Infant Imitation: The Sensory-Motor Agenda

Kenneth Kaye and Janet Marcus
University of Chicago

Six different actions were demonstrated to nine infants once a month between the ages of 6 and 12 months. Each action was presented many times, with each trial contingent upon the infant’s making eye contact with the experimenter. From videotapes, 21 categories of infant behavior were coded continuously. Each category could be considered a component act or feature of one of the modeled actions. All categories were coded during all parts of the session: baseline periods, trials of the tasks of which they were features, and trials of the other tasks. Infants did not confine their performance of features to the relevant trials as construed by the authors, but they did accelerate or introduce features for the first time during those trials. Analysis showed that the infants “worked up to” precise imitations by accommodating themselves to the features in a consistent order over both months and trials.

Traditionally, the question of how imitative capacities vary as a function of the infant’s age and of the nature of the task was addressed by trying a battery of tasks at each of several ages. A given action was demonstrated only a few times to the infants, who were scored as either imitating it or not. The ages when infants appeared consistently able to reproduce the different actions allowed investigators to rank order the tasks by difficulty and then to analyze logically what distinguished harder actions from easier ones. This was the method used informally by Guillaume (1926/1971) and by Piaget (1945/1951), who based much of their reasoning about sensory-motor development upon infants’ successes and failures at immediate imitation. The method has also been used more systematically, with respect to experimental design, by recent investigators (e.g., Abravanel, Levan-Goldschmidt, & Stevenson, 1976; Uzgiris, 1972; Maratos, Note 1).

An implicit assumption of that approach is that infants can either imitate an action right away or else they cannot imitate it at all, and that this is apparent within a few trials. Failing to elicit imitation, investigators have usually switched to another of the modeled acts in their test battery. Infants can often make significant progress, however, if they are given many trials with the same repeated model (Kaye, 1977; Kaye & Marcus, 1978; Valentine, 1930; Kaye, Note 2). One needs only to demonstrate the behavior repeatedly, without contingent response to the infant’s attempts, and allow the infant to control the onset of the presentations and the pauses between them. We previously argued (Kaye & Marcus, 1978) that although our method produced imitation of mouth movements (that the infants could not see themselves making) at an earlier age than had been observed by any other investigators, the results were actually consistent with the main tenets of Piaget’s theory: Approximations to the model were made feature by feature, and it was the assimilation of additional features over the successive trials (rather than a process of reinforcement on the one hand or representational processes on the other) that resulted in accommodation to the model.

In our original study (Kaye & Marcus, 1978), however, we presented our subjects with only one behavior to imitate: rhythmic mouth movements. The results we obtained might have occurred in response to any task sharing similar features of rhythmicity or of opening and closing. If, for example, the infants were imitating “open-and-close” and

This research was funded by the Spencer Foundation.
The authors are grateful to Mary McCauley, Susan Templeton, and the nine mothers for their prolonged involvement in the project.
Requests for reprints should be sent to Kenneth Kaye, 4810 South Dorchester Avenue, Chicago, Illinois 60615.
if their mouths happened to be readily energized, then their imitation of mouth opening and closing would not prove that they had any representation of the experimenter's mouth in terms of a particular part of their own bodies.

To test the generality of the original findings to other parts of the body, to other tasks, and to other ages, we designed a modification of the battery approach. As Uzgiris (1972), Maratos (Note 1), and others did, we selected several kinds of behavior to be modeled in an a priori order. We saw each infant once a month between the ages of 6 and 12 months and presented all the tasks to every infant. Each of the task-presentation periods could then be used as a control for the others, so that imitation of a particular feature of a task could be defined as a significant increase in the appearance of the feature in the infant's behavior during that task when compared with its rate of occurrence at all other times. The major difference between our design and previous studies by other investigators was that we presented each model for as many trials as the infant would allow. Our hypothesis was that we would replicate our earlier (Kaye & Marcus, 1978) findings across all tasks, with two additions: (a) More difficult tasks, such as touching the finger to the ear, would yield data similar to the data we had obtained on mouth opening and closing at 6 months, but they would do so at a later age. (b) The progress we would see over months, in each of the tasks, would be reflected in the progress over a succession of trials at any particular age.

