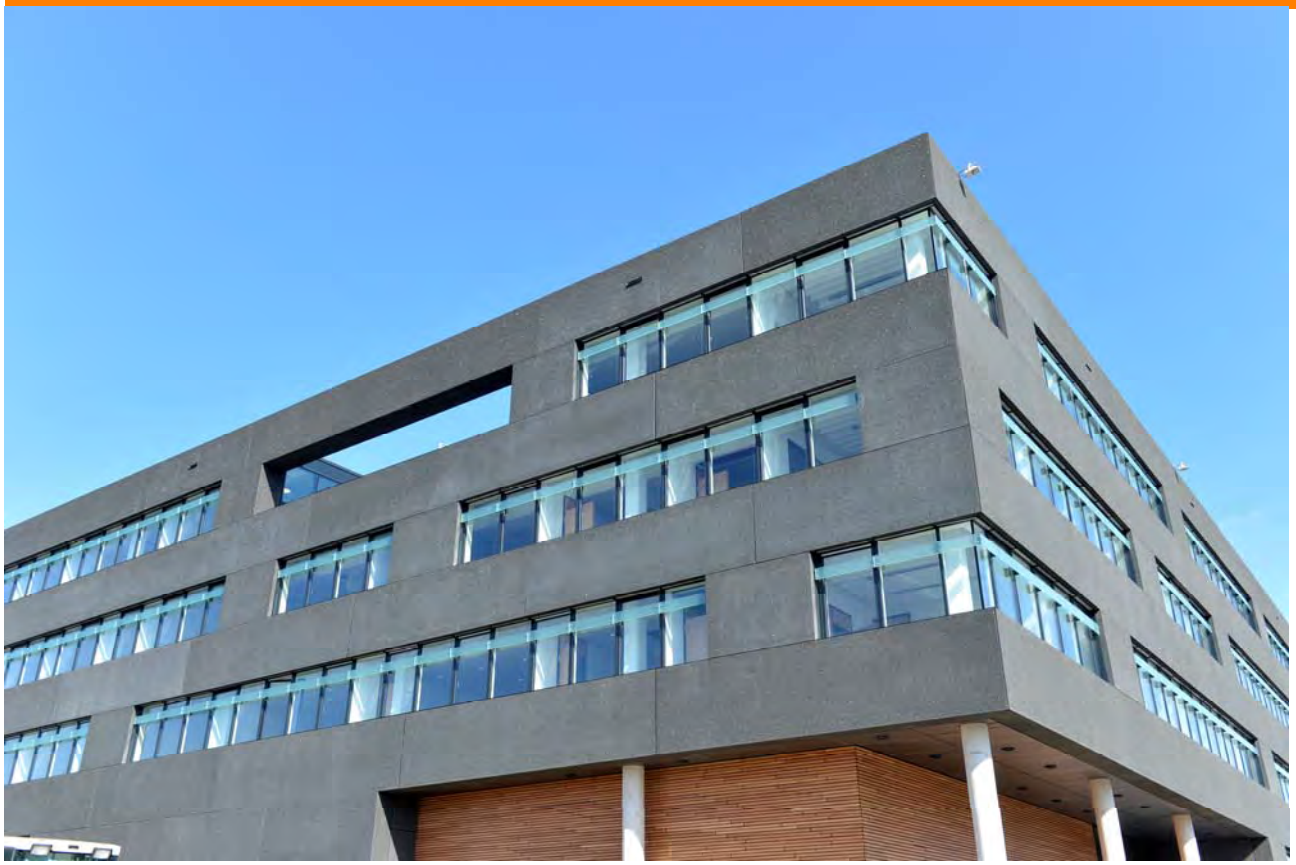
Guiding Visual Search Tasks Using Gaze-Contingent Auditory Feedback



CITEC Building at Bielefeld

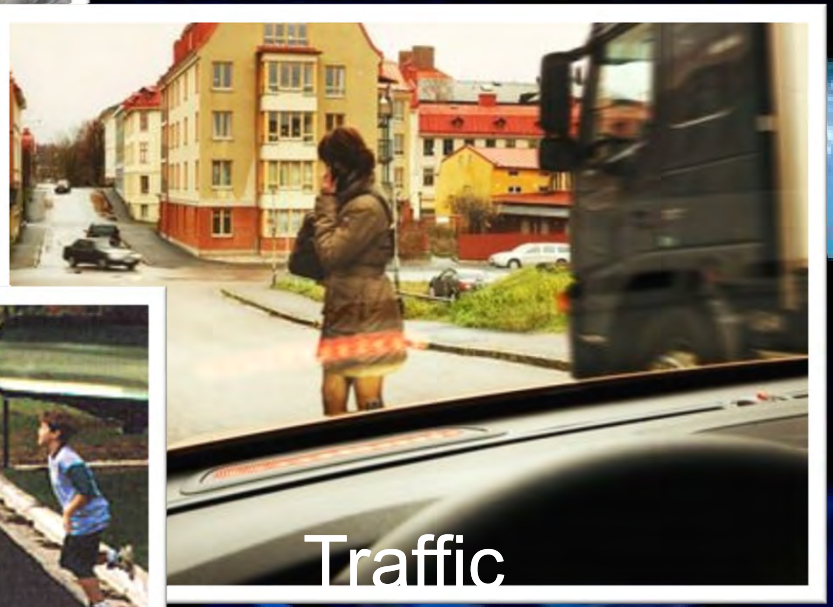
Authors:

- Viktor Losing
- Lukas Rottkamp
- Michael Zeunert
- Thies Pfeiffer

Searching is part of our live....

Sometimes it is important not to miss something…

Control room



Traffic

Traffic

Motivation

- Cognitive Assistance Technology
  - detects events of interest (EoI)
  - monitors the attention of the user
  - alerts user if necessary
  - **guides user's attention towards the EoI ← Focus of this talk**

Target Examples: Decision Support for *Stone-and-Mortar* Shopping

Target Examples:
Tutoring System for Chess

Conventional Approaches for Guiding Attention

Common visual approaches

- Signs, Arrows
- Signal colors (red)
- Animations (flashing, shaking screen)

But sometimes visual signals are not possible or wanted:

- visual augmentation might be impossible
  - no display, black-box target systems, real-world
- visual augmentations might occlude relevant content
  - medical images, driving

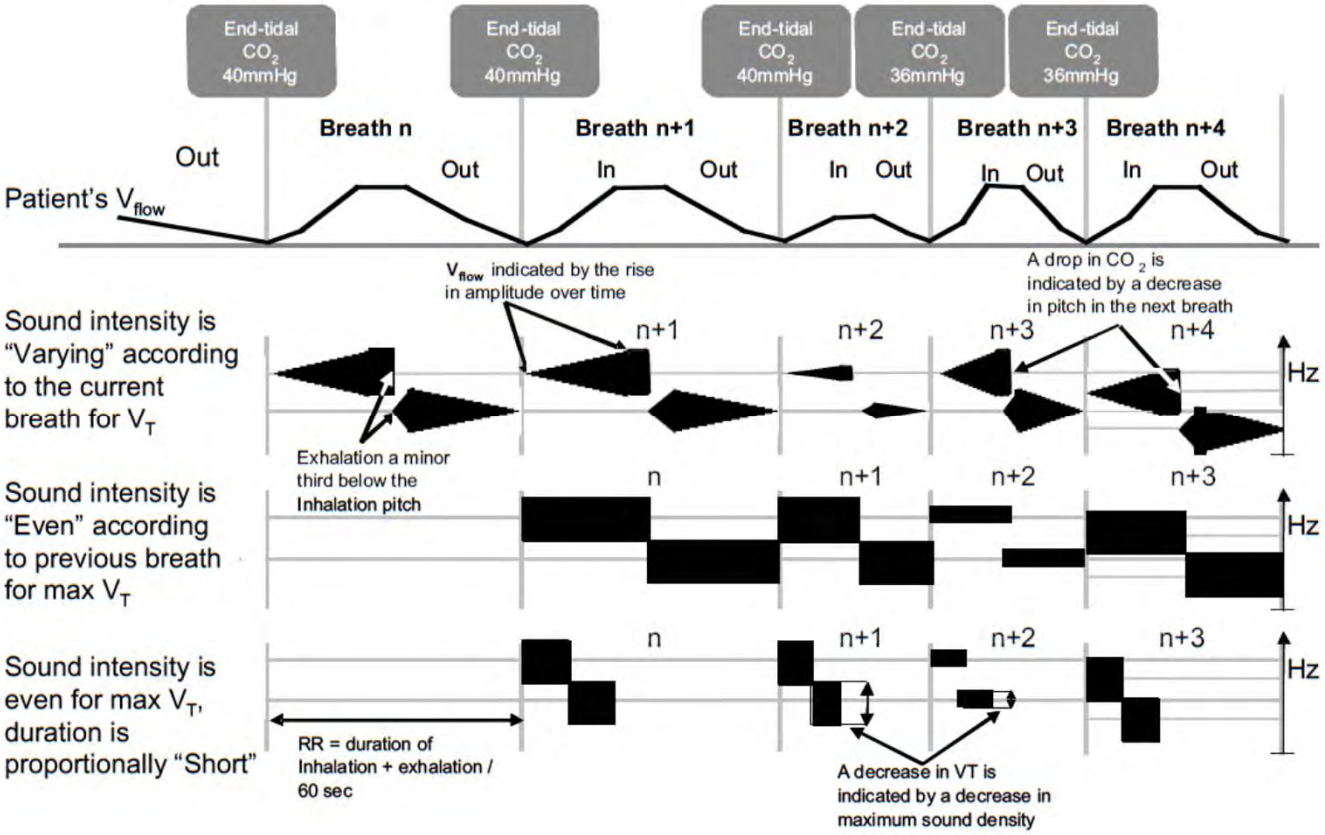
Solution: Cross-Modal Guidance

- Possible Solution:
  - Cross-modal guidance
- Questions
  - How can cross-modal guidance be realized?
  - Can people make us of this cross-modal feedback?
    - How easy is it to use and learn?
    - What is the precision one can achieve?
    - What performance increases can be achieved?

