Published as a separate and in *The Journal of Psychology*, 1972, 82, 287-298.

PLAYING PEEKABOO WITH A FOUR-MONTH-OLD: A STUDY OF THE ROLE OF SPEECH AND NONSPEECH SOUNDS IN THE FORMATION OF A VISUAL SCHEMA* [1]

Center for Cognitive Studies, Harvard University

PATRICIA M. GREENFIELD

A. INTRODUCTION

The role of auditory signals, in general, and speech, in particular, in structuring the visual attention of a four-month-old baby was investigated by systematic study of my own baby in his natural surroundings. The question was whether sound in general and speech in particular would help the baby to segment continuous visual information by calling attention to critical points in a visual sequence. The visual sequences used in the study involved the disappearance and reappearance of objects or mother—the familiar peek-aboo game played with inanimate things, as well as people.

A secondary purpose of the study was to ascertain whether patterns of response would differ according to whether the disappearing objects were human or inanimate. Trevarthen and Richards (10) have found that, from the beginning, infants respond very differently to people than they do to inanimate objects.

The peekaboo game would not be possible if it exceeded the visual capacities of the infants. Haynes, White, and Held (6) have shown, however, that by four months of age, the point at which this study began, infants adjust their visual accommodation quite precisely to the distance of the target. Nor would peekaboo be possible if an object ceased to exist for an infant when it disappeared. But Bower (2) has shown that when things disappear slowly and perspectively, they continue to exist for a seven-week-old baby for a duration of five seconds. By 12 weeks of age, babies have existence constancy for objects disappearing at rates even faster than those used in a peekaboo game. In peekaboo, disappearance is much shorter than five seconds. Thus, at 12 weeks of age, objects are sufficiently permanent for the normal peekaboo game.

* Received in the Editorial Office on August 15, 1972, and published immediately at Provincetown, Massachusetts.

[1] Preparation of the manuscript was supported by Grant No. MH-12623 from the National Institute of Mental Health and Grant No. HD-03049 from the National Institute of Child Health and Human Development to Harvard University, Center for Cognitive Studies.

Charlesworth (4) has very successfully used this game to study cognitive processes in babies from five months up. He showed that even the youngest babies develop clear-cut visual expectancies when objects reappear in a constant locale. But what does it mean to say that expectancies develop, if in a sense the expectancies are there from the outset in the form of existence concepts? The spatio-temporal characteristics of a particular game must be what the infant is learning from experience. In this present study, a hiding-reappearance game was used to determine what interactional conditions facilitate or impede the development of this type of learning.

It is known that newborns are selectively attentive to the frequency range of human vocalization [Eisenberg (5)]. Given its privileged position as a stimulus, can the human voice serve to make spatially coordinated information in another modality—vision—more distinctive for the infant?

Newborns possess a capacity for sound localization as well [Bronshtein and Petrova (3), Wertheimer (11), Semb and Lipsitt (9)]. Although the actualization of this capacity is initially constrained by information-processing limitation [Aronson and Tronick (1)], by the age of three months infants will orient to all visible sound sources [Piaget (8)]. When a four-month-old baby then orients to a sound emanating from a visually informative locus, can this experience help shape a pattern of visual attention? The question of this study is whether a sequence of such experience facilitates the growth of a visual expectancy.

These problems were investigated by comparing the baby's response to the reappearance phase of the hiding game under varying auditory conditions. In games where the disappearing object was inanimate, disappearance and reappearance were accomplished either in silence, accompanied by the object's own sound, or accompanied by mother's speech. When mother was the disappearing object, the same three conditions were used except that a nonspeech sound made by the mother replaced the sound of the inanimate object. Smiling when the object reappeared was used as an index of a confirmed expectancy. According to evidence from Kagan *et al.* (7), four months is an age when the smile functions to indicate assimilation of a visual stimulus to an emergent schema.

B. Procedure

1. *Study 1*

The first study involved the typical sort of peekaboo game, in which I disappeared and reappeared. The baby's response to my reappearance was compared under two conditions: in one condition, the word "peekaboo," said with

the bright intonation normal to the game, accompanied reappearance; in the other condition, I reappeared silently. Under both conditions, I made a sound ("ooh-ooh") to attract his attention before hiding. A series of 10 game sessions, each consisting of seven to 14 trials, was played in a number of different locales around the house. In each game session, the two reappearance conditions—"peekaboo" and silent—were represented an approximately equal number of times; in the average session, each of the two conditions was represented by five or six trials. Fatigue and boredom determined at what point to terminate a game session. The study began at 17 weeks and lasted until 19 weeks of age. Some ancillary sessions were held as late as 23 weeks of age. More than one game session took place in each locale over a period of days or weeks, so that it was possible to trace the development of a peekaboo schema in each of four different locales.