Method

Subjects

Five male and four female infants—born to white, married, English-speaking women—were seen once a month from 6 months through their first birthdays. Three infants missed one visit each, leaving a total of 60 sessions. The sample was relatively homogeneous: All of the infants had been healthy and full-term at birth and were principally cared for by their mothers.

Tasks

Table 1 presents the six actions we selected for modeling and the component acts we coded from the videotapes. The tasks were selected from among those used in previous studies, with the goal of including as varied a set of actions as possible. We increased the differences among them by protruding the tongue only once per trial, shaking the toy twice, vocalizing three times, and so on. Three tasks were silent, three audible; three involved the mouth, three the hands; four required the infants to make movements they could not watch themselves making; one required action on an object. Random orders of presentation of the six tasks were counterbalanced across subjects and sessions. The mothers agreed to avoid these precise actions at home.

After a 10-min acclimatization, the videotaped session began with a free-interaction period between the baby and the female investigator. The infant sat in the mother's lap facing forward (as was true throughout the session, except that from 10 to 12 months some infants preferred to move to the floor for part of the time). During the initial baseline, which lasted approximately 2 min, the investigator simply tried to engage the infant's attention, responding appropriately to gestures and vocalizations but refraining from performing any of the six target actions. The session then proceeded with the six tasks, using the infant's eye contact with the investigator as the trigger for each trial (Kaye & Marcus, 1978). Four of the tasks began with their own initial 2-min baseline period to test cognitive assimilation (Piaget, 1945/1951, pp. 8–18), that is, to give the infant an opportunity to recall the same task from the previous session and to perform part or all of it before the adult did. Throughout these baselines, the investigator held the starting position for the respective tasks: For the ear task, she held up her finger about 20 cm to the side of her head; for the clapping task, she held her hands in front of the baby; for the toy task, she handed the toy to the baby; and for the mouth task, she stared at the baby with a still face. Because the latter pose was indistinguishable from the starting positions for tongue movement and vocalization, we skipped the baseline periods for those tasks to keep the session flowing smoothly.

Each task lasted as many trials as the infant would elicit before permanently turning away or crying, up to a maximum of 10 min. Typically, this meant 20–40 trials over 5 min, interrupted by one or two "breaks" when the infant was allowed to explore an irrelevant object or to turn around to look at or embrace the mother. We must stress that the experimenter was well trained (and carefully monitored) not to respond in any way contingent upon the infant's performance. There simply was no reinforcement for imitation.

Coding and Reliability

The categories listed in Table 1 were coded in real time by the second author and an assistant, who watched the videotape at normal speed and pressed keys on an electronic event recorder that associated a time, to 1.0 sec precision, with each coded event. Five passes through each videotape were required, one for the investigator's behavior and one for the infant's (a) mouth movements, (b) vocalizations, (c) reaching and touching, and (d) grosser arm movements including clapping and shaking the toy. All categories—with the exception of acts that
Table 1

*Tasks and Relevant Behavior*

Clapping—Model brought her hands together and apart in front of the infant four times within 5–6 sec:
- EHAND—touched experimenter’s hand
- SHAND—touched one of own hands with the other, not clapping
- PULL—pulls experimenter’s hands together, in clapping motion
- CLAP—brings palms together and apart
- 4CLAPS—a burst of exactly four CLAPS
- ARMBURST—rhythmic arm movements (Each discrete movement was recorded, but there had to be two or more.)

Shaking toy—Model shook a 5-cm cube containing a bell, twice within 1.5 sec, and handed it to the infant:
- GET—takes toy from experimenter
- SUCK—brings toy to lips
- SHAKE—shakes toy sufficiently to ring bell
- RELEASE—lets toy drop into experimenter’s hand
- 2SHAKES—a burst of exactly two SHAKEs

Mouthing—Model opened and closed mouth five times, like a goldfish:
- MOUTH—any mouth movement to a wider position, excluding smiles (Mouth had to close, at least partly, before another MOUTH would be coded.)
- OPEN-CLOSE—three fingers of either hand opened and closed within 2 sec*
- 5MOUTHS—a burst of exactly five MOUTHS

