Synchrony, Pseudosynchrony, and Dissynchrony: Measuring the Entrainment Process in Mother-Infant Interactions

Frank J. Bernieri, J. Steven Reznick, and Robert Rosenthal
Harvard University

We examined a new method for studying synchrony (i.e., the coordination of movement between individuals in social interactions) in two studies. Raters viewed video clips of interactions and judged the level of synchrony occurring between a mother and a 14-month-old child. Some of the video clips were genuine interactions, but most were pseudointeractions artificially constructed from the genuine interactions via split-screen editing techniques. For mothers interacting with their own children, genuine synchrony was significantly higher than pseudosynchrony, a difference that increased with time. When mothers interacted with an unfamiliar child, however, genuine synchrony was not higher than pseudosynchrony. In fact, mothers with unfamiliar children showed a state of dissynchrony (levels of genuine synchrony significantly lower than levels of pseudosynchrony). Our results suggest that synchrony can be reliably rated, thus allowing future investigations to include such measurements when studying social interaction processes.

There is a small but growing literature concerning the processes involved in negotiating the many face-to-face encounters people experience daily (Davis, 1982; Kendon, Harris, & Key, 1975). The common theme throughout this literature is the treatment of human interactions, not individuals, as the unit of study (Scheflen, 1982). Within this context there is a developing interest in the concept of behavioral entrainment, a term first coined by Condon (Condon & Ogston, 1966). We define entrainment as the adjustment or moderation of behavior to coordinate or synchronize with another, similar to the synchronization occurring between members of an orchestra.

The investigation of behavioral entrainment, or synchrony, has been explored in many contexts and is thought by some to be an absolutely essential aspect of human communication and speech (Wylie, 1985). Condon and Ogston (1966) have demonstrated the existence of intrapersonal synchrony, in which the majority of all body movement is exactly synchronized to one's own speech rhythms. Interpersonal synchrony, the focus of our paper, can occur during social interactions when an individual synchronizes to the rhythms and movements of another person with whom he or she is interacting (Condon & Ogston, 1966, 1967; Kendon, 1970).

There are many possible functions of synchrony. It has been shown, for example, that newborn infants synchronize to adult speech but do not synchronize with non-speech-related sounds such as tapping, street noise, and white noise (Condon & Sander, 1974; Kato et al., 1983). An absence or lack of synchrony has also been observed in people who have dyslexia and other learning disabilities (Condon, 1982). This finding has led some to believe that synchrony is an essential precursor to language development (Wylie, 1985). In adults, synchrony is believed to help regulate and allow for smooth and efficient verbal exchange (Dittman & Llewellyn, 1969; Hadar, Steiner, & Rose, 1985). It might even serve a communicative function in and of itself, signifying interest or approval (Kendon, 1970; Tronick, Als, & Brazelton, 1977).

Synchrony has recently been theoretically linked to social rapport (Tickle-Degnen & Rosenthal, 1987). Some researchers have investigated its diagnostic potential for measuring rapport. The synchrony—rapport relationship has been studied in interviews (Dabbs, 1969), getting-acquainted situations (La France & Ickes, 1981), between doctors and patients (Charney, 1966; Trout & Rosenfeld, 1980), between teachers and students (La France & Broadbent, 1976), and between mothers and their children (Byers, 1982).

Definitions of Synchrony

Given this broad range of theoretical applicability, it is not surprising that synchrony has been measured in many different ways. The particular operationalization employed has depended on the definition used. Although all definitions involve some notion of behavior adjustment or entrainment to another, they can be classified into three broad categories: biological rhythms, simultaneous behavior, and perceived synchrony.

Biological Rhythms

In the biological sciences, synchrony is said to occur when one cyclical process becomes captured by and set to oscillate in rhythm with another process (McGrath & Kelly, 1986). From
this perspective, human behavior is understood to occur rhythmically and therefore can be described in terms of cycles, periods, frequencies, and amplitudes. Depending on the particular behavior and interaction in question, behavioral cycles in interpersonal behavior can range from milliseconds to hours (Mathiot & Carlock, 1982). One line of synchrony research has emphasized the importance of these behavioral rhythms or cycles (Brazielton, Koslowski, & Main, 1974; Bullowa, 1975; Davis, 1982; Stern, 1974). According to this approach, behavioral synchrony is defined as the degree of congruence between the behavioral cycles of two or more people. Much of the developmental literature involving adult–infant synchrony is of this type (Berghout-Austin & Peery, 1983; Bullowa, 1975; Rosenfeld, 1981; Stern, 1971). The behaviors of both individuals are coded separately and analyzed to identify behavioral cycles. Once each individual’s cycles have been determined, they are compared for congruency. Synchrony is the extent to which the cycles appear entrained or synchronized.

For example, Tronick et al. (1977) measured levels of engagement (i.e., particular combinations of body orientation, gaze, vocal activity, and facial expression) of both the mother and the child individually in a frame-by-frame film analysis and found cycles of engagement and disengagement lasting approximately 10 s. Furthermore, these cycles were in phase such that they were marked by periods of mutual engagement and disengagement that prevailed throughout the filmed session.

There are practical difficulties with measuring synchrony in this manner. First, the coding of behavior is time consuming and tedious. In many cases coding has been done with film frame by frame. At a standard rate of 24 frames of film for each second of real time, hours may be needed just to code a few seconds of behavior. Second, what should be coded is a problem. Any number of specific behaviors can be coded and combined for use as a behavioral metric. Issues involving the type, number, and combination of movements to code are present in every study, although in the developmental literature there has been some success in looking at cycles of engagement or attention (Beebe et al., 1982; Brazielton et al., 1974; Stern, 1974).