Idea:
Auditory Feedback

"Topfschlagen" / Hit-A-Pot

Related Work on Sonification

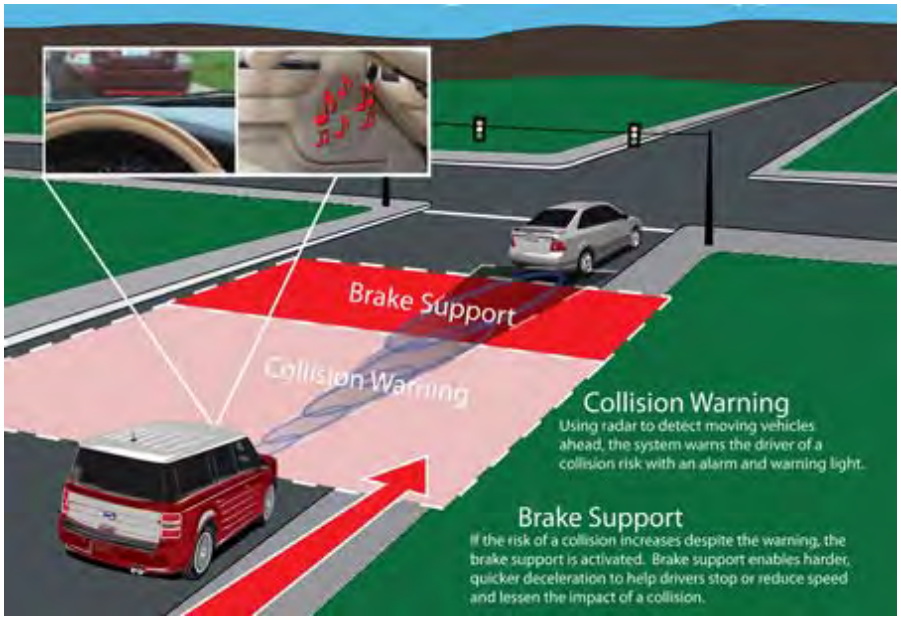


Watson, M., and Sanderson, P. Sonication supports eyes-free respiratory monitoring and task time-sharing. Human Factors: The Journal of the Human Factors and Ergonomics Society 46, 3 (2004), 497-517.

- Sonification of medial data as replacement or addition to visualizations
- Idea: during other tasks, doctors could maintain auditory monitoring
- Sonification helps to timeshare

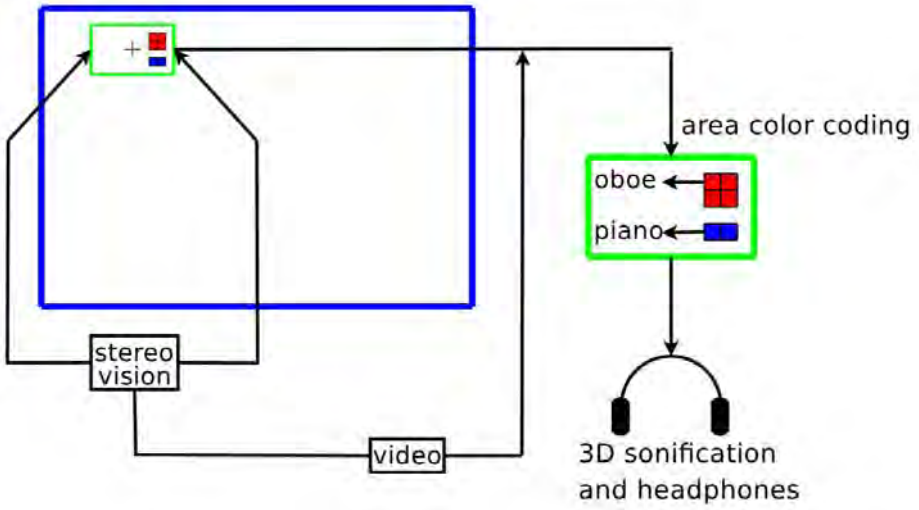
- **Sonification of state or state changes**

Related Work on Sonification



- Sonification of distances (1-dimensional)
- **Human/car in the loop**

Related Work
Interactive Sonificiation for the Blind



- **2D sonification, human in the loop, touch interaction, guidance**

- Deville, B., Bologna, G., Vinckenbosch, M., and Pun, T. Guiding the focus of attention of blind people with visual saliency (2008).
- Hunt, A., Hermann, T., and Pauletto, S. Interacting with sonication systems: closing the loop. In Information Visualisation, 2004. IV 2004. Proceedings. Eighth International Conference on, IEEE (2004), 879-884.
- Hermann, T., Höner, O., and Ritter, H. Acoumotion an interactive sonification system for acoustic motion control. In Gesture in Human-Computer Interaction and Simulation. Springer, 2006, 312-323.

Related Work
Gaze (or better Eye Movements) and Sonification

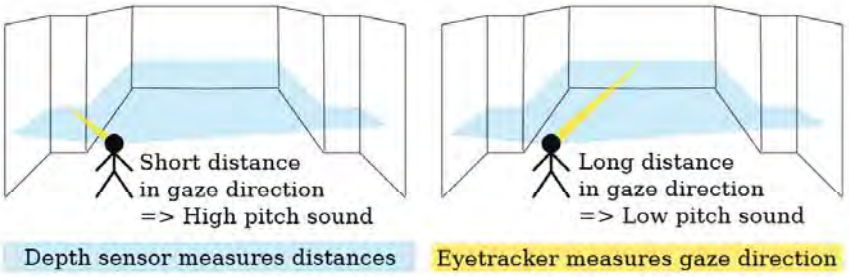
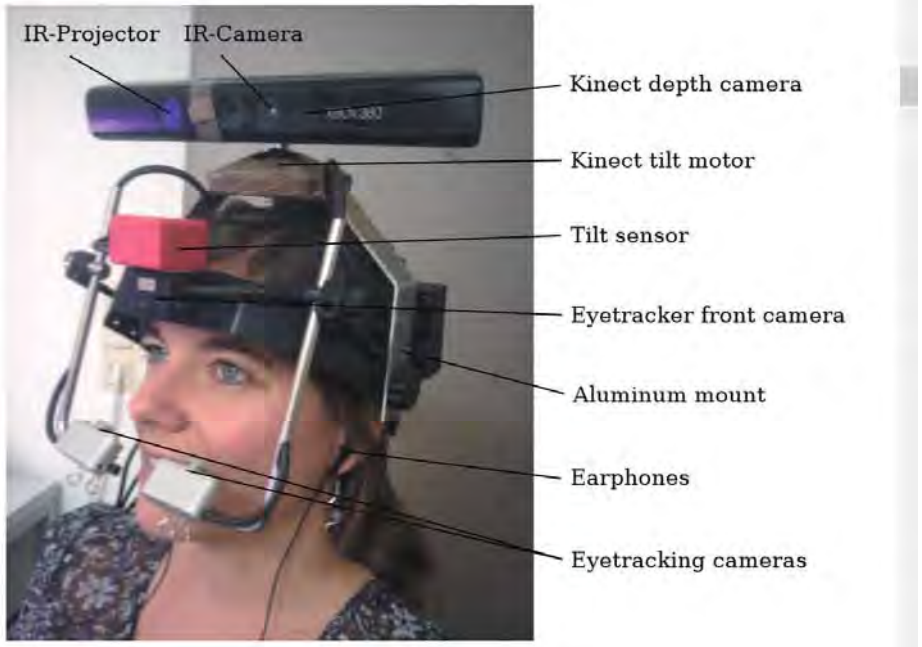


