

Exploring Effective Relationships between Visual-Audio Channels in Data Visualization

Supplementary Figures

Supplementary Figure 1	Pairings between visual and audio stimuli in Exp. 1 and 2
Supplementary Figure 2	Testing trials and the difference between congruent and incongruent pairs, in Exp. 1 and 2.
Supplementary Figure 3	Procedure for a single session in an experiment
Supplementary Figure 4	Pairings between visual and audio stimuli in Exp. 3 - 6
Supplementary Figure 5	Testing trials and the difference between congruent and incongruent pairs, in Exp. 3 - 6

Note: The waveforms presented in the supplementary figure are just an imaginary representation of the audio stimuli. Participants only viewed the visuals in the experiments and listened to the audio. They did not see a waveform of each audio stimulus.

Supplementary Figure 1

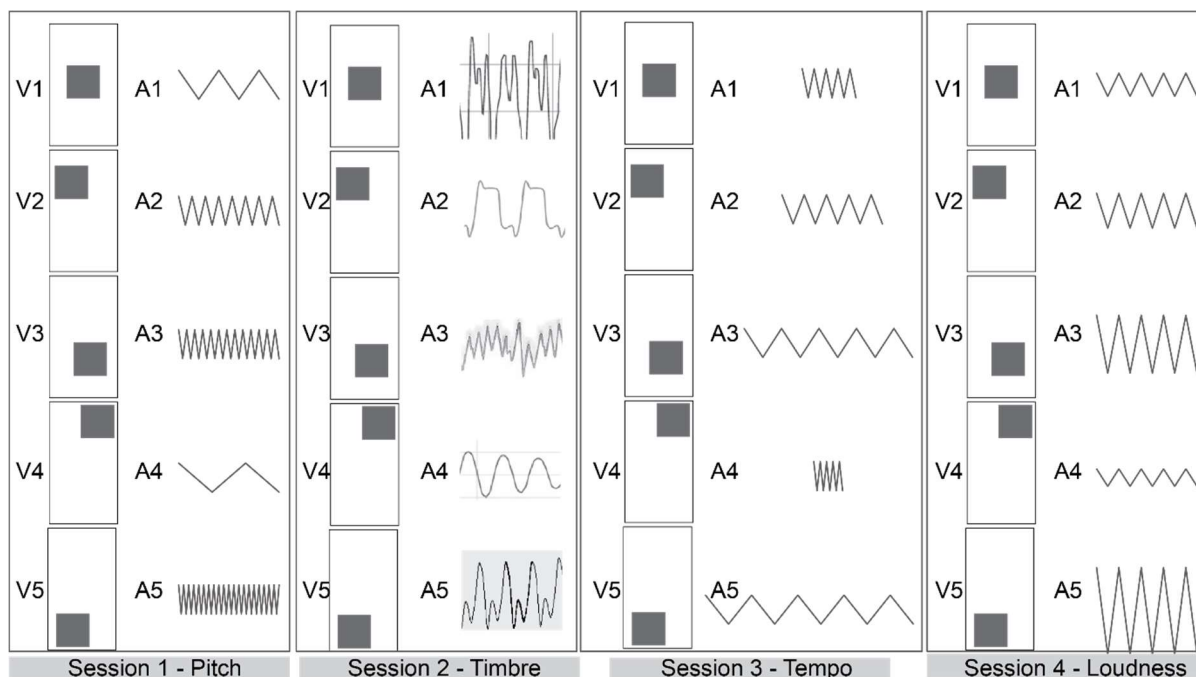


Figure 1 presents five audio-visual pairings in each of the four sessions of experiment 1 (Spatial position). In a session with pitch, five different pitches with the same timbre, tempo, and loudness were used. The same is the case with the loudness and tempo. In the case of timbre, each visual was paired with a different timbre. Pitches, loudness, tempo, or timbre used in the figure do not reflect the actual once. A similar pattern of audio-visual pairing was followed in experiment 2 (color).