**Simultaneous Behavior**

One of the earliest articulations of a concept of synchrony was provided by McDougall (1926), who noted the tendency of spectators to assume the postural strains of dancers or athletes they were watching. Following in this vein, another line of research has focused on behavior mirroring and mimicry (Bavelas, Black, Lemery, & Mullett, 1986; Dabbs, 1969; O’Toole & Dubin, 1968). From this perspective, synchrony occurs when one person directly imitates or mirrors another person’s limb movements or body configuration. A broader definition includes simultaneous affect and attitudes (Siegmans & Reynolds, 1982), as well as simultaneous movement changes (Cappella, 1981). Synchrony is thus operationalized as the quantity of simultaneous behavior, however measured.

The coding required for simultaneous behavior measurements is often long and tedious (Kendon, 1970) and controversial (Cappella, 1981; Gatewood & Rosenwein, 1981; McDowall, 1978). It should be noted, however, that the more macrobehaviors such as gestures and postural orientations have been much easier to code (Dabbs, 1969; La France & Broadbent, 1976), albeit theoretically more restrictive.

**Perceived Synchrony**

The two approaches described earlier are based on mechanistic definitions of the nature of behavioral synchrony. Thus, their focus is on generating objective criteria for identifying, coding, and quantifying exact synchronous behavior. Investigators have been more concerned with the mechanisms supporting synchrony than with its phenomenological effects. Although the two general approaches described earlier have generated the most empirical research to date, a third approach merits exploration.

The essence of the synchrony concept is the sustained nonrandom co-occurrence of two behavioral phenomena. In the biological rhythms approach, the phenomena are behavioral cycles. In the simultaneous behavior approach, they are specific behaviors. There is no theoretical reason favoring either approach as being more important or essential for understanding the concept. The fundamental feature common to both approaches is the apparent unification of two behavioral elements into a meaningfully described whole, synchronous event. The elements of this event may be simultaneous, identical, and in phase or alternating, mirrored, and out of phase. The essential feature is that when the elements are put together, they create a "whole," or perceptual unit (Chapple, 1982). Behavioral synchrony, therefore, can be studied as a perceptual social phenomenon, one that is not normally attended to but observable nevertheless.

We feel it is premature to assume, as others have done, that the subtleties of synchrony are beyond normal conscious perception (Condon & Ogston, 1967; Wolff, 1967). The human perceiver, in fact, may be a valuable measuring device who has been severely underused. When writing to fellow researchers of nonverbal behavior, Scherer and Ekman advised that "at present, perceptual units can only be identified and categorized by human judges, because even the most advanced computers still lack the pattern-recognition ability required for this task" (Scherer & Ekman, 1982, p. 33). Synchrony, therefore, is operationalized here as the extent of gestaltlike harmoniousness or meshing of interpersonal behaviors as judged by a group of raters.

**Purpose**

In the following two studies, we investigated the validity and utility of using human raters to measure behavioral entrainment, or synchrony. We wanted to demonstrate that synchrony exists in mother–infant interactions and is perceptible to outside observers. We examined the mother–child dyad because this dyad type has been the most closely examined for synchrony. The data generated by our synchrony judges represent subjective ratings of synchrony and thus constitute relative rather than absolute scales. Therefore, although these raw synchrony ratings might suggest relative differences across conditions, they do not support inferences concerning the absolute level of nonrandom sustained co-occurrences of behavioral events (i.e.,
bravioral entrainment). To demonstrate the existence of genuine nonrandom entrainment, a meaningful baseline measure of expected random behavioral co-occurrences was needed against which ratings of synchrony could be compared (Cappella, 1981).

A group of judges viewed silent video clips of dyadic interactions and rated the degree of synchrony, or entrainment, present in each unit or clip. Three ratings were collected representing but not precisely corresponding to the three conceptual definitions described earlier: (a) the judges’ perception of simultaneous movement; (b) tempo, or rhythm, similarity; and (c) gestaltlike rating of the harmonious meshing of interpersonal behaviors.

Judges viewed a series of video clips of what appeared to be mothers interacting with 14-month-old infants. Some interactions were genuine, but others paired interactants who were not actually interacting with each other. These pseudointeractions were constructed by isolating the video image of each interactant and then pairing them with the video images of other interactants recorded in other interactions. The resulting synchrony ratings of these pseudointeractions, which we refer to as pseudosynchrony, could not have been due to entrainment because the on-screen interactants being rated as a unit were not interacting with each other.

True synchrony ratings were then compared with the baseline measurements of pseudosynchrony (i.e., expected levels of random behavioral co-occurrences). The inference of genuine behavioral entrainment would be supported by demonstrating significantly higher synchrony ratings for true interaction video clips than for the pseudointeraction video clips.

Study I

Method

Subjects

Eight mothers and their children (aged 14–18 months) participated in this study. Mothers were told that the nonverbal aspects of mother-child interactions were being investigated. Before mothers agreed to participate, we told them that their child would be interacting alone with an unfamiliar woman sometime during the study. Mothers were paid for their participation.

Procedure

Mothers and their children were brought into the laboratory in tetrads consisting of 2 mothers, 1 male child, and 1 female child. Four 3- to 5-min dyadic interactions were videotaped from each tetrad. The 2 mothers each interacted twice, first with their own child and then with a child whom they had just met. Fifty-second dyadic sequences were extracted from the first and third minutes of these four interactions.