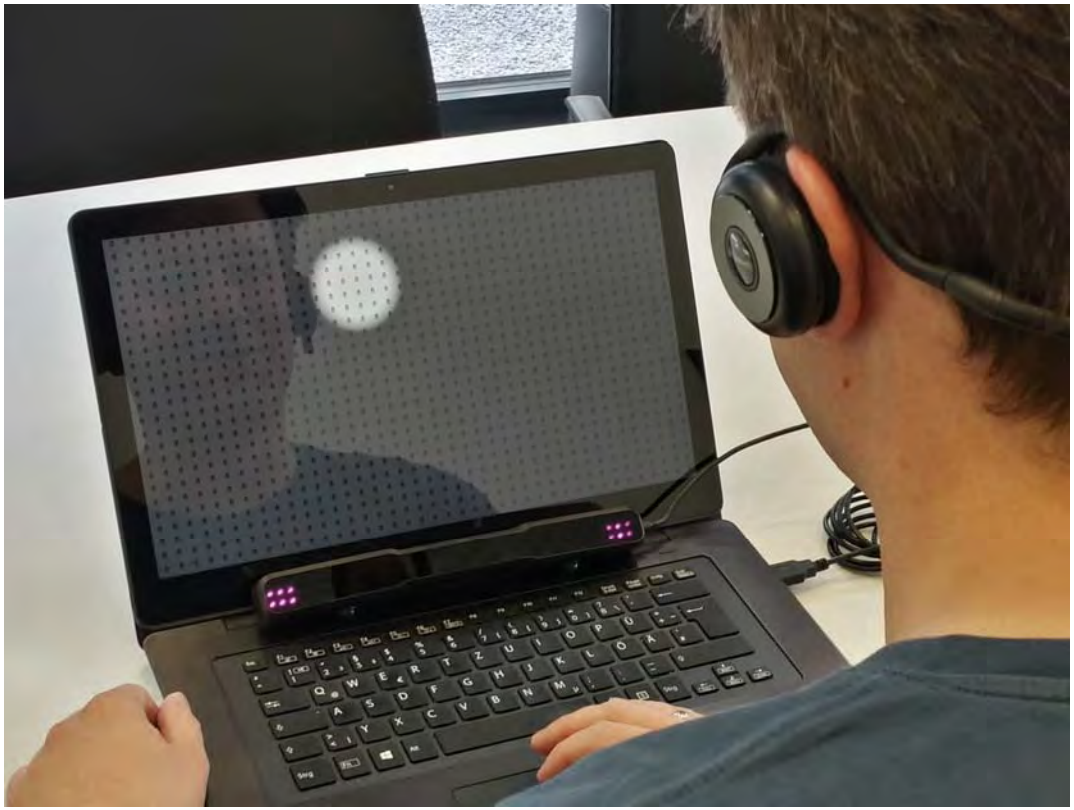
Fig. 1. Functionality of the *Auditory Night Sight (ANS)* from a user's perspective. A sine tone is played via earphones, pitch coding the measured distance (depicted in blue) in gaze direction (yellow).

The *ANS* is intended to help the user orient and navigate. Figure 1 shows the system from a user's perspective. The working principles of the *ANS* are the following.



- **2D sonification, human in the loop, eye movement interaction, no guidance**
  - Twardon, L., Koesling, H., Finke, A., and Ritter, H. Gaze-contingent audio-visual substitution for the blind and visually impaired. In Proceedings of the 7th International Conference on Pervasive Computing Technologies for Healthcare, ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering) (2013), 129-136.

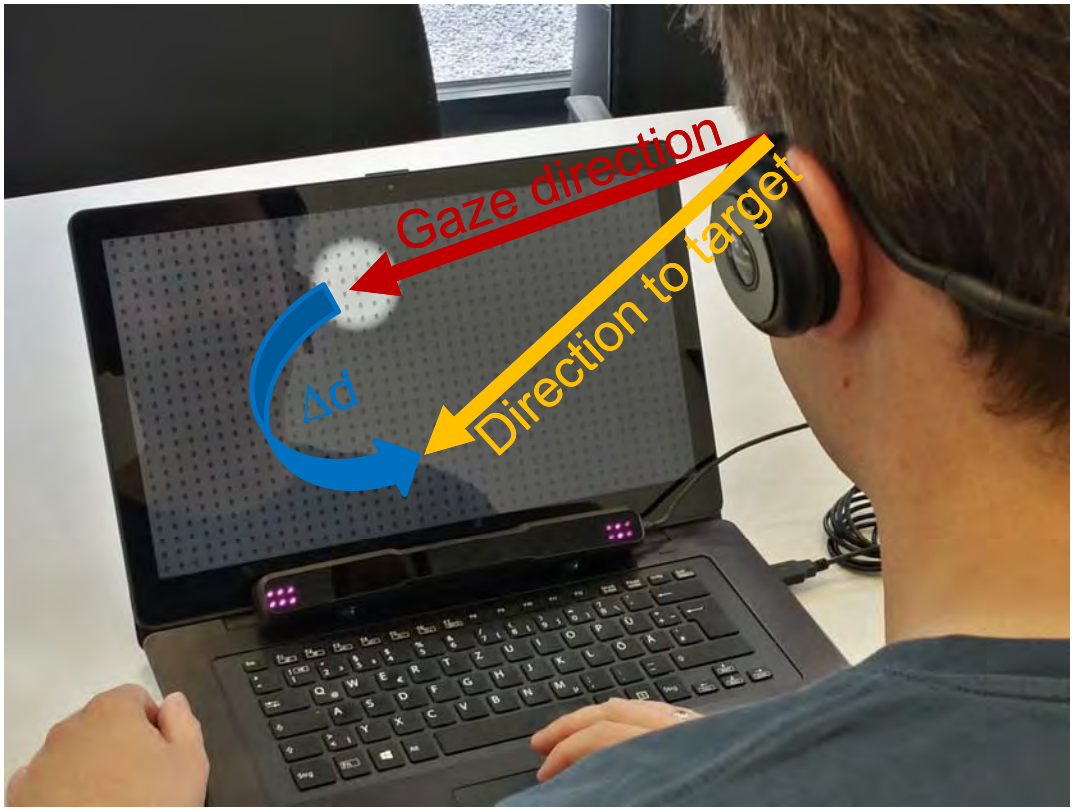
Our approach



Test scenario with a mobile remote eyetracker, display shows fixated area for demonstration purposes.

- **Gaze-contingent guidance of visual attention**
- **Auditory feedback**
- **Sonification of distance to the target in 2D**

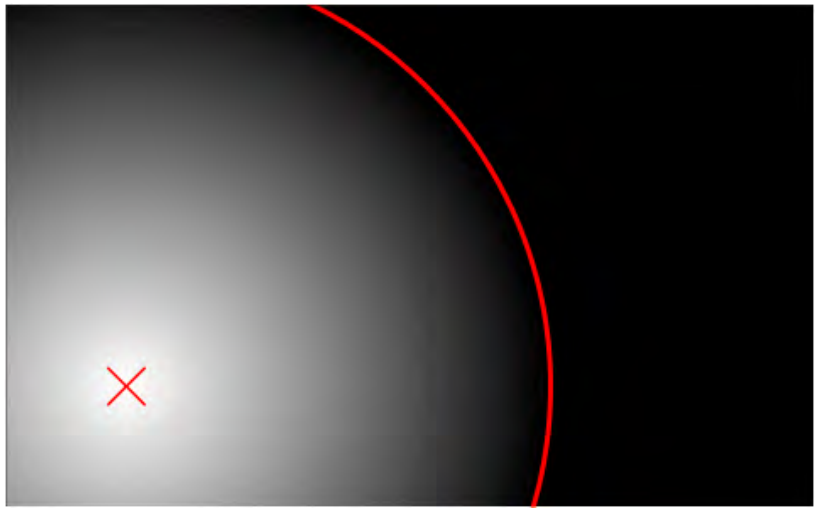
Our approach



Test scenario with a mobile remote eyetracker, display shows fixated area for demonstration purposes.

