Pitch–brightness crossmodal correspondence modulates perceptual selection of the Necker cube

Yeha Cha, Minseong Jin, Chai-Youn Kim
School of Psychology, Korea University, Seoul, Korea

Keywords: crossmodal correspondence, bistable perception, Necker cube

Crossmodal correspondences (CMCs) refer to associated stimulus pairings across different sensory modalities, such as bright visual stimuli paired with high-pitched sounds, and dark visual stimuli paired with low-pitched sounds (Melara, 1989). A previous study demonstrated that CMCs can bias the initial perceptual selection of a bistable stimulus based on figure-ground reversals: the Rubin's face/vase (Zeljko et al., 2021). However, it remains unclear whether such effects of CMC can be generalized to other types of bistable perception. This study used the Necker Cube, which is a bistable stimulus based on perspective reversals.

A Necker cube was shown with the edges of one face in either bright (43.76 cd/m²) or dark gray (23.53 cd/m²), and the remaining edges in the opposite luminance, against the mid-gray background (34.12 cd/m²). During the task, participants (N = 13) indicated which face appeared to be in front using a 2-AFC keypress. Each keypress elicited one of the three auditory stimuli: a high-pitched sound (1800 Hz), a low-pitched sound (600 Hz), or no sound. The tones lasted for 300 ms. Participants were instructed to release the key when their perceptual state changed, and the duration of each perceptual dominance was measured. Dominance durations were normalized by dividing each sound condition's durations by the mean 'no sound' duration for the corresponding brightness condition.

The data were first analyzed using a 2 (pitch: high, low) × 2 (brightness: bright, dark) factorial design. Based on the crossmodal congruency, the results were then examined by comparing the congruent (high-bright, low-dark) and the incongruent (high-dark, low-bright) conditions.

While the interaction between pitch and brightness was not statistically significant ($p = .108$), normalized dominance durations in the congruent condition were significantly longer than those in the incongruent condition ($p = .031$). This effect was stronger for dark than for bright edges, and stronger for low- than for high-pitched sound. These findings suggest that crossmodally congruent stimuli not only enhance perceptual salience but also contribute to the perceptual selection. Such results may build on findings about the effects of eye fixation in Necker cube perception, where the face nearest fixation is typically seen as front (Ellis, 1978); CMCs may help sustain dominance by increasing the duration of attentional or ocular fixation on a given interpretation.

Supported by NRF2023R1A2C2007289