The child was placed in a portable high chair positioned such that the mother and child sat facing each other across a 4-ft (1.21 m) wide table. Mothers were told not to reach out to the child and not to cross the center of the table with their hands. This restriction was necessary because the editing procedure that followed required the video image of each subject to be isolated. Mothers were instructed to talk to the child (either their own or the child whom they had just met) and to keep the child engaged. Mothers were encouraged to be as animated and as talkative as possible.

Two cameras recorded the interactions, one isolated on each individ-ual. Mothers were always facing right. Cameras were positioned side by side with their fields of vision crossed such that a profile shot was produced with just enough frontal angle to expose both eyes of each subject. The pictures of both cameras were combined to form a single video clip of the interaction. As an aid to the synchronization procedure necessary for recombining these two pictures, a time-zero marker was incorporated into the videotaped segments at the onset of the interaction and at the beginning of the third minute.

Editing

The video recording procedure described earlier enabled us to recombine the isolated images via a split-screen generator to re-create the original dyadic interaction. More important, it allowed us to combine and cross individual sequences to form artificially created on-screen dyadic interactions, or pseudointeractions, that had never taken place. In all, four types of dyadic interaction clips were constructed from these individual behavior segments.

True dyadic clips. The first type of interaction clip re-created the originally recorded interaction. This was achieved manually by forwarding the two videocassette recorders (VCRs) to a time-zero point and starting both simultaneously. The sound track provided a good criterion for evaluating the temporal resynchronization. Imperfect synchronization was indicated by an echo, which meant that one VCR was ahead of the other. When the video image time-zero markers were exactly synchronized, the sound tracks were also synchronized.

In addition to the true dyadic interaction clips, three types of pseudodyadic interaction clips were similarly constructed. Because the pairings did not reconstruct the original interaction, the audio portion was not used as a synchronization check.

Altered time frame. These pseudointeraction clips were constructed by pairing the recorded behavior of one interactant in one time period with the behavior of his or her partner in the other time period. For example, a segment of a mother in the third minute was paired with the child she was interacting with in the first minute. Similarly, a segment of her in the first minute was paired with the segment of the child in the third minute. Thus, time was altered but the partners had actually interacted.

Switched partners. In the switched-partners pseudoclip, individual behaviors from different interactions were paired, such that both people appearing on the video monitor were actually interacting with a different partner off-screen. Thus, although the video clip showed a mother and her own child, in reality the mother was interacting with an unfamiliar child and the child was interacting with an unrelated mother. For this type of clip, the time period was held constant: Only behavioral sequences recorded during the same minute of each interaction were combined.

Double-crossed. The last type of pseudoclip was constructed by combining the previous two methods. That is, the apparent interaction clip contained switched partners as well as the behavioral sequences taken at different temporal locations of the interaction.

Design

The editing procedure generated four video clips from each originally recorded dyadic interaction sequence. One re-created the original interaction, and the other three generated pseudointeraction sequences. The

1 The editing procedure caused a slight but detectable decrease in the quality of the video image. We chose to reconstruct the original dyadic interaction rather than record it directly in order to equate the quality of the true dyadic interaction segments with those of the pseudointeraction segments.
three pseudointeraction clips were then compared with the true clips on measures of nonverbal behavioral synchrony.

We began, for example, with a 50-s clip of a mother interacting with her own child in the first minute. This clip was defined as the true related mother–child first-minute clip. Using the child as the reference point, we then constructed three pseudolated mother–child first-minute clips: (a) Altered time frame was the pairing of the original segment of the mother interacting with her own child in the third minute with the segment of them interacting in the first minute. (b) Switched partners was the pairing of a video segment of the mother interacting with an unrelated child during the first minute with the segment of her own child interacting with another woman during their first minute of interaction. (c) Double-crossed was the pairing of the mother interacting with the unrelated child during the third minute with the clip of her own child interacting with the other woman during the first minute.

Eight genuine interaction sequences (mother with a child) were recorded from each of the mother–child tetrads (two mothers interacting with two children in two time periods). Each interaction sequence generated four video clips, one true and three pseudo. Thus, a total of 128 video composite clips from four tetrads were presented to judges for synchrony ratings. Only the video portion of these clips were shown to judges.

Judges

Independent judgments of synchrony were collected from 8 male and 8 female raters between the ages of 18 and 30 years. Judges attended four rating sessions, each occurring at least 24 hours apart, and were paid approximately $20. The judges were randomly divided into four groups, each containing an equal number of men and women. Raters were asked to judge independently and not to talk during the rating sessions. All four groups saw the same 128 stimulus clips, 32 per session, but in four different sequences, creating a Latin square design for the rating schedule.

Rating Form

Judges used a rating form based on three aspects of synchrony. A cover sheet explained what each rating was designed to measure, and judges were told that the rating definitions could be interpreted "loosely and liberally." Each was rated on a 9-point Likert scale. The descriptions provided to the judges were as follows:

1. SIMULTANEOUS MOVEMENT—This reflects the quantity or degree of movement that appears to begin or end at the same moment. For example, if a mother begins to turn her head at the precise moment that a child lifts an arm off of a table, it is an instance of simultaneous movement.

2. TEMPO SIMILARITY—Assume that all people have built-in tempos or speeds at which their behavior is set (much like the tempo of an orchestra follows at a concert). Rate the degree to which the two people in the clip seem to be "marching to the beat of the same drummer."

3. COORDINATION AND SMOOTHNESS—Assume you are viewing a choreographed dance rather than a social interaction. How smoothly does the interactants' flow of behavior intertwine, or mesh evenly and smoothly?

Judges were not given further information on how to rate. In fact, they were told that once the study began, the experimenter would not answer any questions concerning the stimulus tapes. In order to make their task easier, judges viewed each 50-s silent video clip twice in succession while making their ratings.

