Space Perception in Early Infancy: Perception within a Common Auditory-Visual Space?

Abstract. *Human infants aged 1 to 7 months were exposed to modifications of the normal spatial relationship between their mothers' face and voice. There was no evidence that such modifications were experienced by the infants as violations of a preexisting expectancy for face and voice to occupy the same spatial location.*

In a study of the origins of audiovisual integration, Aronson and Rosenbloom reported that infants as young as 30 days showed marked distress on observing their mothers speak to them at the same time as the mother's voice was spatially dislocated (via a loudspeaker) from her mouth (*1*). It was concluded that such dislocation represented a violation of the unity of the infant's perceptual world, a world in which speaker and voice are expected to share the same spatial location. Such results and conclusions, if reliable and valid, would have had profound implications for our understanding of early perceptual and cognitive development. Unfortunately, however, the study had a number of serious procedural and methodological deficiencies. For example, the authors failed to control for the sequence of presentations of normal and displaced face-voice conditions; the displaced condition always succeeded the normal condition. It is a matter of common observation that infants are more likely to become fretful the longer they remain in an experimental situation. Second, they discarded from the study those infants who showed distress during the beginning stage of the experiment, thus maximizing the probability that any change in distress behavior would be in the direction predicted by their hypothesis. Third, for their "normal" condition, Aronson and Rosenbloom assumed that an infant located midway between two stereo speakers would experience the mother's voice as coming from the front—that is, from her mouth—when the speakers were in balance (and from right or left when one or the other speaker dominated): such an assumption is questionable and its validity has never been assessed. Finally, the principal dependent variable reported by Aronson and Rosenbloom—the frequency of tongue protrusions—has never been validated as an index of infant distress.

As part of a larger, more extended study of audiovisual integration in early infancy, a replication of Aronson and Rosenbloom's experiment was undertaken, one which incorporated appropriate modifications to control for the deficiencies listed above. It was also decided to study the phenomenon developmentally. Infants, both male and female, were seen at three age levels: 1 month (25 to 31 days), 4 months (110 to 126 days), and 7 months (182 to 217 days). There were 11 subjects at 1 month and 12 subjects at each of the other two ages. Each infant was tested individually and was seated in a semiupright infant seat. For younger subjects, a broad body belt afforded support for the torso while the head and neck were supported by a U-shaped foam cushion. Limbs and arms could move freely.

The experiment was conducted in a small room, 3.5 by 4.6 m, which was partitioned into two sections at the center. Each section had an independent door on one side and a one-way viewing mirror on the other. The central panel of the divider was made of transparent plexiglass. The infant seat was located on the floor in front of the plexiglass. Two loudspeaker cabinets were located on the floor 101 cm from either side of the infant seat and a third loudspeaker was positioned against the plexiglass panel, immediately in front of the seat. A hand microphone, connected to the speakers, was located in the other section of the room. In a test of the apparatus, each blindfolded adult who was seated on the floor at the position occupied by the infant seat reported that the voice of an experimenter addressing the microphone in the other section of the room came from straight ahead, left, or right, depending on which speaker was activated at a particular time and regardless of the location of the experimenter within the other section.

The mother sat on a low chair, facing the infant through the panel, and spoke to the infant through the hand microphone. The distance between mother and infant was approximately 125 cm from face to face. The mother was instructed to speak continuously to the infant in a normal tone for a period of 2 minutes: if she thought she was in danger of having nothing to say she was to recite nursery rhymes or describe the furnishings of the room. Mothers seemed capable of complying with such instructions without great difficulty.

From an observation room one experimenter controlled whichever of the three loudspeakers was activated at a particular time. Other experimenters independently observed and recorded the infant's behavior from a television monitor. The camera was mounted immediately behind and above the mother and was focused on the infant's head and shoulders.

Each subject experienced four contiguous experimental episodes, each of 30 seconds duration. During episodes 1 and 4, the mother's voice was relayed over the center speaker. On episodes 2 and 3, her voice was relayed over right and left speakers; here, half the subjects in each group experienced a right-left sequence and the remainder experienced the opposite. Five indices of the infant's affective response to experimental treatments were recorded during each episode; these were frequency of smiling to the mother, frequency of vocalization, frequency of frowning, frequency of fretting or crying, and, in view of Aronson and Rosenbloom's report, frequency of tongue protrusions. In addition, three measures of head turning and visual regard were obtained; these were frequency of looking forward to mother, frequency of turning and looking to left or right, and frequency of looking down from mother. With trained observers these various indices yielded reliability coefficients of between .80 and .90.

The data for each response index were analyzed separately. There were no indices for which all subjects manifested the relevant behavior, and in a number of instances only a minority of subjects exhibited the relevant behavior. Accordingly, for each variable, only the data for those subjects who had exhibited the relevant behavior at least once during the entire experimental session were included for analysis. This was an admittedly nonconservative procedure, one which would increase the probability of our obtaining significant results. On the other hand it was a procedure that enabled us to detect such discriminations as did occur even when they were restricted to a minority of subjects. The procedure, therefore, favored the Aronson and Rosenbloom hypothesis. At each age level, Friedman analyses of variance by ranks were applied to the data, with experimental episodes (that is, spatial location of the mother's voice) being used as the independent variable. All the Friedman analyses reported herein have 3 degrees of freedom.