1. **Measure gaze direction**
   (MyGaze system, 30Hz, 0.5°)
2. **Calculate Δd to target**
3. **Map to sound**

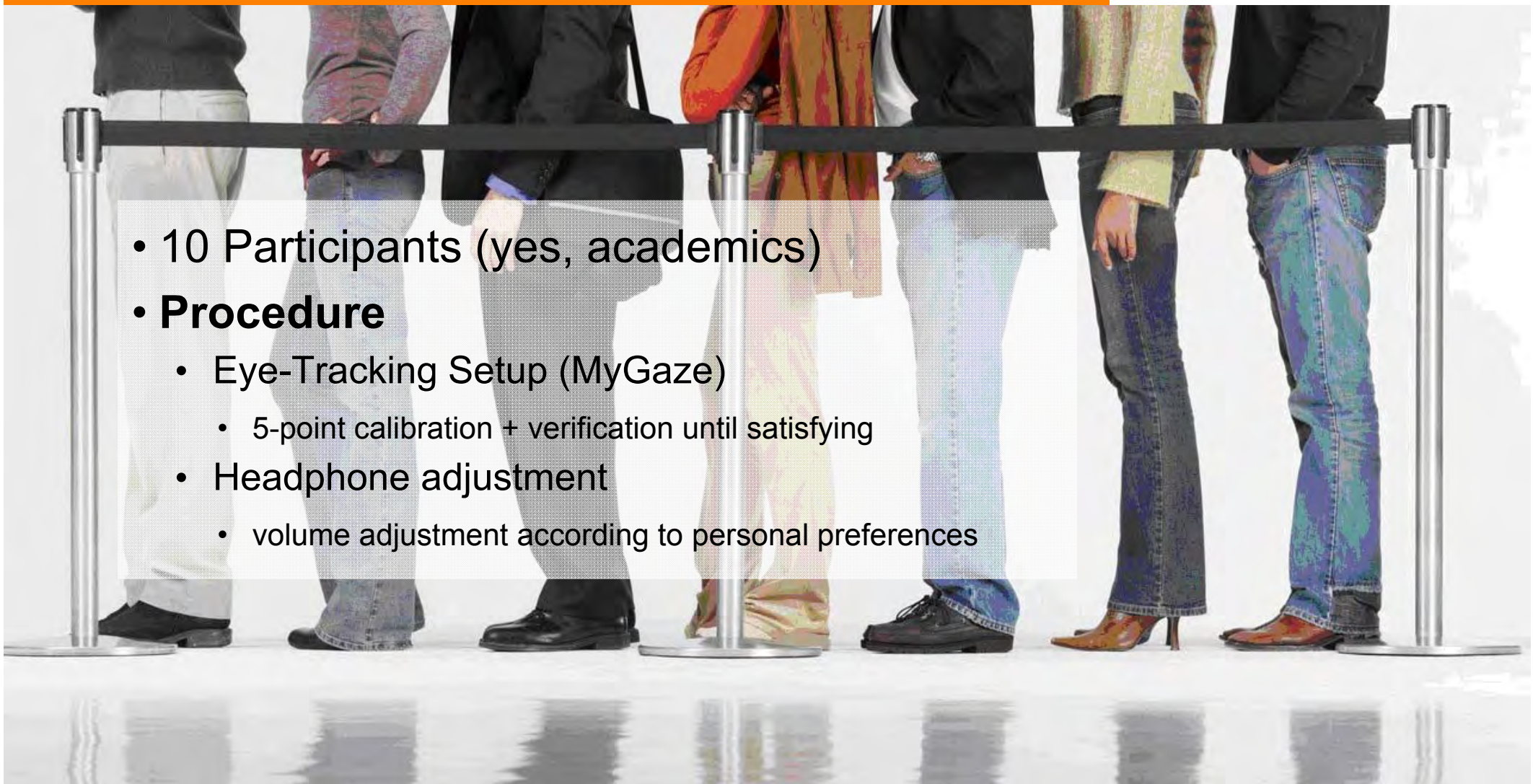
Mapping Distance/Error to Sound



Linear mapping in 1/4th of the screen real-estate, constant sound everywhere else.

- Pilot Study:
  - tried linear mapping, capped exponential, bell-curved over the whole screen real-estate
  - Rejected, either to little information (linear) or unnatural impression (user feedback)
- Final approach:
  - Localized Linear mapping
    - good slope to produce audible differences
    - requires “probing”

The Study

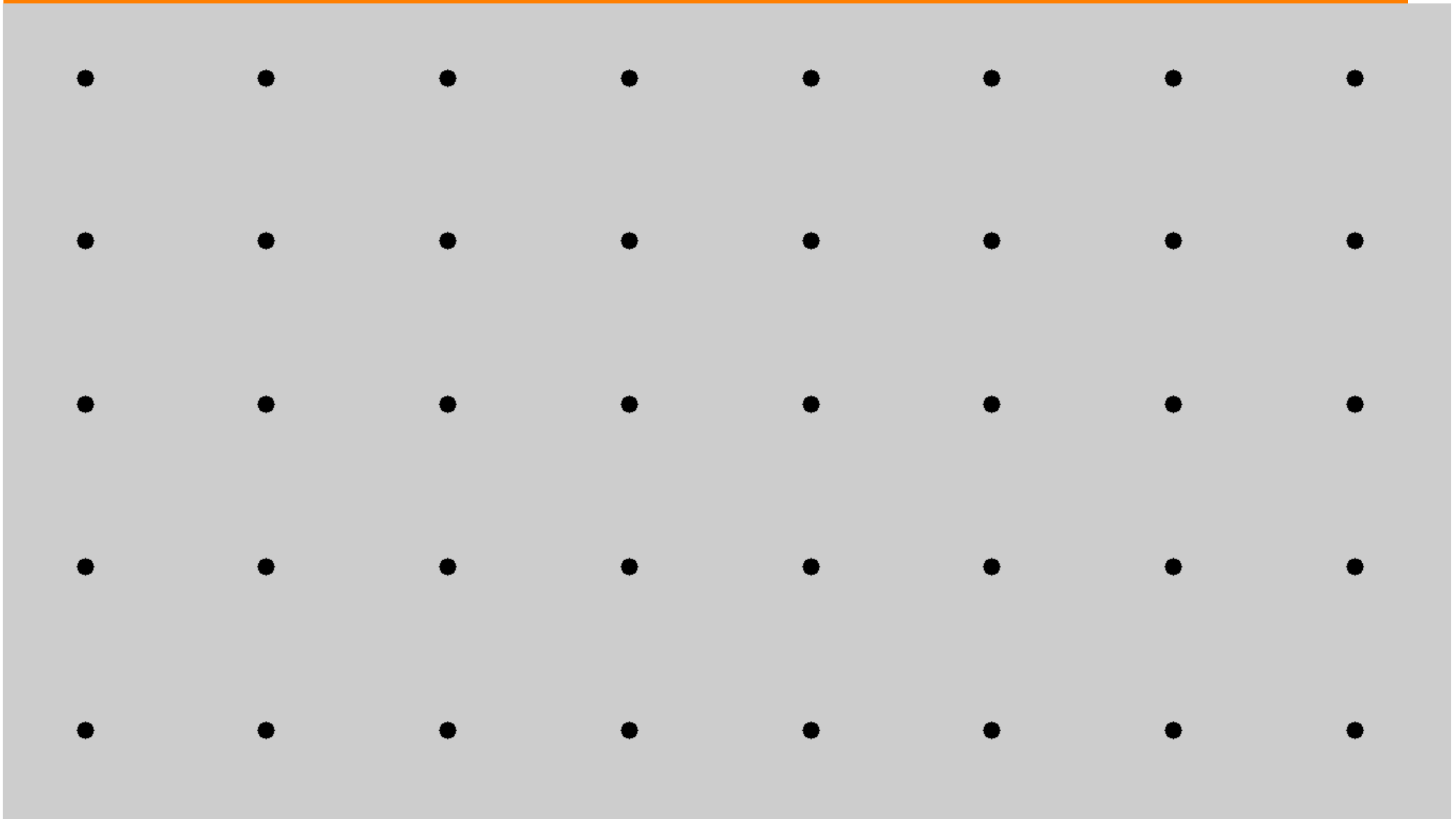


- 10 Participants (yes, academics)
- **Procedure**
  - Eye-Tracking Setup (MyGaze)
    - 5-point calibration + verification until satisfying
  - Headphone adjustment
    - volume adjustment according to personal preferences