Rating Procedure

The aim of the rating procedure was to have the judges rate the chrony in all 128 stimulus clips without realizing that they were viewing artificially constructed pseudointeractions. Several measures were introduced to achieve this aim.

First, the stimulus tapes, each containing the 32 clips, were constructed such that all of the clips from a particular tetrad of mothers and children were of the same type (i.e., either true, altered time frame, switched partners, or double-crossed). This prevented the judges from seeing any of the originally isolated target video clips more than once during a video session.

For example, on Day 1 a group of judges rated all of the true interaction clips from Tetrads 1 and 2. On Day 2 they rated all of the altered time-frame clips from that tetrad; on Day 3 they rated switched partners; and on Day 4 they rated the double-crossed. On each day the judges also rated all four tetrads, each in a different type of reconstructed clip. On one of the session tapes, for instance, the eight clips involving Tetrads 1 and 2 were true, the eight clips of Tetrads 2 and 3 had an altered time frame, the clips of Tetrads 3 and 4 had switched partners, and the clips of Tetrads 4 were double-crossed. As a consequence, the judges rated clips from all four tetrads (16 subjects) and all four types of clips (true and pseudo) without ever viewing the same exact behavior of any interactant twice in a rating session.

We thought that by requiring at least 24 hours to pass before seeing the next series of 32 clips, the judges would be less likely to remember exact video image pairings. We expected judges to remember seeing some of the same isolated behaviors the day before, which they were told would happen. But given that they had no expectation of pseudointeraction construction, perhaps they would attribute the familiarity of any clip to intentionally planned repeated trials. If they did have memories of different pairings, we hoped that they would attribute this to faulty memory, a reasonable assumption given that there were several similar clips.

Also, we gave special instructions to raters to help avert suspicions of pseudointeraction clips. On the first day, judges were informed that they would be seeing several video clips of some mothers interacting with their children. The mother would either be with her own child or with an unrelated child. They were also told that due to the novelty of the task and the possible practice effects that might systematically affect their ratings over the course of four rating sessions, they would be rating many clips more than once. We explained that the rationale was to get many measurements of the same interaction throughout the study in order to measure how their ratings of this novel task was affected by practice.

The counterbalancing of the stimulus materials, together with the rating schedule and the instructions given to the raters, resulted in the judges' belief that each video clip they rated was of a genuine interaction. When prompted at debriefing, only 1 of the 16 judges mentioned the possibility of artificially paired clips. Also, when the true nature of the tapes was revealed to the judges, they were asked to guess what proportion were artificially constructed clips. No one made an estimate greater than one third of the total when, in fact, 75% of all clips rated were pseudointeractions.

Results

Preliminary Analyses

Composite variables. We constructed our rating scales to represent three aspects of synchrony: simultaneous movement, tempo similarity, and gestalitike smoothness or coordination. However, the three ratings were highly correlated, ranging from $r = .72$ to $r = .80$. Two explanations are possible: Either all thn
hypothesized components of synchrony were highly interrelated or our judges were unable to distinguish between and assess the three components. In either case, we decided to average the three variables to form a composite variable of perceived 

**global synchrony.** The results were based on this composite measure.

**Reliability of judges’ ratings.** The reliability of the judges’ ratings of global synchrony was computed by using intraclass correlations (Rosenthal, 1982, 1987). Because we were concerned about the level of agreement the judges attained across the 128 rated stimulus clips, the average interjudge correlation was computed across these 128 units. The intraclass r of .23 yielded a Spearman-Brown effective reliability coefficient for the mean of our 16 judges of r = .83.

**Data reduction.** Tetrads were chosen as the unit of analysis for significance testing because we wanted to generalize the results to the synchronizers, not to the synchrony observers. Tetrads were used instead of interaction dyads because the tetrads were the only truly independent sampling unit. The resulting design consisted of four tetrads that were each measured 32 times for synchrony. Preliminary analyses involving sex of child did not yield significant results, so we dropped this from the analysis by averaging across the appropriate cells. This resulted in 16 repeated measurements for each tetrad.

**Statistical analyses.** A limitation of this design was its set of relatively low-power 3-df error terms. According to the $F_{\text{max}}$ criterion, a set of fifteen 3-df error terms could vary by a factor as much as 150 to 1 before being significantly different (Walker & Lev, 1953). To provide a better estimated and more stable error term for all of the possible effects in the study, we pooled the error terms for all of the repeated measures factors by nomination. That is, we refit the model to a one-way, 16-level repeated measures analysis of variance (ANOVA). We derived the subsequent analyses from a series of planned 1-df contrasts performed on a 4 Tetrads x 16 Repeated Measures ANOVA (Rosenthal & Rosnow, 1984a). All significance tests reported in this section were computed via planned contrasts performed on a 15-df omnibus $F$ for the repeated measures effects across the 16 cells in the design. This procedure yields a better and more stable estimate of error variance as long as the mean squares that are aggregated do not differ significantly in magnitude (i.e., $F_{\text{max}}$ criterion).

Except for the main effect of relatedness, which had a relatively larger mean square error term, all of the error terms were well within the accepted $F_{\text{max}}$ criterion. Pooling the excessively large error term for the main effect of relatedness, which was not an effect of interest, with all of the others made its significance level slightly too low and the rest of the significance levels too high (i.e., conservative). Although this procedure increased the likelihood of Type II errors for all of the effects of interest, it seemed less arbitrary than nominating only a subset of the error terms for pooling. As a result, the contrast Fs or ts based on this generalized error term will differ from those based on the local error terms but in a generally unbiased fashion (Green & Tukey, 1960; Rosenthal & Rosnow, 1984b).

