Visual and auditory stimulus features, and their crossmodal correspondence, affects perceptual selection in the bounce/stream illusion

Jeon, S., & Kim, C-Y.

Various factors influence perceptual selection in bistable perception. For example, two identical disks that approach each other, overlap, then move apart, can be perceived as either streaming past or bouncing off each other (Metzger, 1934). Dubbed the bounce/stream illusion, it can be biased towards bouncing perception by introducing a sound near the point of visual coincidence (Sekuler et al., 1997). There is also evidence that an auditory tone can modulate perception of a static bistable image to favor the visual interpretation crossmodally corresponding to the tone (Zeljko et al., 2021). This study investigates the impact of visual (light/dark) and auditory (high/low) stimulus features, along with their crossmodal correspondence (light-high/dark-low; Melara, 1989), on perception of the bounce/stream illusion. Black or white disks were paired with either a high-pitched (1800Hz) or low-pitched (600Hz) tone. Stimulus onset asynchronies (SOAs) were set between -300ms to +300ms, where negative values indicate that the tone precedes the visual overlap of the disks. No-sound trials were also included. Binary responses of bounce/stream perception were collected from fifteen participants, and percent bounce responses for each lightness/pitch combinations were calculated for analysis. Consistent with prior studies, there was a main effect of SOA irrespective of auditory and visual stimulus features ($p<.001$). Of more relevance to our purpose, lightness modulated perception; participants exhibited more bounce responses when the disks were black than when they were white ($p=0.008$). Moreover, perception was further biased when a tone crossmodally corresponding to the disks' lightness was presented (i.e., low–black or high-white) than when an unrelated tone was presented (i.e. high-black or low-white. $p=.008$ and $p=.001$, respectively). These findings suggest that beyond the temporal relationship between audiovisual information, the individual features of auditory and visual stimuli, and the crossmodal correspondence between them, shape perceptual selection of an ambiguous motion stimulus.