Tongue protrusion—Model protruded her tongue as far as she could, held it out for 2 sec, and drew it back into her mouth, once:
- TONGUE—extends tongue beyond lips*
- SMOUTH—touched own mouth
- EFACe—touched model anywhere on the face

Vocalization—Model said “gi, gi, gi” within 4 sec:
- VOC—any discrete vocalization except a cry or a laugh
- 3VOCS—a burst of three vocalizations

Touching ear—Model held index finger of left hand to the side of her head, pointing upward; when the infant made eye contact with her eyes, she stuck the finger in her ear, held it for 1 sec, and removed it:
- SFACE—touched self anywhere on the face except mouth or ear
- SEAR—touched own ear

* Unreliable category, not analyzed.

could only occur when the toy was present—were coded in all parts of the session.

Throughout the months of coding, randomly selected tapes were reinserted into the randomized list of tape numbers so that they would be coded a second time, usually without the coders being aware of when this was the case. For all but two categories, coefficients of reliability were calculated as percent agreement within a maximum time lag of 3 sec (the number of events “matched” within this criterion divided by the total number recorded). These coefficients averaged .83 for the infant’s behavior and nearly 1.0 for that of the investigator. The actual response lag between two codings of the same event averaged less than 1 sec.

Because mouth movements and vocalizations occurred more rapidly than did the other categories of behavior, a different kind of reliability coefficient was appropriate for MOUTH and VOC. For each of these categories, we correlated the numbers of events seen by each of the coders in each segment of tape (i.e., reliability of the counts for each different task and baseline). These coefficients averaged .87 for MOUTH and .93 for VOC.

Intracoder and intercoder reliability was established above .80 for all categories except two: the infants’ opening and closing of their hands and tongue protrusion. Because the video image included the whole baby as well as the investigator’s head and upper body, resolution was not sufficient to distinguish an infant’s tongue protrusion, for example, from a mouth opening where the tongue was visible inside the mouth. Therefore, the tongue protrusion task was not analyzed beyond the summary in Table 2, though that period of the session was still coded for the other behavior.

Because it was not always possible, when coding the baby, to cover the half of the monitor where the investigator appeared (so as not to be biased by knowledge of the task), we selected a subset of tapes on which this was strictly possible as a result of the performers remaining within separate zones. Reliabilities were nearly as high under these conditions as they had been in the
actual codings, and the same was true when we trained a new assistant to code selected infant behavior while she remained unaware of the actions that had been modeled in the study.

Results

Once we had intercalated the different modes of behavior in the computer and produced transcripts of various kinds, we took several different approaches to analyzing the sequential relation among events. For the sake of brevity we shall first summarize the findings, then report them specifically task by task.

The intention of our original design was to pick out the months, for each behavioral category, when the infants’ rates of producing that behavior were significantly higher during the trials of the task of which it was a component than during the rest of the session. Unfortunately, progress in the various tasks was shown by the nine infants in different months. Nor could any individual be said to progress steadily. Furthermore, the infants did not share our understanding about the division of the session into separate tasks: They did not stop producing features of one task just because we had gone on to another task. What they did do, however, was nearly always perform a feature for the first time during the appropriate trials. After this occurrence, a feature was often performed later in the session, perhaps with even greater frequency. It then also might occur in later sessions, even before the experimenter had modeled the task of which it was supposed to be a feature. In the case of categories like SHAND—simply touching one hand with the other, which occurred at a low but nonzero rate in the baseline periods—there was nearly always a session in which the rate at least doubled during and following the trials (in the case of SHAND, the clapping trials) before any session in which there was a substantial rate of SHAND prior to the trials. We were able to use this sharp increase as a criterion for the month of “first occurrence” of SHAND, MOUTH, VOC, SPACE, and SEAR (see Table 1 for definitions). For all other categories, the rate increase was from zero (i.e., an absolute criterion was used).

Our first set of findings (reported below, by task) shows that in every task there was a consistent order, across infants, in the first occurrence of the categories we had defined a priori as approximations of “features” of the tasks.