**Overall Test of Global Synchrony.**

Our main goal was to determine whether judges perceived greater synchrony within the true interaction clips than within the pseudointeraction clips. A planned contrast testing the difference between the true interaction clips against the mean of the three pseudointeraction clips across all interactions (including both mother-unfamiliar child dyads and mother-own child dyads) was not significant, overall $F(1, 45) = 1.97, r(3) = .63$. When the effects of relatedness and time were analyzed, however, a more interesting pattern emerged.

The mean ratings for global synchrony, broken down by relatedness and time, appear in Table 1. Mothers interacting with their own children showed a significantly higher level of rated synchrony in the true interaction clips than in the pseudointeraction clips, $F(1, 45) = 5.58, r(3) = .81, p < .025$. Furthermore, this difference increased from the first to the third minute, $F(1, 45) = 4.94, r(3) = .79, p < .05$. Mothers interacting with unrelated children, however, evidenced no overall true-pseudo difference, $F(1, 45) < 1$. In fact, these unrelated dyads experienced a significant decline in the true synchrony over pseudosynchrony advantage over time, $F(1, 45) = 4.57, r = .78, p < .05$. The contrast directly testing the interaction effect of relatedness on the true-pseudo difference was marginally significant, $F(1, 45) = 3.75, r(3) = .75, p < .10$.

Although there was no main effect of time on the true synchrony over pseudosynchrony advantage, there was a significant three-way interaction involving relatedness, $F(1, 45) = 9.51, r(3) = .87, p < .005$. An inspection of the raw means showed that although the true synchrony over pseudosynchrony advantage increased over time for related mothers and children, the advantage actually reversed for the unrelated dyad by the end of the third minute. The reason for this significant decline seems to be the peculiar finding that in the third minute of the unrelated dyadic interactions, the true synchrony ratings dropped so low that the pseudosynchrony ratings were actually higher than the true interaction synchrony ratings, $F(1, 45) = 4.54, r(3) = .78, p < .05$. For related mothers and children, however, the third minute yielded the greatest true synchrony over pseudosynchrony advantage, $F(1, 45) = 6.60, r(3) = .83, p < .025$.

**Comparison of Pseudointeraction Clip Types**

The next series of analyses explored the individual effects of each type of pseudointeraction clip. Although the power of this

---

1 It was possible to use mothers (i.e., Mother A and Mother B) as a variable in place of sex of child. The present data set was analyzed both ways, and neither the sex-of-child variable nor the mother variable generated any significant results. In addition, the arbitrariness of the two levels of the mother variable supported our decision to drop it.

2 The computation of the effect size, Pearson r, from significant tests was done using the following formula (Rosenthal & Rosnow, 1984b, p. 347):

$$r = \sqrt{\frac{F_{1, n}}{F_{1, n} + df_{\text{error}}}}$$

The magnitude of the effect is indicated by $r^2$, which is an estimate of the variance accounted for. When we computed an effect size for contrasts on repeated measures, the degrees of freedom of the sampling units, rather than the degrees of freedom of the error term, were used (Rosenthal & Rosnow, 1984a, p. 73). Thus, all effect sizes reported are based on 3 df, derived from four tetrads.
Table 1
Mean Ratings of Global Synchrony

<table>
<thead>
<tr>
<th>Type of stimulus clip reconstruction</th>
<th>Related mother-child dyad</th>
<th>Unrelated mother-child dyad</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First</td>
<td>Third</td>
</tr>
<tr>
<td>True interaction</td>
<td>5.48</td>
<td>5.97</td>
</tr>
<tr>
<td>Mean of pseudointeraction</td>
<td>5.14</td>
<td>5.29</td>
</tr>
<tr>
<td>True-pseudo difference</td>
<td>.34</td>
<td>.68***</td>
</tr>
</tbody>
</table>

Note. The values are means computed across 18 dyads (4 tetrads, each with a male and female child dyad). The significance of true-pseudo differences was tested using planned 1-df contrasts on a 15-df within-tetrad effect term. All p values were derived from contrast Fs(1, 45).

*p < .05. ** p < .025. *** p < .005.

study precludes any definite conclusions about the possible differences among the three controls, these findings may help to guide future research in this area. The hypothesis we tested in each instance was that rated global synchrony would be higher in true interaction sequences than in a particular pseudointeraction control clip. A series of matched-pair t tests, computed via the contrast method described earlier, pitted the global synchrony rating of true interactions against each of their three controls.

Overall. Across all types of interactions, only the switched-partners pseudointeraction clips showed lower ratings of synchrony than the true interaction clips, F(1, 45) = 3.34, r(3) = .73, p < .08. Surprisingly, the double-crossed pseudointeraction clips that had both altered time frames and switched partners showed no robust effect, F(1, 45) < 1.

Relation and time. A more detailed analysis of the differences between each pseudocontrol and true interaction clips appears in Table 2, where synchrony scores for each pseudointeraction type have been subtracted from the appropriate true interaction. Within the related mother-child interactions, the double-crossed and the switched-partners control clips showed significantly less synchrony than the true interaction clips, F(1, 45) = 4.41, r(3) = .77, p < .05, and F(1, 45) = 9.03, r(3) = .87, p < .001, respectively. In the unrelated mother-child interactions, none of the individual control clips differed from the true clips.

