sections:

- Left Section:** A spectrogram visualization of a sound sample. Below it is a playback control bar showing '0:00 / 0:10' and a volume slider.
- Right Section:** A list of 30 possible sound categories, organized in a 10x3 grid. The categories are:

Category		
Small-sounding engine	Small/medium rotating saw	Ice cream truck
Medium-sounding engine	Large rotating saw	Other/unknown music
Large-sounding engine	Other/unknown saw	Small group talking
Other/unknown engine	Car horn	Person shouting
Rock drill	Car alarm	Crowd
Jackhammer	Siren	Amplified speech
Hoe ram	Reverse beeper	Dog barking/whining
Pile driver	Other/unknown alert signal	Other/unknown human or animal vocalization sound
Other/unknown impact sound	Stationary music	Other/unknown construction sound
Chainsaw	Mobile music	Other/unknown sound

At the bottom right of the category list, it says 'Showing 30 of 30' and 'Clear filters'. A green 'Done' button is located at the bottom center of the interface.

# Summary of Next Steps

---

- Begin Zooniverse Annotation Campaign in the next couple of weeks
- Initially experiment with variations in task design
- Continue with best task design
- Eventually transition to an active sampling scheme to potentially reduce the number of overall annotations (collecting at least 1000 example per class in total)

# Questions?

---

- Come see Ana Elisa's poster later to learn more details about the Zooniverse project and its associated upcoming experiment!
- Thanks to help from Ana Elisa Mendez, Ayanna Seals, Justin Salamon, Graham Dove, Juan Pablo Bello, and Oded Nov