One locale was Matthew's crib, beneath which I would disappear while he was in it. This was the only one in which he had played peekaboo before the onset of the study; the three other hiding places were a large bathroom hamper, a bed, and a sofa. Locales for hiding around the house were thus sampled to a reasonable degree.

Reappearance consisted of popping my face up (or sideways, in the case of the hamper) into Matthew's visual field. The amount of time in hiding was not controlled. In this way I felt that I would obtain a "representative sampling," in the Brunswikian sense, of amount of time hidden while maintaining natural conditions. My experience has indicated that artificial timing constraints in infant experiments often destroy the phenomenon one wishes to study. Nevertheless, when simulated game conditions were timed in the laboratory, temporal characteristics turned out to be remarkably consistent: time spent hidden averaged 1.3 seconds; the range was from 1.2 to 1.4 seconds. Rate of disappearing averaged 41 cm/second; the range was from 38 to 42 cm/seconds. Thus, "natural" timing in disappearance-reappearance games turns out to be uniform without any experimental constraints.

As a more general point about methodology, it seems that two qualities of infant behavior make it necessary to develop systematic yet natural observation conditions; in comparison with adults, infant behavior is (*a*) much more sensitive to situational context and inner state, and (*b*) much less detachable from specific goals. These considerations are most important when one purports to study how commerce with the social or physical environment affects the growth of some aspect of development.

Although it is necessary to avoid arbitrariness in designing experimental conditions, it is also necessary to avoid "experimenter effects" which confound

results. In the present case, an obvious source of such an effect could be the mother's smiling. Therefore, a conscious rule not to smile first but always to smile back when the baby smiled was initiated after the first game session. I thought that this rule would duplicate, in a systematic way, the naturally occurring contingency *vis-à-vis* smiling. Still, this procedure introduces the possibility that smiling is being operantly conditioned. This hypothesis will be evaluated later in the light of the actual results.

2. *Study 2*

This series of hiding games, begun after the main part of Study 1 when Matthew was 19 weeks old and continuing until 22 weeks of age, was designed to compare the effectiveness of speech and nonspeech auditory signals in structuring a response to the visual peekaboo game. A second complementary purpose was to see whether the pattern of response differed when an inanimate object, rather than a human being, disappeared and reappeared.

In this study, two parallel series of games were played with mother and with objects. In the object games, where three toys constituted the objects, experience with a given toy was the variable analogous to experience with a locale in the "mother" games. Thus, in both types of game the growth of a specific expectancy could be traced over time. During the object games, Matthew was always on the couch with me in front of it. The toy started in front of my face, disappeared downwards, and reappeared in front of my face. Thus, my presence was a constant stimulus in the object games. When the object game was simulated later in the laboratory its temporal conditions turned out to be extremely consistent and identical with those obtaining when the object was a person: disappearance occurred at an average rate of 41 cm/second; the range was 38 to 44 cm/seconds. The object stayed hidden an average of 1.3 seconds; time hidden ranged from 1.2 to 1.4 seconds.

Each game session consisted of nine peekaboo trials equally divided among three conditions. No condition was presented on two trials in a row. In addition to a speech and silent condition as in the first study, there was also a nonspeech sound condition. In game sessions where a toy was hidden, the speech signals were distinctive from those used when mother hid: "Hey, Matthew" to signal impending disappearance, and "Here it is" when the toy popped back. When mother hid under the nonspeech condition, the sound used to herald disappearance and reappearance was a sort of squeak, difficult to describe but often used with babies. This sound had in fact been a favorite for Matthew, but when tested before the study began, it no longer elicited a smile by itself. When a toy was hidden under the nonspeech sound condition,

the toy's own sound—squeak or chime—signalled disappearance or reappearance. Under the silent condition, both mother and toy disappeared and reappeared in silence. This was a departure from the first study in which an orienting signal was given even under the silent condition. In both studies, however, a trial was not begun until Matthew focused visually on whatever was about to disappear.

In this study, two new hiding places were used when mother hid—a table and a chair in their normal locations in the house. Two game sessions took place at the table, one at the chair. For inanimate disappearance, locale was constant—the couch used in the first study— but the object itself was variable. One of two rubber squeakers, a deer, was already familiar; a duck squeaker and chime rattle were not. Altogether six sessions of inanimate disappearance took place. One was excluded from the data analysis because no smiles were elicited under any condition, probably because Matthew was tired.