The type of control clip made no difference when time was considered. All three pseudocontrols showed no differences across time periods. All three, however, picked up the significant interaction effect of Relation X Time on the true-pseudo difference, showing relatively higher gains in true synchrony over pseudosynchrony over time within the related mother-child dyadic interactions compared with unrelated mother-child dyadic interactions. Use of the double-crossed and switched-partners controls showed a strong Relation X Time interaction effect on the difference, F(1, 45) = 7.66, r(3) = .85, p < .01, and F(1, 45) = 7.51, r(3) = .85, p < .01, respectively. The altered time-frame interaction effect was not as strong, F(1, 45) = 4.19, r(3) = .76, p < .05.

In general, the pattern of results from this exploratory analysis does not suggest any great differences between the types of control clips in nature, only in relative magnitude. The pattern of the results involving true interaction synchrony over pseudointeraction synchrony seemed to be strongest with the switched-partner control and weakest with the altered time-frame control.

Discussion of Study 1

To justify a new method of measurement, one must address issues in both reliability and validity. We now evaluate the use of naive raters in measuring dyadic coordination, or synchrony.

Reliability

Although the average interjudge correlation of synchrony was low, it was comparable to the level attained and used in other studies concerned with the subtleties of nonverbal behavior (Babad, Bernieri, & Rosenthal, 1986; Blanc, Rosenthal, &

Table 2
Difference Ratings (True Interactions Minus Pseudointeractions) of Global Synchrony

<table>
<thead>
<tr>
<th>Type of pseudointeraction clip control</th>
<th>Related mother-child dyads</th>
<th>Unrelated mother-child dyads</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First</td>
<td>Third</td>
</tr>
<tr>
<td>Double-crossed</td>
<td>.66*</td>
<td>.45</td>
</tr>
<tr>
<td>Switched partners</td>
<td>.31</td>
<td>1.22****</td>
</tr>
<tr>
<td>Altered time frame</td>
<td>.05</td>
<td>.31</td>
</tr>
</tbody>
</table>

Note. The means were averaged across eight dyadic interactions (four tetrads, each with a male and female child dyad) and across 16 raters. All significance testing was performed using planned 1-df contrasts on a 15-df within-tetrad effect term. All p values were derived from contrast Fs(1, 45).

*p < .10. ** p < .05. *** p < .01. **** p < .001.
Vannicelli, 1986). As is true in other areas of nonverbal investigation, it is important for investigators to collect synchrony ratings from a relatively large group of judges. In this study, 16 judges were needed to raise the effective reliability of their group mean to $r = .83$.

Validity

To establish the use of raters as a viable method for measuring the synchrony within interactions, we needed to show that the judges rated synchrony higher when it should have been higher, and lower when it should have been lower or even nonexistent. This was achieved in two ways: (a) measurement across different kinds of true interactions and (b) use of pseudointeraction controls.

**Measurement across different kinds of true interactions.** On the basis of past research we would expect mothers interacting with their own children to synchronize often. A typical 1-year-old has considerable physical contact with its mother since birth, and this dyad would have had ample time to become intimately familiar with each other’s mannerisms and behavioral rhythms. Also, because they are biologically related, there is a greater probability that natural biological rhythms will be synchronous. On the other hand, it is less likely that a woman will synchronize with an unrelated child whom she had just met, with the likelihood of biologically based or practice-based synchronous behavioral rhythms being much less. Clearly, the latter situation is less conducive to synchrony than the former, and as predicted, judges rated mother-child dyads higher in synchrony than unrelated mother-child dyads, $F(1, 45) = 4.98$, $r(3) = .79$, $p < .05$.

**True interactions versus pseudocounters.** It is possible that judges may have had implicit theories of synchrony, expecting or assuming that synchrony would or should be more pronounced in interactions in which people are related, know each other, and look happy. Although judges were not told which child belonged to which mother, they may have correctly deduced it from their physical appearance or behavior. Thus, an alternative explanation for the effect of relatedness could be that judges, unable to perceive the phenomenon directly, deduced it indirectly.

The rating of pseudointeraction clips directly controlled for and tested these alternative explanations. For example, suppose there was a tendency for judges to rate synchrony high because the interactants appeared related. Judges rating the switched-partners pseudointeraction control clip should rate the related on-screen interactants, whose behaviors were recorded in other interactions, just as high as when they rated the true interaction clip of the same two interactants. This did not happen. As reported earlier, the ratings of synchrony for true interaction clips were higher.

The pseudointeraction clips allowed us to completely control for all of the attributes of each interactant. The same 50-s video clip of an isolated individual was rated four times (paired each time with a different clip of their partner) by the same judge. Only once was the judge rating the true synchrony of the dyadic interaction, brought about through a mutual entrainment process. For the other three times, the judge was rating pseudosynchrony, the apparent behavior coordination (synchrony) that occurred just by chance, projection, or intuition. These results suggest that for related mother-child dyads, true synchrony was significantly greater than pseudosynchrony. If judges had been using only the individual attributes or behaviors of each interactant to intuit or calculate synchrony, there would have been no measurable synchrony-pseudosynchrony difference.

A Further Validation

A final validity check was made to test directly whether certain identifiable attributes of the individuals’ behavior affected ratings of synchrony. Specifically, we tested whether judges used the interactants’ affect to indirectly deduce their ratings of synchrony. The problem with testing this hypothesis is that there may be a preexisting link between rapport and synchrony in normal human interaction. Investigators have hypothesized that interactant rapport and level of synchrony go hand in hand (Kendon, 1970; Tickle-Degnen & Rosenthal, 1987), and there is some supporting empirical evidence for this. A recent meta-analysis has indicated that posture mirroring, which is believed to be an aspect of synchrony, is positively related to feelings of togetherness and rapport (Tickle-Degnen, Rosenthal, & Harri- gan, 1987). If there is a relationship between synchrony and rapport, then for any sample of true interactions, judges’ ratings of synchrony should be positively correlated with ratings of the interactants’ positive affect and rapport. Across a sample of pseudointeractions—where there is no genuine interaction and hence no genuine rapport—this relationship should not hold.