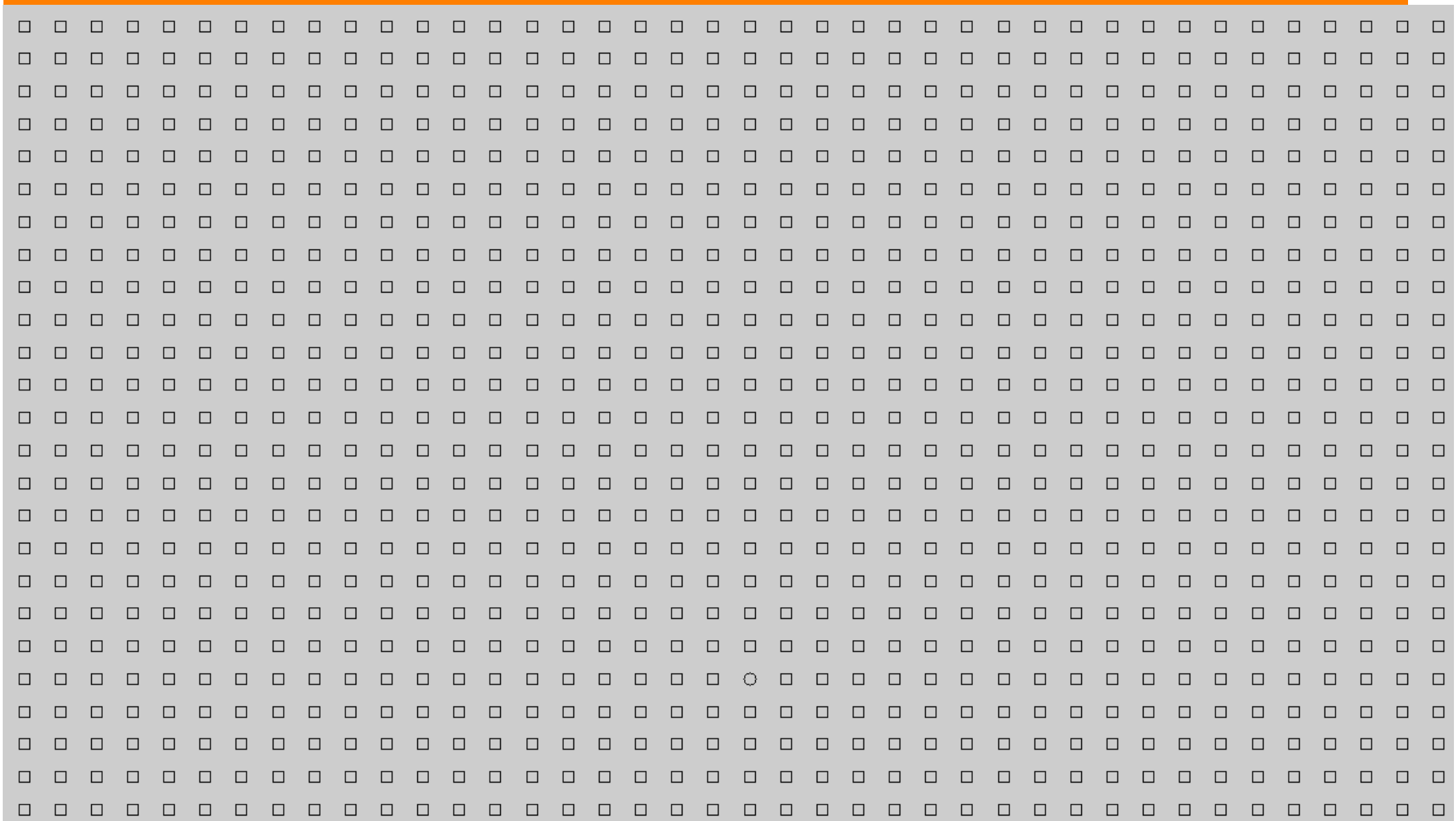
The Study Trials

- Training: 10 training trials
  Identify the target point based on acoustic guidance alone
- Visual search tasks
  - 20 Shape Trials, 20 Number Trials, 9 Cartoon Trials
  - User clicks on target item to complete trial
  - Presentation **either with or without acoustic guidance**
  - Half of the participants got acoustic guidance on even trias, the other on odd trials
  - Time per trial restricted to 90 seconds

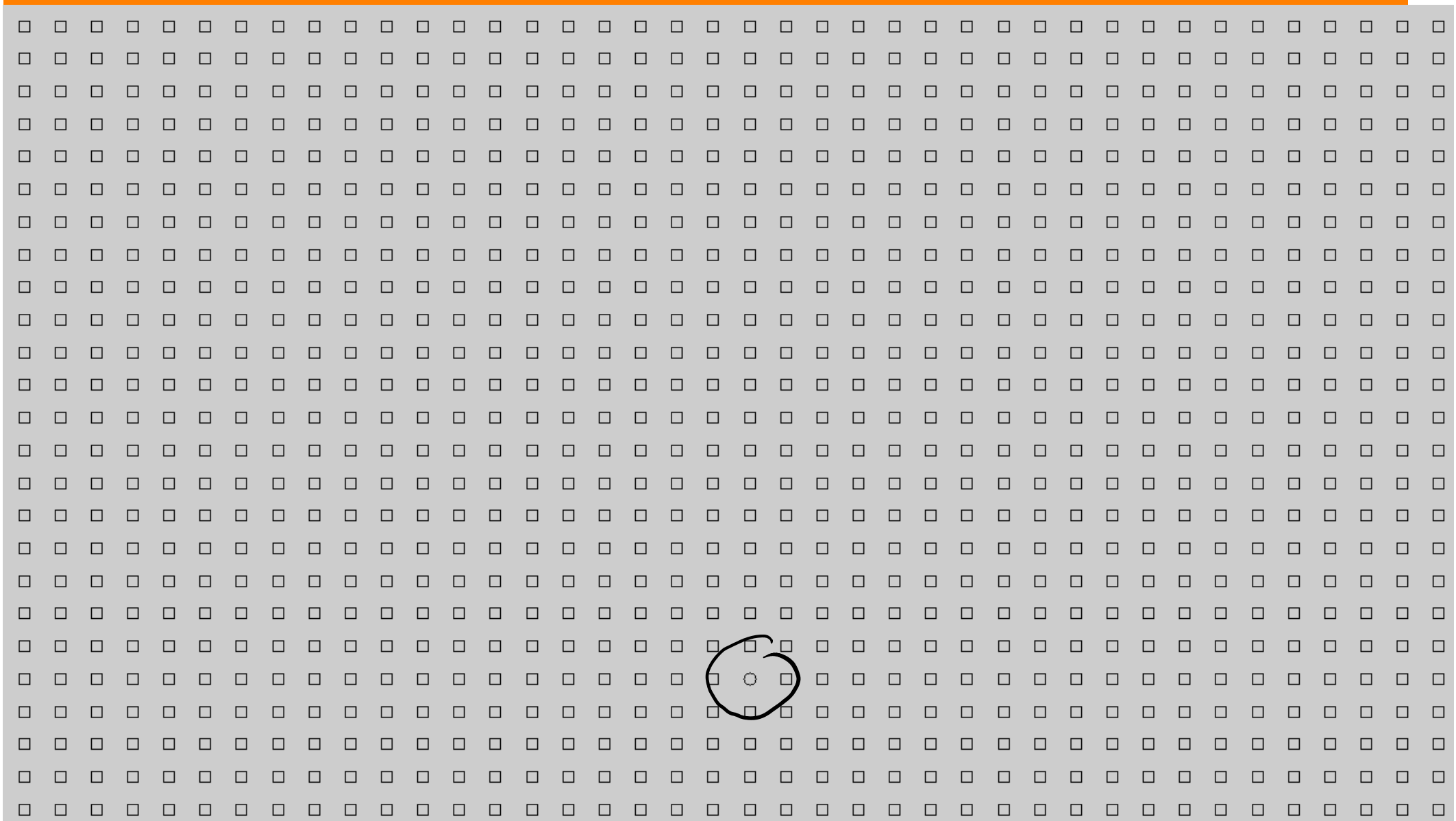
Training Set



Shapes Task: Search for Circle in Squares



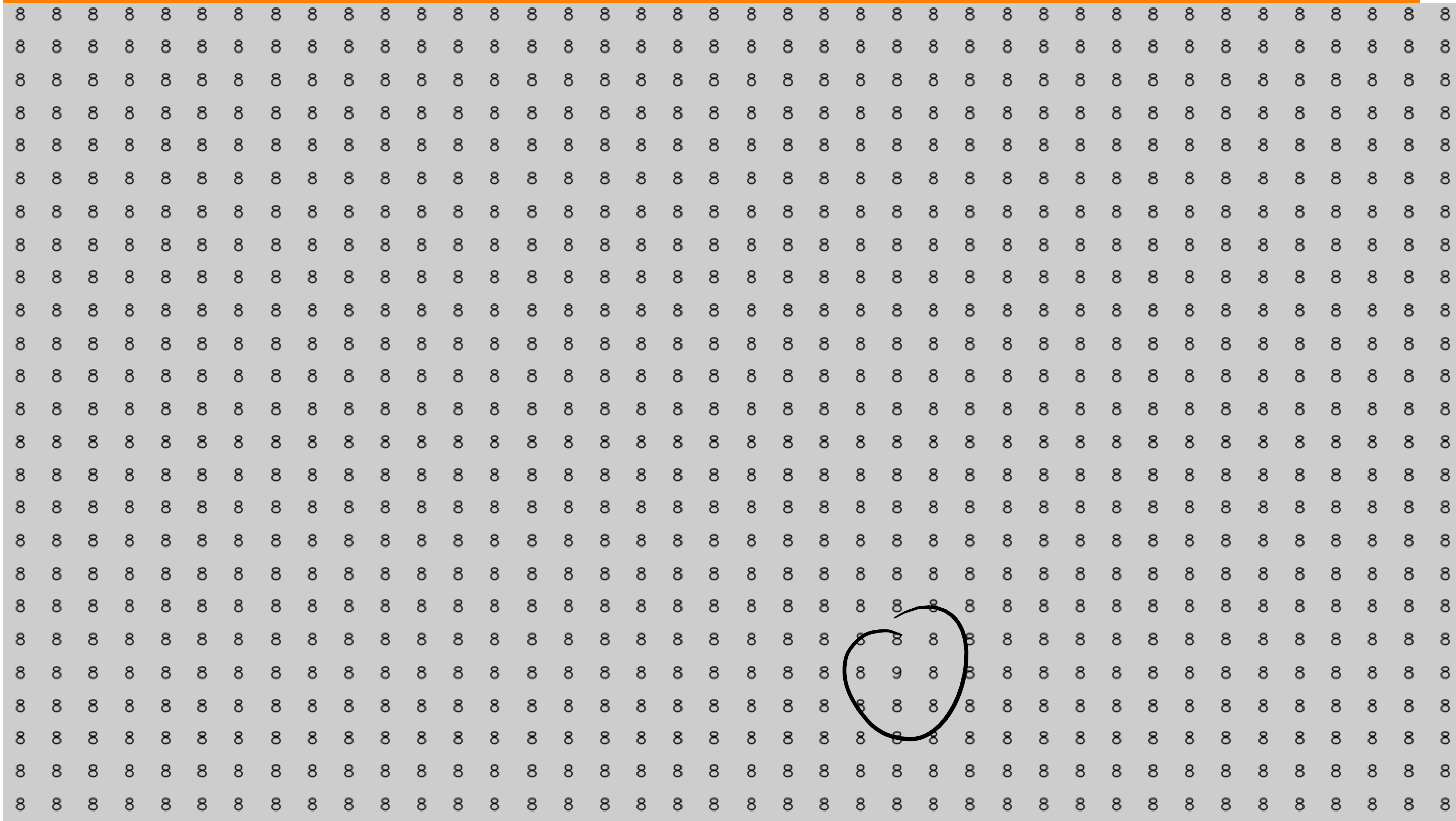
Shapes Task: Search for Circle in Squares