C. Results

1. *Study 1*

At first, speech constitutes a powerful cue in the peekaboo game with mother; her reappearance rarely elicits a smile without "peekaboo" marking the event. With increasing experience with the game in a given locale, however, the speech signal loses its privileged position, and the visual sequence alone suffices to elicit a consistent smiling response. Table 1 makes this pattern clear: in the first game in given locales, silent reappearance elicits a smile only six out of 17 times, or 35% of the time. When the game is repeated in the same locales at a later date, silent reappearance elicits a smile 50% of the time; when it is repeated a third time at a still later date, silent reappearance elicits a smile 100% of the time. The results from the "peekaboo" condition are in sharp contrast; smiling occurs 17 out of 19 trials or 89% of the time the very first time the game is played in a given locale. This response remains constant: smiling occurs 11 out of 12 trials (92%) for the second game in a given place, four out of four trials (100%) for the third game in the same place. Whereas the difference between the "peekaboo" and silent conditions is large and statistically significant ($p < .01$, one-tailed Fisher Test) during the first session in a given locale, all differences evaporate by the third series of trials: the visual sequence alone elicits smiling on every trial, exactly as it does in combination with the "peekaboo" signal.

In the results presented thus far, experience is somewhat confounded with maturation, especially by the third game session in a given locale. In order to

TABLE 1
RESPONSE TO MOTHER'S REAPPEARANCE DURING GAME SESSIONS VARYING IN FAMILIARITY OF LOCALE: STUDIES 1 AND 2

Reappearance marked by	Game sessions: First in given locale		Second in same locale		Third in same locale	
	Smile	No smile	Smile	No smile	Smile	No smile
Study 1[a]						
"Peekaboo"	17	2	11	1	4	0
Silence	6	11	6	6	3	0
Study 2[b]						
Human speech sounds ("peekaboo")	6	0	3	0		
Human nonspeech sounds	3	3	3	0		
Silence	0	6	3	0		

[a] Locales in Study 1 were as follows: First game session—hamper, sofa, bed; second session—hamper, sofa; third session—hamper.
[b] Locales in Study 2 were as follows: First game session—table, chair; second session—table.

see whether experience can play an independent part, let us turn to the data collected on the second day of the study in the one locale where peekaboo had been played before the systematic investigation began. We may then compare these results with those collected at the same time in two new locales. In this comparison, we hold maturation constant while varying amount of game experience in a specific locale.

The second day of the study we played peekaboo by Matthew's crib, the place where he had previous peekaboo experience. Reappearance elicited a smile on five out of five "peekaboo" trials and five out of five silent trials. The day before and the day immediately following, games were played in two new locales—hamper and couch. The results obtained on silent trials in new hiding places were uniformly different from those obtained in the familiar hiding place. When I reappeared silently, smiling occurred one in five times at the hamper, one in seven times at the sofa. In sharp contrast, silent reappearance at the crib had elicited a smile on every trial, that is, five out of five times. (This result at the crib locale was replicated five days later.) In other words, experience can promote the development of the visual disappearance-reapearance schema at any chronological point within the one-month range tested in this study.

It is interesting that at this point in development, expectancies seem extremely specific. This fact emerges when one notes that performance fails to

improve during a second game session when that session takes place in a *different* locale, whereas there is great improvement the second time when the game is played in the *same* locale. By the end of the month, either because of maturation or because of a learning-to-learn phenomenon, the initial game in a new locale (bed) elicits a much higher rate of smiling responses on silent trials (four out of six, or 67%) than was true at the beginning of the month (two out of 10, or 20%).

Thus, one can conclude, first, that the speech signal enables a consistent pattern of response to emerge earlier than is possible on the basis of visual cues alone. Second, experience that includes vocal articulation of the critical visual event promotes the future development of a consistent pattern of response in the absence of this auditory cue.

The inference is that the smile results from the confirmation of an emergent expectancy and that this consistent pattern of response therefore signifies the actualization in a specific instance of existence constancy. Some independent indication that the development of an expectancy constitutes the underlying process is furnished by more data collected one week after the last games reported thus far. In these two games, anticipatory smiles occurred on every trial at the alerting signal, so that recognitory smiles could no longer be scored. If prediction is the hallmark of an expectancy, then this anticipatory response constitutes extremely strong evidence for the reality of such a development.