Therefore, we tested two propositions: (a) There is a relationship between rapport and synchrony within true interactions; and (b) because judges are not influenced solely by individual affect when rating synchrony, there is no relationship between rapport and synchrony within pseudointeractions. These two hypotheses together address the differential validity of the construct of synchrony and its measurement.

**Study 2**

**Method**

**Procedure**

The original 64 isolated interactant video clips (i.e., 8 mothers in 2 interactions at 2 time periods and 8 children in 2 interactions at 2 time periods) used to construct the paired interaction video segments were shown to a different group of raters for judgments of affect. A stimulus tape was generated by randomizing the 32 clips of the women in the first half of the tape and the 32 clips of the children in the second half. Independent ratings were collected of warmth, activity, anger, enthusiasm, frustration, happiness, passivity, boredom, engagement, and vocal activity.

**Judges**

Four male and four female undergraduate judges rated all 64 clips in one rating session and were paid approximately $6 for their task. Half of the judges rated the clips of the mothers first, and the other half rated the clips of the children first. Raters were instructed to make their judgments independently and not to comment.
Table 3  
Correlations Between Segment Synchrony and Rating Composites

<table>
<thead>
<tr>
<th>Composite ratings</th>
<th>True segments</th>
<th>Pseudo segments</th>
<th>True-pseudo difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child positivity(^a)</td>
<td>.47*</td>
<td>.12</td>
<td>.35*</td>
</tr>
<tr>
<td>Child interest(^b)</td>
<td>-.03</td>
<td>-.09</td>
<td>.06</td>
</tr>
<tr>
<td>Child happy/vocal(^a)</td>
<td>.13</td>
<td>-.07</td>
<td>.20</td>
</tr>
<tr>
<td>Child passivity</td>
<td>.21</td>
<td>.06</td>
<td>.15</td>
</tr>
<tr>
<td>Mother positivity(^a)</td>
<td>-.09</td>
<td>.11</td>
<td>-.20</td>
</tr>
<tr>
<td>Mother effort(^\ast)</td>
<td>.17</td>
<td>.06</td>
<td>.13</td>
</tr>
</tbody>
</table>

*Note. The median intercorrelation between composites is \(r = .12\).
\(^a\) Child positivity = mean of warmth, not angry, and not frustrated.
\(^b\) Child interest = mean of active, enthusiastic, engaged, or not bored.
\(^\ast\) Child happy/vocal = mean of happy and vocal.
\(^\ast\) Mother positivity = mean of happy, not angry, and not frustrated.
\(^\ast\) Mother effort = mean of warm, active, enthusiastic, or not passive, not bored, engaged, and vocal.

\(^*\) \(p < .01\). 

Preliminary Analyses

Intraclass reliabilities for all 10 rating variables were computed separately for child ratings and mother ratings. For the mother video clips, the median intraclass \(r\) across the 10 rating variables was .37. Taking into account that the means of eight judges were used in the analysis, the effective reliability of our dependent measure via the Spearman-Brown correction was \(r = .82\). For the clips of the children, the median intraclass \(r\) was much lower, \(r = .10\). The effective reliability via the Spearman-Brown correction was \(r = .48\).

To reduce the redundancy in the data and to increase the reliability of the child ratings, a principal-components analysis was performed to help form composite rating variables (Rosenthal, 1987). The analysis was done separately for the mother and child ratings. From these analyses, 4 composite variables were formed from the initial 10 child ratings, and 2 were formed from the mother ratings. Unit weighting was used to construct the following 6 variables: (a) mother positivity—composed of happiness, reversed frustration, and reversed anger; (b) mother effort—composed of warmth, activity, enthusiasm, engagement, vocal activity, reversed passivity, and reversed boredom; (c) child positivity—composed of warmth, reversed anger, and reversed frustration; (d) child interest—composed of activity, enthusiasm, engagement, and reversed boredom; (e) child happy/vocal—composed of happiness and vocal activity; and (f) child passivity—composed of passivity.

Results

The independent ratings of the mothers and children were paired in exactly the same manner as the isolated subject video clips in Study 1. This yielded 128 pairings of mother-child affect. Each video clip pairing in Study 1 had a corresponding set of mother-child affect ratings from Study 2.

We hypothesized that within true interactions there would be a relationship between the interactant’s rapport (i.e., individual affect) and their level of synchrony and that this relationship would not exist within the pseudointeraction pairings. Table 3 shows the relevant set of correlations. As predicted, the positive affect of the child (i.e., warmth, absence of anger, and absence of frustration) correlated with ratings of global synchrony within genuine interactions. More important, child positivity did not correlate with synchrony across the pseudointeraction clips. Positive affect of the mother, however, did not correlate with synchrony ratings in either true or pseudo clips. None of the remaining four rating variables showed any relationship to synchrony in either the true interaction or pseudointeraction clips.

Because the rating of synchrony depends on the simultaneous viewing of both interactants, it is possible that a relationship existed between synchrony and the joint affect of mother and child. To explore this possibility, the geometric mean of the two partners’ ratings were computed for each interaction for each of the 10 original ratings. The original ratings were used because the composite variables formed for the mothers and children were not comparable and therefore could not be directly combined. Rated synchrony was correlated with the 10 newly formed joint affect ratings. The results are shown in Table 4. Once again, although some ratings appear related to genuine synchrony, none of the 10 ratings appear significantly related to pseudosynchrony.