Numbers Task: Search for 9 in 8s

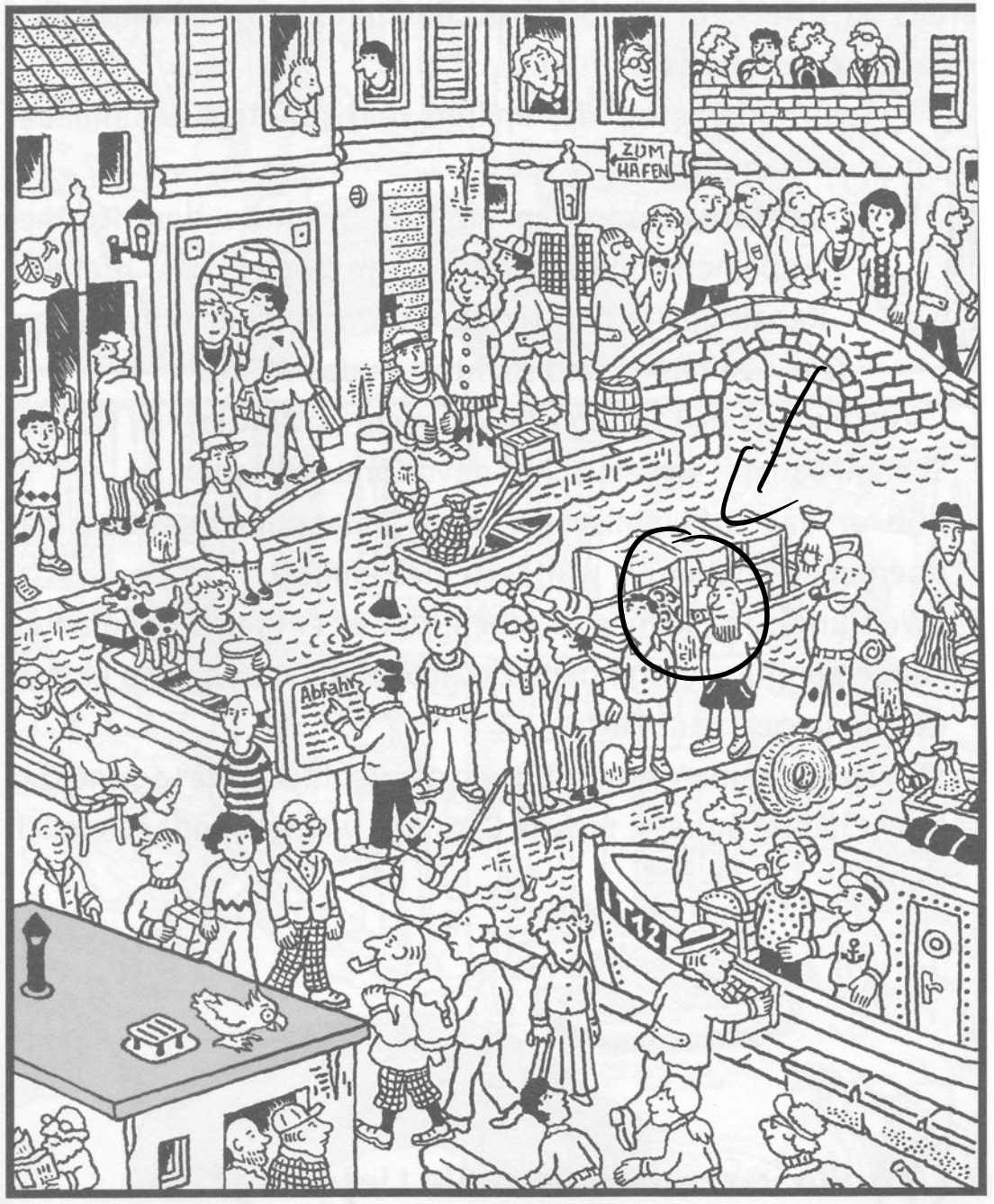
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	9	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8

Numbers Task: Search for 9 in 8s

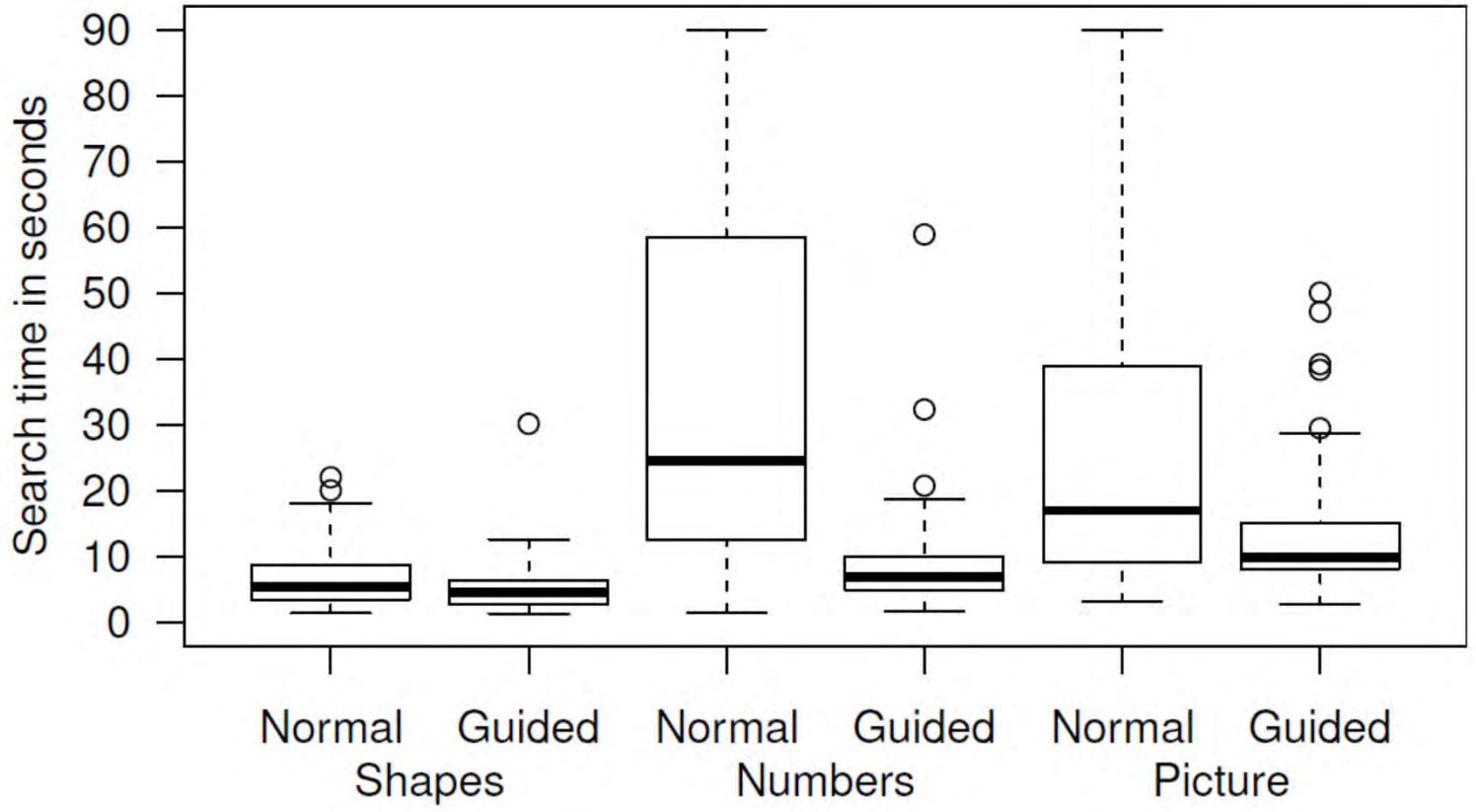


Finde den oben abgebildeten Mann rechts im Bild.