Supplementary Figure 2

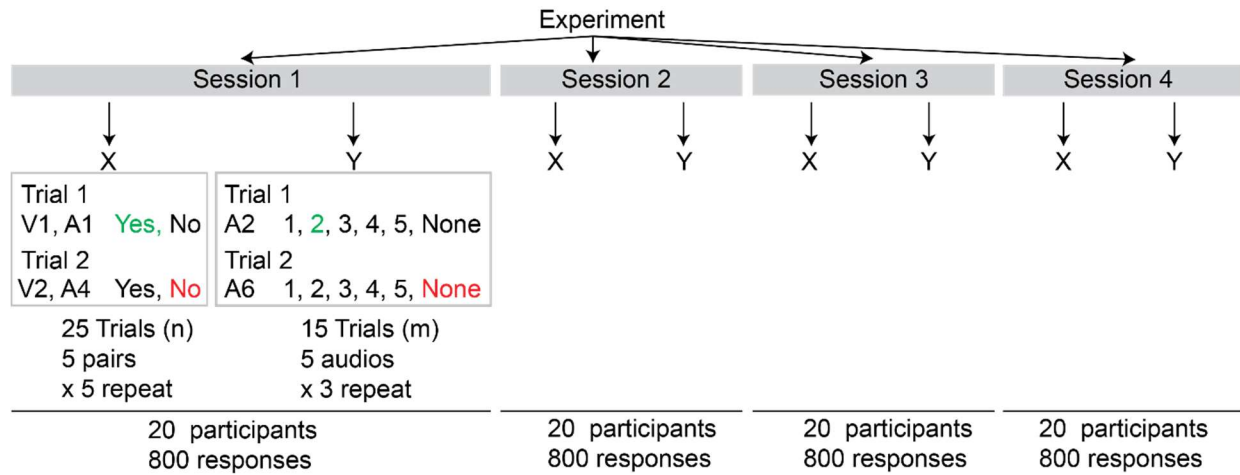


Figure 2 presents the organization applied in experiments 1 and 2 and the difference between congruent and incongruent pairs. The trials with the answer no or none present incongruence. We followed a similar pattern in the trials of all sessions of the experiments. In the trials, we repeated audio-visual pairs (see Supplementary Figure 1) randomly. In stage X, a single visual, in its multiple repetitions, could be presented with a matched or mismatched audio.

Supplementary Figure 3

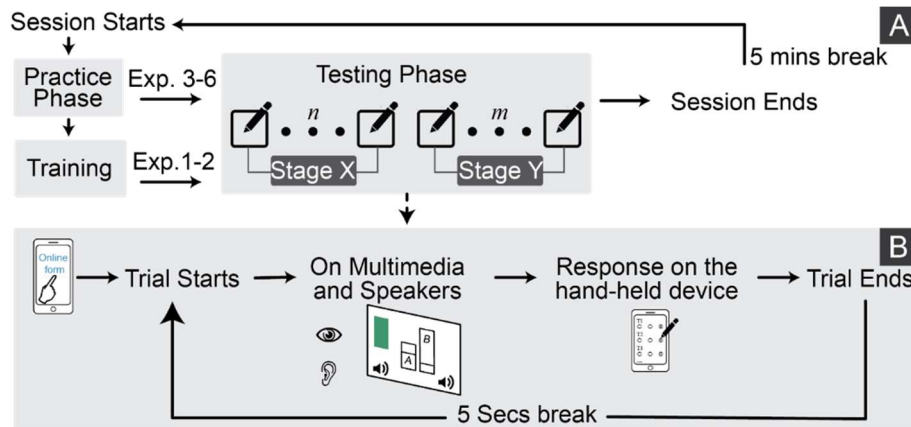


Figure 3 presents the procedure for a single session in an experiment (A) and the trials in a session (B). In the testing phase, n and m represent the numbers of testing trials in stages X and Y (see Supplementary Figures 2 and 5), respectively (A). All participants were sitting in the same room. They viewed visuals on the projector screen and listened to the audio on the speaker. In each trial, participants, after viewing/listening to the stimulus, had responded on the online form, which they had opened on their hand-held devices at the start of the testing phase (B).