The temporal patterning of the smile response would seem to offer the best evidence against reinforcement as an explanation of the results. Since I would smile back any time Matthew smiled, reinforcement in no way explains why, initially, he smiled only at the end of a trial (reappearance) and never at the beginning of a trial (before disappearance).

2. *Study 2*

The results of Study 2 replicate those of Study 1; Table 1 shows that the first time mother disappears and reappears in a particular place, reappearance elicits a smile when accompanied by "peekaboo." By the second game session in one of the same locales four days later, smiling occurs whether or not there is any sound to signal disappearance and reappearance. The nonspeech signal starts out exactly intermediate between speech and silence in eliciting a smiling response (three out of six trials, or 50% of the time). But experience either improves the effectiveness of the nonspeech signal or obviates the need for any auditory signal; during the second game in a given locale, smiling occurs on every trial in *every* condition.

The results of the parallel object-hiding games are displayed in Table 2

where three degrees of experience with game and object are distinguished. Degree of experience is a composite of two factors: (*a*) familiarity with an object before it was used in any hiding game, and (*b*) experience with an object in a game session. At all three levels of experience, speech-trials virtually always elicit a smile, silent trials never. Thus, Matthew never seems to learn what to expect from a hidden toy unless he is given some auditory cues. The visual schema seems harder to establish for things than for mother, and auditory cues become correspondingly more crucial in directing the pattern of visual attention.

As for nonspeech sounds, they gradually become more effective in structuring visual attention. In the first game with the unfamiliar toys (duck squeaker, chime rattle), the nonspeech sound condition elicited a smile on only on out of six trials (17%). In the games of intermediate familiarity (deer and duck squeakers), smiling occurred two out of six times (33%). In the games at the highest level of familiarity (second session familiar deer squeaker), smiling occurred in the object-sound condition two out of three times (67%). This trend toward the increasing effectiveness of nonspeech sounds is statistically significant at the .05 level when one considers the results both for "people" and for "thing" games (one-tailed Fisher Test).

In the object games, Matthew never smiled at my face alone (during the disappearance phase), but smiled only when face and object were visible together (reappearance). This pattern of response is inconsistent with a reinforcement explanation of the observed behavior, for "smiling back" as a reinforcer could occur at any point in an object game, not exclusively upon the object's reappearance.

Additional data would have been desirable, but Matthew refused to play

TABLE 2
RESPONSE TO OBJECT'S REAPPEARANCE DURING GAME SESSIONS VARYING IN FAMILIARITY OF OBJECT

Disappearance and reappearance marked by:	Degrees of Experience: Level 1 (Initial game session with unfamiliar object)		Level 2 (Initial game session with familiar object; second game with unfamiliar object)		Level 3 (Second game session with familiar object)	
	Smile	No smile	Smile	No smile	Smile	No smile
Human speech sounds	5	1	6	0	3	0
Object sound	1	5	2	4	2	1
Silence	0	6	0	6	0	3

any more. Since Charlesworth (4) has found persistence at a peekaboo game to be inversely related to the infant's ability to predict reappearance, this behavior constitutes additional evidence that reappearance had become predictable, that Matthew had indeed formed a visual schema.

In sum, the pattern of results is very striking and consistent for inanimate and animate objects. When Matthew starts to play a new game—that is, one involving a new locale or a new object—speech holds a very privileged position; it is the only condition that reliably elicits the recognitory smile. With experience, the sound of an inanimate object becomes almost as effective in producing this smile. Silent reappearance of an inanimate object never elicits a recognitory smile. As for games involving mother's disappearance and reappearance, nonspeech sounds are initially somewhat effective in eliciting the smile, although much less so than speech. After experience with the game in a given locale, Matthew shows consistent appreciation of my reappearance under all three conditions. At this point the visual event alone suffices to trigger an expectancy and to induce its confirmation.

D. Discussion

The greater difficulty of establishing a disappearance-reappearance schema for things than for a person who is a mother may reflect strong general expectations about what happens to a mother or to any person in the baby's experience when she hides. This would be in line with a general conclusion of the two studies: the more unpredictable a visual sequence, the greater the role of auditory cues in delineating its critical points. Because I was the only person to play peekaboo with Matthew, we cannot distinguish a specific "mother" effect from a general "person" effect. The inanimate objects, by contrast, represented a sampling of different noise-making toys; generalization about inanimate objects therefore seems warranted.