Our second set of findings (also reported by task) involved the ordering of those features within a series of trials. The question was whether infants tended to “work up to” a new feature over trials, beginning with the approximations they had already mastered in earlier sessions. To test this statistically, we reduced the computer transcripts to the infants’ behavior during trials, confining ourselves to just those sessions in which a new feature was produced for the first time, but listing “old” as well as new features the first time they were produced during the trials. Thus any act could appear no more than once in the ordered list for that session, but it could in principle appear either before or after the feature that was being produced for the first time. We found that “old” features, whenever they appeared in these sessions, overwhelmingly appeared before the new features, and that when two new features appeared in the same session, they appeared in the order corresponding to the typical order over months.

These analyses allowed us to disregard the disparities among infants with respect to the particular ages at which they chose to improve upon their imitations of the model. The median months in which they first achieved success are shown in Table 2. These successes were achieved after many trials within the session. It is interesting to observe, however, that infants did not necessarily maintain their highest level of success in later months.

In the results that follow, a “burst” of repetition is defined as any series, no two of which were separated by 2 sec or more. Thus, we can distinguish between CLAPS, for example, and exactly 4 CLAPS.

Clapping Task

Accommodation to the model’s clapping was reflected in the consistent order of performance, during clapping trials, of EHAND, SHAND, PUT, CLAP, and 4 CLAPS (see Table 1 for definitions). EHAND occurred before
Table 2
Number of Infants Imitating at Top Level

<table>
<thead>
<tr>
<th>Task</th>
<th>Criterion</th>
<th>Number of infants</th>
<th>Median month</th>
<th>Median trial*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Touching ear</td>
<td>SEAR</td>
<td>9</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>Vocalization</td>
<td>3VOCS</td>
<td>9</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>Shaking toy</td>
<td>2SHAKES-&amp;-RELEASE</td>
<td>8</td>
<td>9½</td>
<td>4</td>
</tr>
<tr>
<td>Mouthing</td>
<td>MOUTHBURST†</td>
<td>9</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>MOUTHS</td>
<td>4</td>
<td>8½</td>
<td>12</td>
</tr>
<tr>
<td>Clapping</td>
<td>CLAP</td>
<td>7</td>
<td>11</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>4CLAPS</td>
<td>3</td>
<td>12</td>
<td>8</td>
</tr>
<tr>
<td>Tongue</td>
<td>(obvious, prolonged protrusion:)</td>
<td>0</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

Note. n = 9.

* Trial number within the session in which criterion was achieved.
† Two or more, but not five, mouth movements within 2 sec of each other.

SHAND in seven of the nine infants; seven of the nine performed SHAND before PULL; nine performed PULL before CLAP, or never performed CLAP; and the three who achieved 4CLAPS all previously performed a simple CLAP and then a burst of two or three CLAPS.

The median ages of first occurrence were: EHAND, 6 months; SHAND, 7 months; PULL, 10 months; CLAP, 11 months; and 4CLAPS, 12 months.¹

In the first month in which an infant produced one of the a-priori-defined approximations to clapping, we looked at the order in which that feature appeared relative to all other clapping features produced in that session's clapping trials. As explained above, all possible ordered pairs were considered: For example, PULL appeared in 12 ordered pairs including sessions in which it was produced for the first time and sessions in which another feature was produced for the first time. In 10 of those 12 (p = .02, binomial sign test), the order of occurrence corresponded to the order reported in the preceding paragraph: after EHAND or SHAND, before CLAP or 4CLAPS. This consistent order was true of all five levels of approximation, up to 4CLAPS, for a total of 47 out of 61 ordered pairs (p < .001).

Toy-Shaking Task

The median age at which the infants first shook the toy sufficiently hard to ring the bell was 7 months; the median age was 7½ months when they shook it exactly twice within 2 sec and stopped; 9 months when they shook the toy and released it, as the investigator had done; and 9½ months when they shook it exactly twice and released it. Few infants offered the toy directly to the investigator. We considered them to have reached the top level if they let the toy drop into the investigator's hand after 2SHAKES.