The data for the infants' turning the head and looking to left and right were as follows. The 11 infants aged 1 month, 10 of the 4-month-olds, and 11 of the 7-month-olds showed some evidence of this behavior. At all ages most turning and looking occurred during episodes 2 and 3. However, only at 4 and 7 months was the effect significant ($\chi_r^2 = 8.04$, $P < .05$, and $\chi_r^2 = 14.11$, $P < .01$, respectively). At 1 month, only 57 percent of head turns during episodes 2 and 3 were in the direction of the active loudspeaker; the corresponding proportions at 4 and 7 months were 84 and 73 percent, respectively. The difference between proportions is significant ($\chi_2^2 = 7.01$, $P < .05$). These data indicated that our changing the location of the mother's voice from the frontal to the lateral plane significantly influenced the frequency of head turning, at least at 4 and 7 months, and that at these ages, also, sound localization was relatively efficient. This is not to say that youngest subjects did not discriminate phenomenally between the different experimental conditions, but only that any such discrimination was not detected by this particular response index.

The infant's looking toward the mother was in evidence at all ages during all four episodes. Neither frequency of looking toward the mother nor frequency of looking down from the mother was significantly influenced by the direction from which her voice originated. Aronson and Rosenbloom, however, reported that infant distress associated with face-voice separation was such that subjects could not be induced to look again at the mother, even when the straight ahead condition was resumed.

Three of the 1-month-olds, two of the 4-month-olds, and four of the 7-month-olds cried or fretted at least once during the experiment. Most instances of such behavior were brief and transient and there was only one infant—a 1-month-old—who could not be consoled during the experimental session by verbal comforting from the mother. At no age level did differences between episodes approach statistical significance. Thus, nonsignificant results were obtained for the most direct measure of distress employed in this study.

At all age levels the frequency of frowning was unrelated to the spatial location of the mother's voice during the different experimental episodes.

At 1 month, seven subjects manifested overt protrusion of the tongue at least once during the experimental procedure; at 4 months, three of the subjects, and at 7 months, none of the subjects protruded the tongue. Neither at 1 month nor at 4 months did the number of tongue protrusions discriminate between episodes.

Vocalizations, other than those associated with fretting or crying, were recorded for three of the 1-month-olds, eight of the 4-month-olds, and seven of the 7-month-olds. At all ages, differences between episodes were small and were not statistically significant.

Smiling was recorded on the part of four of the 1-month-olds, ten of the 4-month-olds, and nine of the 7-month-olds. At 1 and 7 months the frequency of smiling was relatively even across episodes but at 4 months there was more smiling during the first episode than during the other three episodes ($\chi_r^2 = 11.61$, $P < .01$).

These data, therefore, afford no support for the hypothesis that the very young human infant lives in a perceptually unified audiovisual world, distortion of which is the occasion for distress. The incidence of distress observed here was no greater than that normally witnessed in experimental studies involving very young subjects. Moreover, such upset as did occur was unrelated to the spatial relationships between the mother's location and the direction from which her voice emanated during the different experimental episodes. It should be noted also that the incidence and frequency of smiling were higher than those of fretting or crying. The frequency of tongue protrusions bore no relationship to any of the other indices or affectivity nor was it in any way related to experimental conditions.

The nature and development of audiovisual coordination during early infancy remains an open question. The assumption that such coordination initially occurs within a unified audiovisual space is unsupported by our findings. Similarly, there is no evidence from our study that modifications of the normal spatial relationship between face and voice are experienced by the young infant as violations of a preexisting expectancy for face and voice to occupy the same spatial location.

HARRY MCGURK
Department of Psychology, University of Surrey, England
MICHAEL LEWIS
Educational Testing Service, Princeton, New Jersey 08540

References and Notes

1. E. Aronson and S. Rosenbloom, *Science* **172**, 1161 (1971).
2. This research was supported by grant GBZ8105 to M.L. from the National Science Foundation. Preparation of this report was supported by a grant from the United Kingdom Social Science Research Council to H.McG.

21 December 1973; revised 26 April 1974

The Asteroid Belt: Doubts about the Particle Concentration Measured with the Asteroid/Meteoroid Detector on Pioneer 10

Results obtained with the Sisyphus asteroid/meteoroid detector on Pioneer 10 have been presented by Soberman and co-workers (*1*). Auer and Northrop (*2*) criticized these results and concluded that no significant fraction of the events reported as real meteoroid events between 1.0 and 3.3 A.U. is, in fact, due to cosmic meteoroids.

Recently, Soberman *et al.* (*3*) reported the particle concentration in the asteroid belt (that is, from 2.0 to 3.5 A.U.) which is shown in Fig. 1. As will be pointed out, the brightness of the interplanetary light (also known as zodiacal light and gegenschein) measured at high elongations from both the earth and the Pioneer 10 spacecraft is incompatible with such a high concentration of matter. Since both kinds of data depend on optical detection of reflected sunlight from particles, several uncertainty factors, such as particle albedo or mass density, cancel out in the comparison.

The asteroidal debris cannot reflect more light than is observed. On that basis, and attributing all interplanetary light observed from the earth to particles in the asteroid belt, Kessler (*4*) derived an upper limit for the concentration of particles of any single size, a_1, which is

$$N_{max} = 2.7 \times 10^{-19} a_1^{-2} \quad (1)$$

where N_{max} is number per cubic meter and a_1 is in meters, for an albedo $p = 0.2$ and solar distances 2.0 A.U. $\leq R \leq 3.5$ A.U. This expression is not a size distribution such as given by Sober-