Finde den oben abgebildeten Mann rechts im Bild.



Does acoustic guidance increase performance?



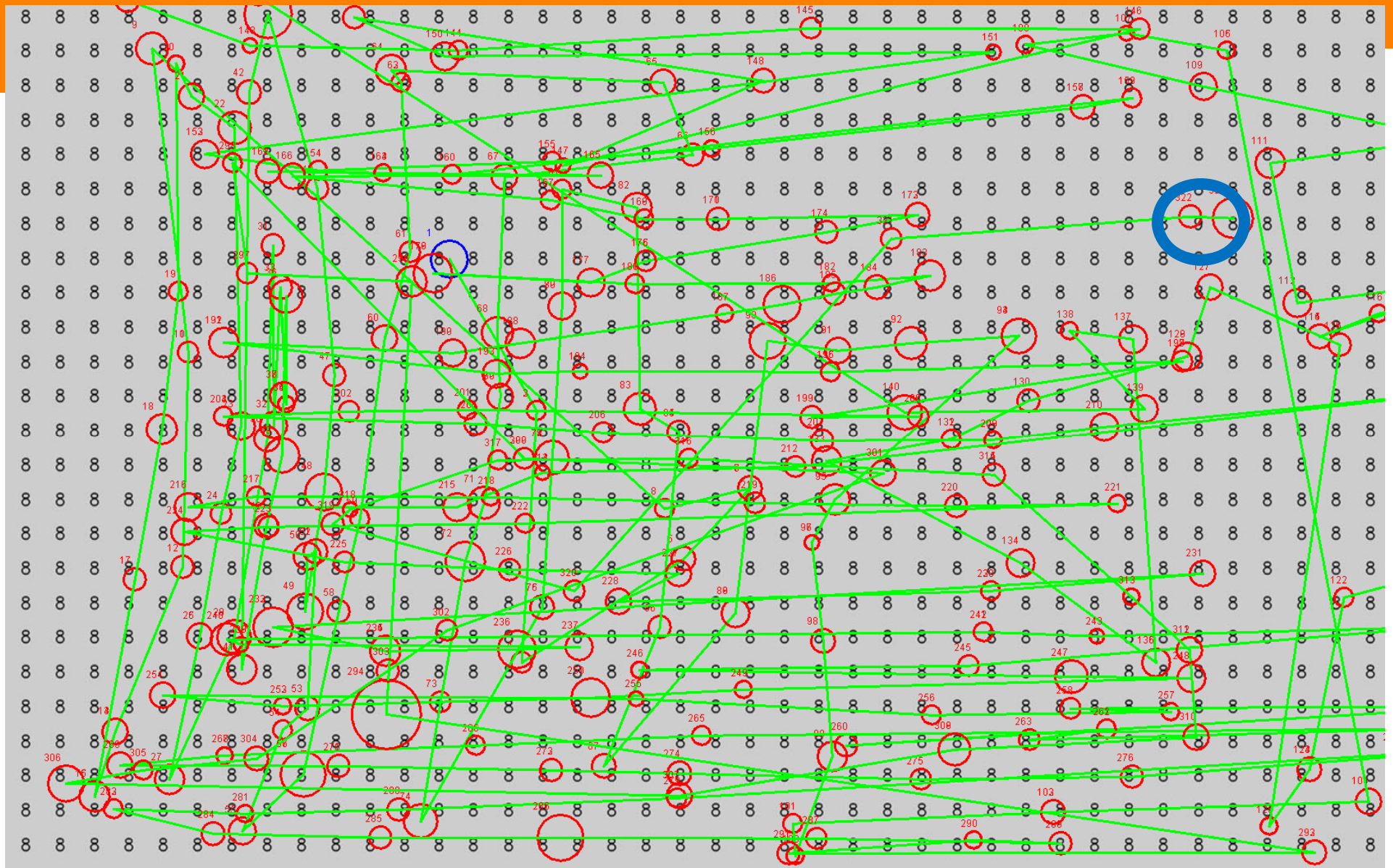
Does acoustic guidance increase performance?

Results

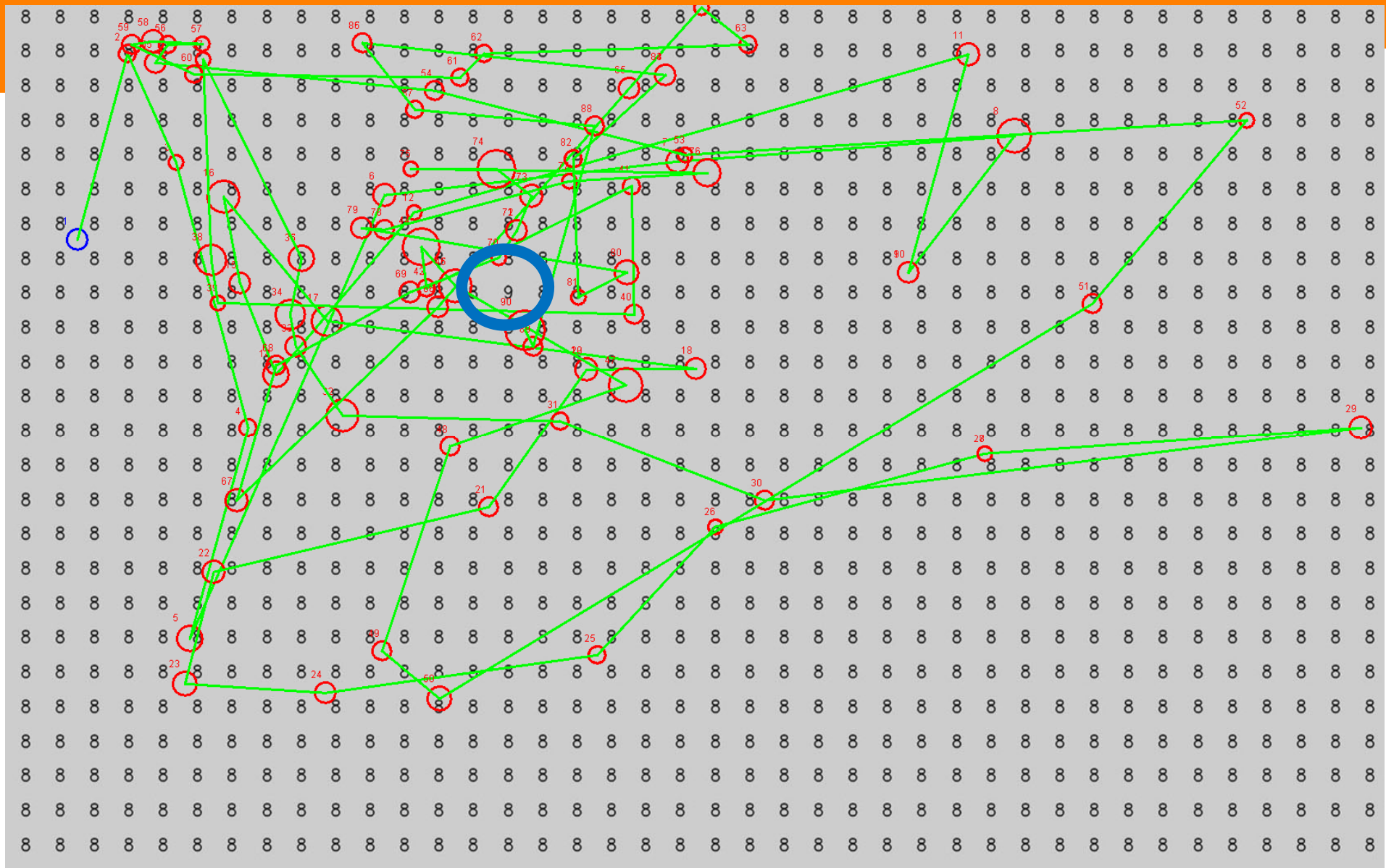
User #	Shapes OFF	Shapes ON	Numbers OFF	Numbers ON	Picture OFF	Picture ON
1	4.0	4.0	17.6	*5.9	18.0	15.6
2	10.2	*4.5	61.7	*7.8	38.9	9.0
3	7.9	5.9	36.1	16.6	31.4	24.9
4	5.1	4.1	39.0	*9.2	22.9	11.4
5	9.2	8.9	45.1	*8.0	30.7	27.3
6	4.9	4.1	53.3	*9.0	22.9	x6.2
7	6.0	4.5	31.7	*7.4	12.9	19.0
8	7.1	5.9	32.5	*6.5	29.2	8.6
9	5.1	5.1	19.7	*6.4	10.3	16.9
10	7.4	4.9	33.2	*7.2	46.7	8.6
MEAN	6.7	5.2	37.0	8.4	26.4	14.7