All nine infants first performed a simple SHAKE earlier than they did SHAKE-&-RELEASE; eight did SHAKE-&-RELEASE earlier than 2SHAKES-&-RELEASE; seven produced 2SHAKES earlier than 2SHAKES-&-RELEASE. There was no significant order between 2SHAKES and SHAKE-&-RELEASE: After imitating the simple SHAKE, some infants went on to match the correct number, then imitated the release, then put it together, whereas others first added the RELEASE, then worked on 2SHAKES, then put it together. There was evidence that they built up to higher accommodations over the series of trials, insofar as the above orders were preserved in 37 of 53 ordered pairs of features within the sessions when features appeared for the first time (p = .003). Although this task was different from the others in that the rest of the session could not be used as a check against the spontaneous production of the acts we considered imitative (because the toy was not available then), the baseline period with the toy did serve that purpose. Closer approximations (beyond the simple SHAKE) always appeared during the trials.

¹ Median ages are for the purpose of comparing the different criteria only; they have no absolute significance because the subjects were not followed beyond 12 months until all of them had met the highest criterion.
before being produced in later sessions on the task baseline.

**Mouthing Task**

All of our 6-month-olds produced mouth movements, though usually not in bursts and not necessarily in excess of the movements they made in other parts of the session. The median age for mouthing bursts during mouthing trials was 7 months. Only four infants ever achieved precisely 5MOUTHS; the median age at which they did so was 8½ months. Bursts of two or more mouth opening and closings were almost entirely confined to the trials. All nine infants produced such bursts and continued to do so over many trials and months—bursts of 3, 4, 6, 7, and so forth—so it was inevitable that some of them would eventually hit upon the number 5. In those sessions in which a new level was attained, 19 of 20 sequences preserved the order MOUTHING—BURST OF BURST—5MOUTHS (p < .001).

**Vocalization Task**

The "gi-gi-gi" was never achieved with precision, though there were many trials in which a glottal consonant was combined with /j/. We were not able to distinguish the /g/ reliably, so we confined our analysis to vocalizations, bursts, and 3VOCS. These occurred for the first time at median ages of 6, 6, and 8 months, respectively; seven of the nine infants preserved this order, progressing from vocalizations to bursts to exactly 3VOCS, and 26 of 34 ordered pairs preserved the order within a series of trials (p = .002).

**Ear-Touching Task**

The experimenter's precise movement was to hold her index finger up near the side of her head, bend her elbow, and touch the tip of the finger to her ear. No infant ever imitated the form of this movement. They did, however, increase their rate of touching their own ears (sFACE) during the ear trials (compared with the baseline) at 6 months (seven of the eight infants, p = .035) and of touching their own eyes (SEAR) during the trials (compared with the baseline) at 7 months (five of the nine infants, ns). It is possible that earlier attempts at imitation, prior to 6 months, would have involved EFACE or EHAND as in the clapping task; but within the 6- and 7-month sessions we found no evidence that these acts preceded SFACE or SEAR. The relation between the latter two events within trials (sFACE preceding seAR) did emerge but not as strongly as was the case for the features of other tasks.

The most striking finding was that three infants responded to the ear-touching trials by clapping: one at 9, one at 10, and one at 11 months. None of the three had ever clapped during a task, and all did so on the very next clapping trials. The significance of this is discussed later.

**The Number Feature**

Clearly 4 and 5 were more difficult numbers than were 2 or 3. All infants accomplished and persistently produced 2SHAKES on the toy-shaking task. Seven infants produced 3VOCS during the vocalization task. In the bursts of mouth movement, even the four infants who "counted to 5" once did not consistently do so in subsequent bursts. Only three infants achieved 4CLAPS during the clapping task.

In principle, infants could have been considered to be imitating a feature of the clapping task if they produced a burst of four movements of any kind, or to be imitating the shaking task if they produced a burst of two movements (not necessarily shakes). We included the category ArmBursts—waving the hand or arm a given number of times—for that reason, but we failed to discover any correspondence between the number of movements in these bursts and the number in the tasks being modeled beyond what could be attributed to chance.

**Touching the Experimenter's Face or Hand**

At every age, the infants' rates of EFACE were significantly higher (by sign tests, requiring eight out of nine or seven out of eight infants to be higher) during the three tasks involving the face (mouthing, vocalization, and tongue protrusion) than during the other three tasks. Also at every age, their rates of EHAND were significantly higher during the three tasks performed by the hands (clap-
ping, toy shaking, ear touching) than during the three facial/oral tasks. It was particularly striking that infants did not touch the investigator's ear during the ear task; they touched her hand. Therefore, they were not attempting to perform the action on the model. What they may have been doing—and this would apply as well to the other tasks when they touched the investigator—was obtaining information about the action she was performing.

