

# Infant sensitivity to audio-visual discrepancy: A failure to replicate

SANDRA M. CONDRY, MAURICE HALTOM, JR., and ULRIC NEISSER  
*Cornell University, Ithaca, New York 14853*

It has been suggested that very young infants perceive in a common auditory and visual space. Aronson and Rosenbloom (1971) attempted to demonstrate this commonality by showing that infants become distressed by discrepancies between the visually and aurally specified locations of a speaker. However, this finding has not proved easy to replicate, and the present study also failed to confirm it. There are reasons to believe that the method of Aronson and Rosenbloom does not provide a strong test of their hypothesis.

The results of Aronson and Rosenbloom (1971) seemed to show that even very young infants (30-55 days old) are aware of the relationship between aurally specified and visually specified locations. Their subjects behaved as if they knew that the place where one sees a person (her visually specified location) should coincide with the direction from which her voice is heard as coming. When a speaker's voice was artificially displaced 90 deg to the right or left, all of the experimental infants exhibited distress. In particular, there was a significantly greater incidence of tongue protrusions during the auditory displacement than during control periods. This suggested the existence of an innate (or at least quickly learned) spatial coordination among sense modalities: a common auditory-visual space. However, McGurk and Lewis (1974) were unable to replicate this result. Their subjects exhibited no distress when a speaker's voice was displaced. The study reported here was performed to explore this issue further.

## METHOD

The eight subjects of the present experiment (two girls and six boys) ranged in age from 39 to 58 days. Their mothers served as speakers. Mothers and infants were not separated by a Plexiglas window, as in Aronson and Rosenbloom's (1971) study; they were in the same 10 x 10 ft curtained cubicle, 24 in. apart. The subject was placed in a semi-reclining infant seat. To his right and left were two loudspeakers, 35 in. apart. The mother spoke into a tiny sensitive microphone, held 1 in. from her mouth by a headset. She spoke relatively softly; the amplification assured that the sound heard in the infant's position was localized entirely on the basis of the loudspeakers (as judged by adult

pilot subjects). The mothers stood directly behind a waist-high periscope device attached to a television camera, which thus obtained a full-face view of the infant. The entire experimental session was videotaped.

Each mother talked to her infant for 4 min. During the first and third minutes both loudspeakers were set at equal volume, so that her voice appeared to be coming from a normal central location. During the second and fourth minutes one speaker was turned off (order was counterbalanced across infants), so the sound came either from the left or the right rather than from the mother directly. This procedure differed from that of Aronson and Rosenbloom; they presented the mother's voice in its normal location for either 2 or 5 min and then displaced it only once for a single minute.

## RESULTS

Two trained observers, blind to the manipulations, rated the videotapes on three measures: whether the head was to the left, center, or right (scored -1, 0, +1, respectively), whether or not the tongue was protruding (1 or 0); whether the infant's general emotional state was positive, neutral, or negative (+1, 0, or -1). Judgments were made every 5 sec. Agreement between the ratings of the two judges was 76% for head position, 86% for tongue protrusions, and 95% for emotional state. The analyses reported below are based on the algebraic means of the two judgments for each infant over each 1-min episode.

Analyses of variance with repeated measures (Winer, 1971, p. 268) were used to compare the four conditions: left speaker on, right speaker on, and two separate 1-min

Table 1  
Mean Ratings of Head Position, Tongue Protrusions, and Emotional State per Baby per 5-sec Interval for Each Loudspeaker Setting

Condition	Head Position (-1 < HP < +1)	Tongue Protrusions (0 < TP < +1)	Emotional State (-1 < ES < +1)
Left Speaker	.16	.20	-.06
Right Speaker	.34	.26	-.09
Both Speakers (first minute)	.25	.26	-.06
Both Speakers (third minute)	.20	.24	-.13

episodes with both in balance. A separate analysis was conducted for each measure. None was significant. For head position,  $F(3,21) = .57$ ; for tongue protrusions,  $F(3,21) = .39$ ; for emotional state,  $F(3,21) = .20$ . The most tongue protrusions occurred during the first episode (with sound central) and the fourth (with sound on one side).

### DISCUSSION

This study, like that of McGurk and Lewis (1974), failed to replicate the findings of Aronson and Rosenbloom (1971). While several of our infants may have noticed the new location of their mothers' voices (the nonsignificant trend in our data was toward appropriate head positions), none was distressed by the discrepancy between its auditory and its visual location.

Although the method developed by Aronson and Rosenbloom does not seem to provide reliable evidence for the notion of a common auditory and visual space, the unreliability of the method does not constitute evidence against that notion either. The method makes two assumptions: that infants combine auditory and visual information, and that they are distressed by a discrepancy between them. Failures to replicate their results may only mean that the second of these assumptions is wrong. The evidence of other studies suggests that the first may nevertheless be correct (Mendelson & Heath, 1976; Spelke, 1976; Wertheimer, 1961). The consistent results obtained in these studies, which show that infants will seek visual information about events they have heard, contrast markedly with the

difficulty of replicating Aronson and Rosenbloom's work. Future research on infants' intermodal perception should probably focus on exploratory behavior rather than on surprise or distress reactions. The anticipations that guide perceptual activity (Neisser, 1976) may be revealed more clearly in the course of normal functioning than when they are unexpectedly violated.

### REFERENCES

- ARONSON, E., & ROSENBLOOM, S. Space perception in early infancy: Perception within a common auditory-visual space. *Science*, 1971, **172**, 1161-1163.
- MCGURK, H., & LEWIS, M. M. Space perception in early infancy: Perception within a common auditory-visual space? *Science*, 1974, **186**, 649-650.
- MENDELSON, M. M., & HAITH, M. J. The relation between audition and vision in the human newborn. *Monographs of the Society for Research in Child Development*, 1976, No. 167.
- NEISSER, U. *Cognition and reality*. San Francisco: W. H. Freeman, 1976.
- SPELKE, E. Infants' intermodal perception of events. *Cognitive Psychology*, 1976, **8**, 553-560.
- WERTHEIMER, M. Psychomotor co-ordination of auditory-visual space at birth. *Science*, 1961, **134**, 1962.
- WINER, B. J. *Statistical principles in experimental design*. New York: McGraw-Hill, 1971.

(Received for publication February 14, 1977.)