- Trend to increase performance
- Significant for number task

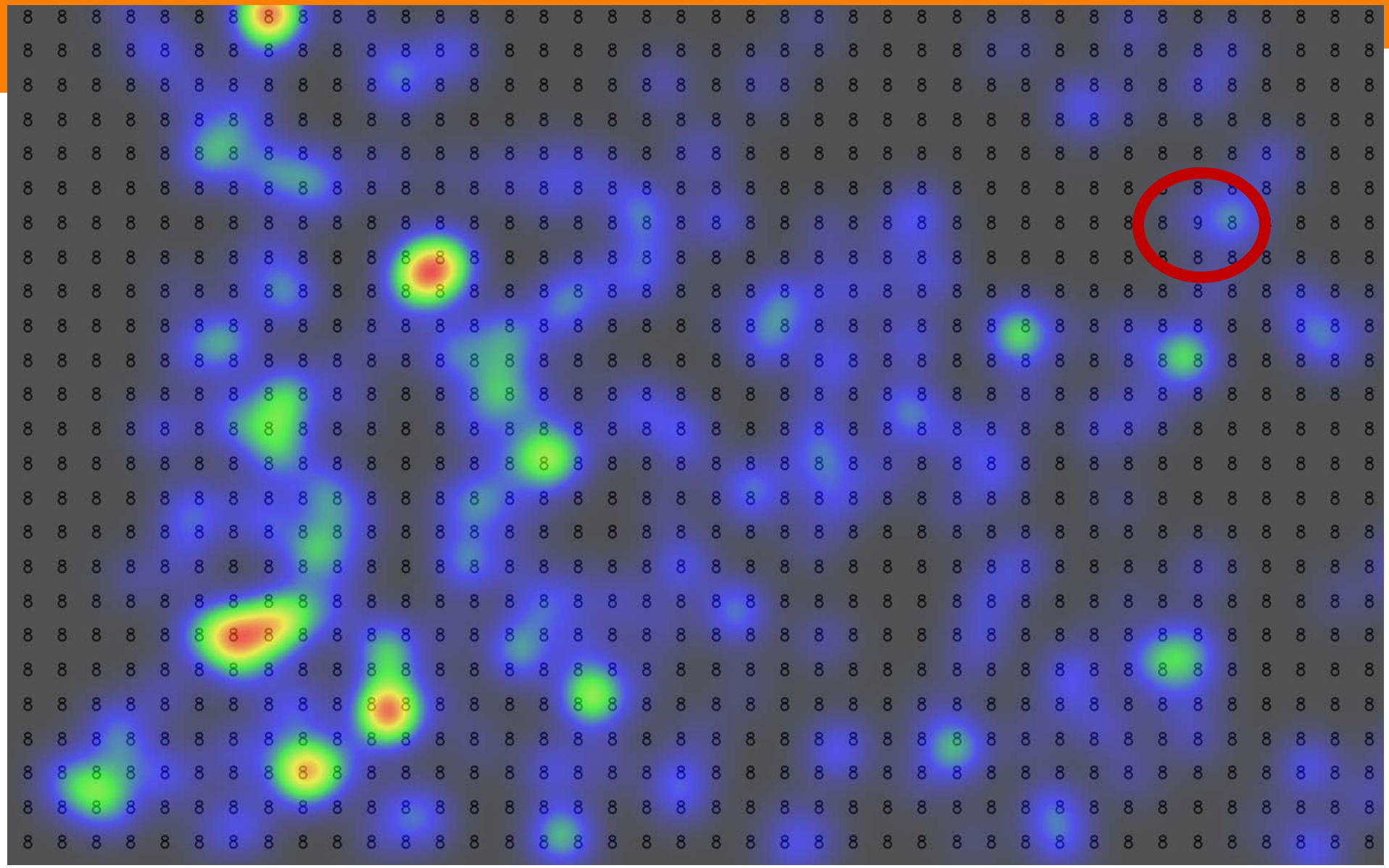
Scanpath for example trial without acoustic guidance



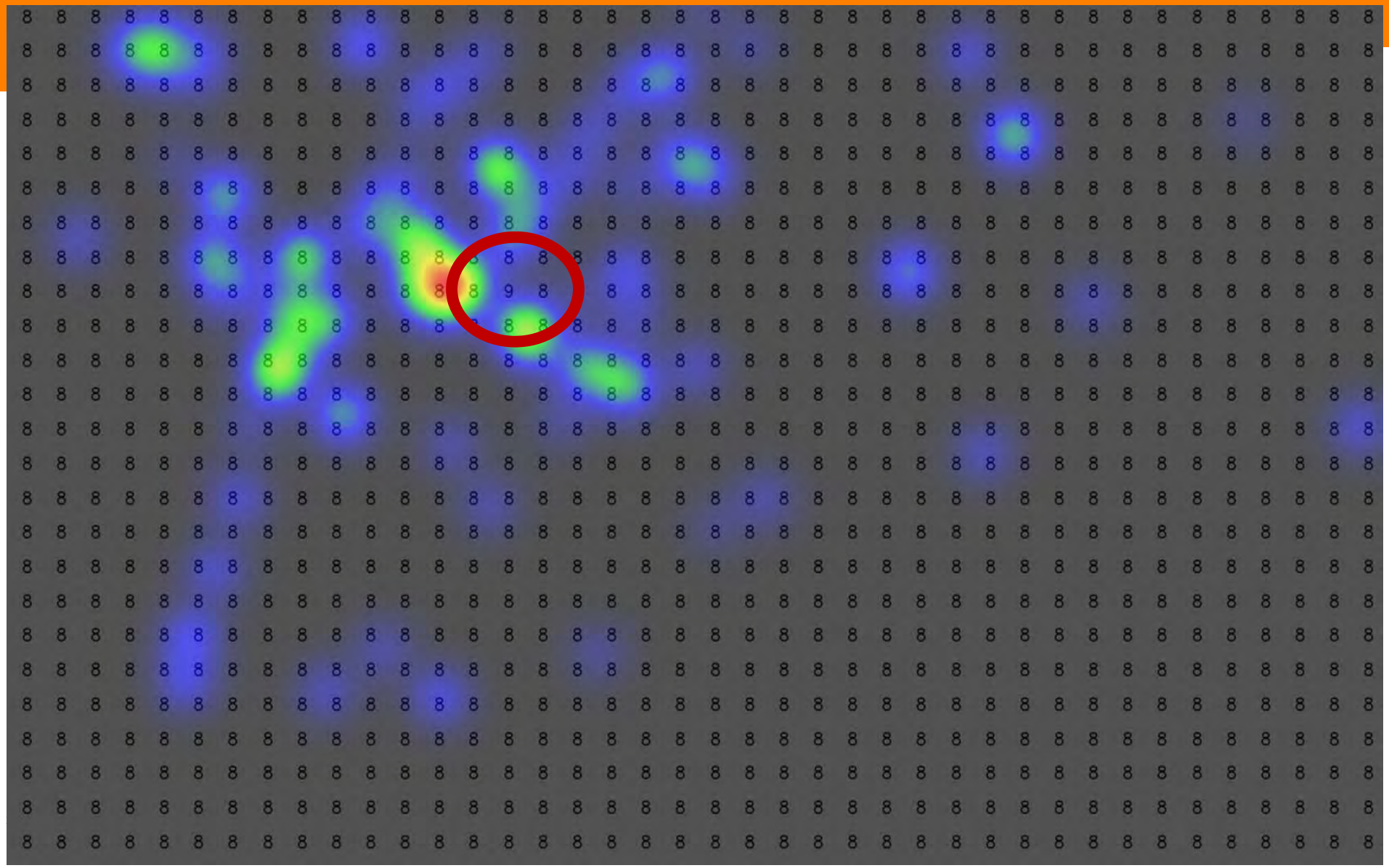
Scanpath for example trial with acoustic guidance



Heatmap for example trial without acoustic guidance



Heatmap for example trial with acoustic guidance



Participant feedback

- Participants **noticed no audio delay** (~150ms)
- Most participants reported **no problem with accuracy**
- Participants strategy (according their own words):
  - Use acoustic guidance for localizing area
  - Without guidance following a row- or column-wise search pattern
- All participants reported the system to be helpful and well-functioning
- Some participants suggested to improve the sonification tone, e.g. by using more natural sounds

Conclusion

- Promising first results
  - Even with simple mapping approach
  - Improvement even in very simple tasks
- Interesting strategy
  - Global orientation using acoustic guidance
  - Localized search using vision
  - => Switching from one modality to the other
  - However: Strategy most likely depends on the context and mapping

Possible next steps

- Potential for improvements
  - Here we used a mapping from 2D (position) to 1D (volume)
    - Could also use other parameters (pitch, 2D or 3D position)
  - Different sounds (samples (Hot/Cold ;-))
  - Instead of absolute location: relative location, directional
  - Multiple targets encoded differently (samples)
- Validation
  - Comparing with visual feedback/guidance
- New target area
  - 3D mobile eye tracking (shopping, chess, infoviz)