Discussion

Although our subjects did not cooperate tidily with our design, they did strongly confirm both of our predictions. Accommodation proceeded by approximations, feature by feature, in those tasks just as we had previously found in the mouthing task (Kaye & Marcus, 1978). In the present study, the order of acquisition over months was also usually reflected within the series of trials in any month when a better approximation to the model was made. It was an exceptional case when a higher approximation appeared on an early trial without the infant having to “work up to it” from cruder forms he or she had previously mastered.

Nevertheless, we also had some unexpected findings. After seeing all the infants make substantial progress over time on our videotapes, we were disappointed to find that when their data were combined at each age and plotted month by month, there appeared to be no pattern at all. This was because (a) some infants, having achieved part or all of a task in one month, would achieve less the next month; (b) there was great variation in the specific months in which specific improvements were made; and (c) infants often produced the components of a task later in a session—when we thought we were engaged in another task—after they had refused to go on with the first task.

These findings support our earlier conclusion (Kaye & Marcus, 1978) with respect to the methodology of studying imitation in infancy. One cannot expect to reveal in one or two trials what an infant is capable of imitating. The infant’s agenda is unlikely to match that of the investigator. Fortunately, parents normally provide dozens or even hundreds of trials, allow infants to elicit repeated demonstrations of actions they are working on, and are happy to acknowledge sensory-motor accomplishments outside of the contexts in which they have directly requested them (Kaye, 1981).

The study does not permit us to draw conclusions about the relative difficulty of imitating different kinds of action. It was apparently easier for babies to shake a toy twice and release it to the experimenter than it was to clap their hands four times. But that conclusion cannot be generalized to other possible variants of the two tasks, let alone to actions with an object versus actions without an object.

Inferences about the infant’s ability to imitate depend upon the criteria we adopt: All infants touched their ears during the ear-touching trials, but none did so in a manner that matched the model’s movement, and many touched their ears with just as high a rate during some other parts of the session. Perhaps another way of saying this is that the relative difficulty of the tasks depends upon the criterion the infant adopts in each case. It is reminiscent of Kaye’s (1977) study of detour reaching. Six-month-old infants whose mothers demonstrated the path around a Plexiglas barrier often imitated the hopping motion that the mothers used in order to make their hand movement salient: The mothers knew, and the experimenter knew, that this motion was irrelevant to obtaining the toy from behind the barrier; but how was the infant supposed to know?

Some intriguing moments in our videotapes were those when three different infants who had not previously clapped their hands did so during the ear-touching trials. The two actions are topologically equivalent: Two appendages are brought together. (Similarly, 1- and 2-year-olds will imitate a person doing pushups by simply placing their two hands on the floor—not by lying prone on the floor, which looks to an adult like a better match but is topologically different.) These infants, however, had already touched their ears at earlier sessions; we do not regard the clapping as a mistaken accommodation to ear touching. Instead, the model
seems to have been assimilated to an item in the infants' own sensory-motor agendas. When the experimenter touched her ear, it suggested the clapping schema they had worked on in previous sessions; they were now ready to perform it (perhaps they had done homework), and all did so again readily on the subsequent clapping trials.

Imitation by these infants was advanced relative to what others have reported (for example, shaking a bell or rattle appears on all of the major infant tests at 11 or 12 months, with a less strict criterion than what our infants achieved at 9 or 10 months); but it was less advanced than in our earlier study (Kaye and Marcus, 1978). The infants in that study, given only one task, achieved mouthbursts at 6 months (22 of 34 infants), whereas the infants in the present study with six different tasks to work on produced mouthbursts at a median age of 7 months. Our conclusions are the same: Active imitation plays a major role in the differentiation of schemas, but imitation should not be defined at the convenience of an experimenter. As Piaget (1945/1951) argued, accommodation to modeled actions proceeds by a highly selective accommodation of features, with infants defining for themselves what matches what. Their imitation over a period of time is not as capricious as it may appear, though the process itself remains mysterious.

Reference Notes


References