A potential source of weakness in the studies was the failure to assess whether "peekaboo" and "here it is" would elicit a smile by themselves at the outset of the investigation. Even if they had, however, this fact could not explain why the silent and nonspeech conditions became equally effective in eliciting a smile after experience with a given visual sequence. The effectiveness of the silent condition in eliciting a recognitory smile the first time peekaboo was studied in the familar crib locale shows that previous experience with a silent event is not a necessary requisite for the development of a purely visual schema, for there had been no silent hiding before the study began. While seeing that "peekaboo" trials are sufficient for the development of the hiding-reappearence expectancy, we do not know if the silent trials also made a con-

tribution to this learning. But if the pattern of attention is the crucial aspect of learning a visual expectancy, then it makes sense to think that experiences in which attention is directed to the critical loci at the right times would be much more valuable in organizing a temporal pattern than would experience where attention might well be randomly or irrelevantly distributed, as on the silent trials.

If Bower (2) is correct in claiming that infants of this age have existence constancy at the outset, what does it mean to say that a child "learns" a simple hiding game in the course of certain kinds of experience? It must be the spatiotemporal pattern of this particular game that is being learned. In relation to timing, it is probably significant that many of the earliest games babies play have a strong rhythmic component, for example, "pat-a-cake." It is often this rhythmic component which seems to contribute most to the game-like quality of the interchange; rhythm comes to be a rule-of-the-game that the baby learns first. As for spatial learning, our results show that peekaboo is initially defined as a game-in-a-particular-place. Later, we have seen, Matthew's expectancy transcends the limits of any specific locale. It could well be that when the expectancy also transcends the limits of a specific temporal rhythm, the whole game loses it appeal for the baby.

If these results are valid, they suggest an important way in which caretakers can control and develop the deployment of an infant's visual attention through the use of sound. Sound can contribute informational redundancy *vis-à-vis* the visual event by encoding it in a second modality. Moreover, sound, under normal conditions, has two design characteristics—spatial localization and temporal organization—which may be important in guiding visual attention. The spatial localization of an ordinary auditory stimulus plus the omnidirectional capacity of human sound receptors means that visual attention can be guided to the informative locus. Temporal organization, on the other hand, brings with it the possibility of imposing points of articulation upon the relatively continuous flow of the visual world. It appears that a mother's speech, and probably speech in general, hold a privileged position among auditory stimuli in fulfilling these functions. Since the role of speech in the enviroment is a key source of subcultural variation in cognitive development, and since skillful deployment of attention may well be the crucial cognitive accomplishment in the first year of life, these phenomena could constitute the ontogenetic beginnings of an important relationship between language and attention. Many more infants must be studied if such developmental phenomena are to be fully understood.

E. Summary

This study investigated the role of auditory signals in general, and speech in particular, in structuring the visual attention of a four-month-old baby.

The first study involved the typical sort of peekaboo game, in which mother disappeared and reappeared. The baby's response to reappearance was compared under two conditions: in one condition, the word "peekaboo," said with the bright intonation normal to the game, accompanied reappearance; in the other condition reappearance occurred in silence. The results showed, first, that the speech signal enables a consistent pattern of response to emerge earlier than is possible on the basis of visual cues alone. Second, experience that includes vocal articulation of the critical visual events promotes the future development of a consistent pattern of response in the absence of this auditory cue.

The second study was designed to compare the effectiveness of speech and nonspeech auditory signals in structuring a response to the visual peekaboo game. A second complementary purppse was to see whether the pattern of response differed when an inanimate object, rather than a human being, disappeared and reappeared. The pattern of results is very striking and consistent for inanimate and animate objects. In a novel hiding game—that is, one involving a new locale or a new object—speech holds a very privileged position; it is the only condition which reliably elicits the recognitory smile. With experience, the sound of an inanimate object becomes almost as effective in producing this smile. Silent reappearance of an inanimate object never elicits a recognitory smile. As for games involving mother's disappearance and reappearance, nonspeech sounds are initially somewhat effective in eliciting the smile, although much less so than speech. After experience with the game in a given locale, the baby shows consistent appreciation of mother's reappearance under all three conditions. At this point the visual event along suffices to trigger an expectancy and to induce its confirmation.

References

1. Aronson, E., & Tronick, E. Implications of infant research for developmental theory. In J. Eliot (Ed.), *Human Development and Cognitive Processes.* New York: Holt, Rinehart & Winston, 1971.
2. Bower, T. G. R. The development of object permanence: Some studies of existence constancy. *Percept. & Psychophys.*, 1967, 2, 411-418.
3. Bronshtein, A. I., & Petrova, E. P. The auditory analyzer in young infants. In Y. Brackbill & G. G. Thompson (Eds.), *Behavior in Infancy and Early Childhood.* New York: Free Press, 1967. Pp. 163-172.