Supplementary Figure 4

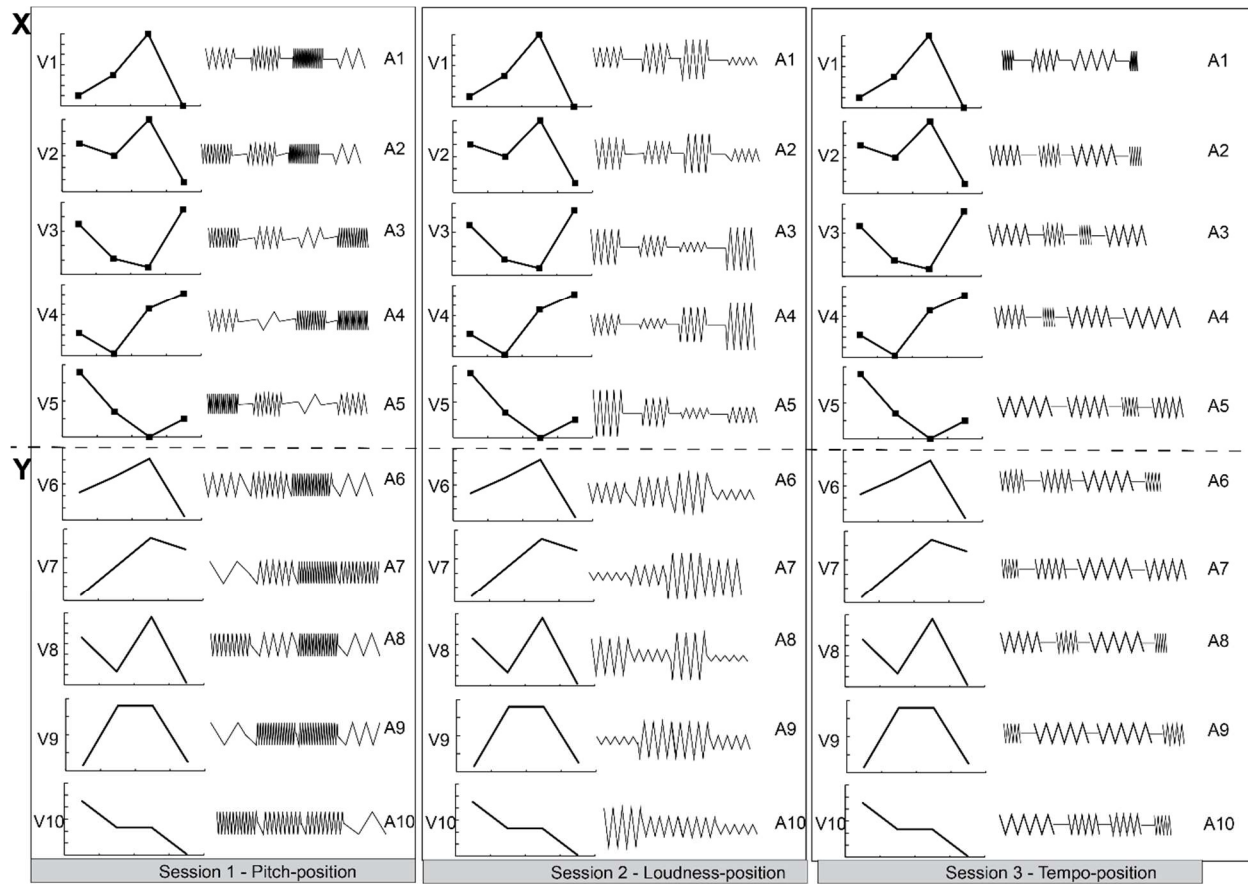


Figure 4 presents audio-visual pairing in the three sessions of experiment 3. These are samples, not the actual representations used in the experiment. Each line chart of experiment 3 comprised 20 data points. In stage X, the audio representations for pitch and loudness were in the form of sound + silence, and Y were continuous streams, i.e., no silence. Whereas, in either stage, the tempo stimuli were in the form of sound + silence. We used the same dataset to generate the audio and visual stimuli in a pair. We followed a similar approach to generate all the audio-visual pairs in experiments 4 to 6. However, all audio-stimuli of the last three experiments had sound + silence format.

Supplementary Figure 5

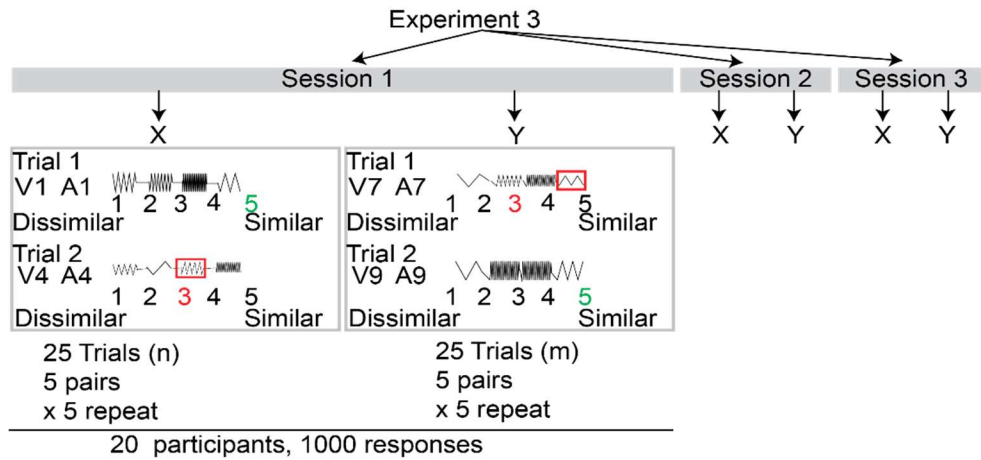


Figure 5 presents the organization applied in experiment 3 and the difference between the congruent and incongruent stimuli. The trials with similarity ratings of less than 5 show incongruence. The audio and visual stimuli presented in a trial were from the same pair (see Supplementary figure 4). However, either an exactly matched or mismatched audio was presented. Participants rated the perceived similarity between the audio and visual representation. A similar pattern for the trial generation was followed in all other sessions of experiment 3 and experiments 4 to 6.