Discussion of Study 2

This set of results rules out the alternative explanation that judges may have been rating something other than synchrony. Had they been using any characteristics of the individual interactants (e.g., happiness, enthusiasm, or activity) to rate the level of synchrony, there should have been a significant correlation between these characteristics and synchrony regardless of whether the pairings represented true interactions. It is clear from Tables 3 and 4 that no aspect of either the mother, the child, or the dyad was significantly correlated with synchrony ratings across the pseudointeraction clips. This was true even when the original 20 items (before the formation of the composite variables) were considered. Although our list was not exhaustive, we did choose those that seemed the most theoretically linked to the concept.

Unfortunately, some of the child variables may not have been reliable. Even with eight raters, the effective reliabilities of these variables were all less than \(r = .50\), much lower than the reliabilities found for all of the other rating variables, \(r > .70\). Consequently, results concerning these less reliable child ratings are not conclusive.

Table 4  
Correlations Between Segment Synchrony and the Geometric Mean of Child and Mother Ratings

<table>
<thead>
<tr>
<th>Geometric mean rating</th>
<th>True segments</th>
<th>Pseudo segments</th>
<th>True-pseudo difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Warm</td>
<td>.44*</td>
<td>.12</td>
<td>.32</td>
</tr>
<tr>
<td>Active</td>
<td>.14</td>
<td>.03</td>
<td>.11</td>
</tr>
<tr>
<td>Angry</td>
<td>-.15</td>
<td>-.18</td>
<td>.03</td>
</tr>
<tr>
<td>Enthusiastic</td>
<td>.21</td>
<td>.01</td>
<td>.20</td>
</tr>
<tr>
<td>Frustrated</td>
<td>-.21</td>
<td>-.11</td>
<td>-.10</td>
</tr>
<tr>
<td>Happy</td>
<td>.39*</td>
<td>.16</td>
<td>.23</td>
</tr>
<tr>
<td>Passive</td>
<td>.12</td>
<td>.04</td>
<td>.08</td>
</tr>
<tr>
<td>Bored</td>
<td>-.23</td>
<td>-.07</td>
<td>-.16</td>
</tr>
<tr>
<td>Engaged</td>
<td>.19</td>
<td>.00</td>
<td>.19</td>
</tr>
<tr>
<td>Vocal</td>
<td>.04</td>
<td>-.14</td>
<td>.18</td>
</tr>
</tbody>
</table>

\(^*\) \(p < .05\).
The lack of a relationship between any of the mother rating variables and synchrony was a surprise because synchrony, by definition, is mutually dependent on both interactants within an interaction. Even when the effect of both the mother and the child was combined and used to predict synchrony, the effects were weaker than for the child ratings alone. One possible explanation is that for the present type of interaction, the child was somehow more responsible for determining a maximum potential for synchrony. In other words, if a child was unable or unmotivated to synchronize, nothing that the mother could have done would have increased the synchrony within the interaction. Another possibility is that the infants' affect and spontaneous behavior were more diagnostic of synchrony because of their relatively increased spontaneity. It may be that children at this age are more affectively reactive to states of synchrony than are adults.

General Discussion

We can conclude from these results that judges rated synchrony directly. By using artificially constructed pseudointeraction clips, we showed that judges were not simply projecting their expectations of synchrony onto the stimulus tapes. If they had, the synchrony ratings generated from these pseudointeraction clips (pseudosynchrony) would not have differed from those of true interaction clips (entainment synchrony). We also know from Study 2 that subjects were not rating synchrony indirectly from more easily observed characteristics of the mother and child.

We now have evidence that a group of judges can reliably rate a valid measure of synchrony in dyadic interactions. Although the interactions used here were of a special type (i.e., mothers and 14-month-old children), there appears to be no a priori reason why this technique would not generalize to other interactions.

The Differential Effects of Three Types of Pseudosynchrony

We have some evidence to suggest that at least three types of pseudointeraction clips are usable. The altered time-frame, switched-partners, and double-crossed techniques all yielded a similar pattern of results, although they differed in their sensitivity.

The pseudosynchrony measured from altered time-frame interaction clips consistently showed smaller differences from true synchrony. This was not surprising because these clips were characterized as closest to the original interaction clips. The only difference was a time slip of one of the interactants. By their nature they pose the most stringent test of an entrainment hypothesis.

Some investigators view synchrony as a cyclical phenomenon. If this view is correct, it is possible that a time slip could effectively induce a 360° phase shift in one of the interactants' cycle that could result in relatively undisturbed overt synchrony. Because only two time periods were studied here, this hypothesis was not testable. Future investigations involving many time periods at unequal intervals are needed to address this issue.

The switched-partners condition showed the largest differences from true synchrony. These pseudointeraction types were far removed from the originals, as the on-screen interactants were not really interacting with each other. Any synchrony occurring in these clips not attributable to chance would have to stem from some biologically based behavioral cycle present in all humans. Although this is an intriguing idea, its investigation was beyond the scope of this project.

We had expected the double-crossed pseudoclip to show the largest difference from the true interactions because it had combined two different methods for destroying the appearance of any entrainment synchrony. For example, it was possible that the first minute of a mother–child interaction would be different (e.g., more or less active) from the last minute. If all mother–child interactions were characterized by similar universal differences in behavior from the first to the last minute, switching the time frames and switching partners could further disrupt the appearance of entrainment synchrony.

When the two control methods were combined, the double-crossed controlled synchrony was slightly lower, although not significantly, than the switched-partners controlled synchrony (see Table 2). Therefore, the disruptive effects of the switched-partners and altered time-frame controls on the appearance of entrainment synchrony were not additive and were not entirely independent of each other. These results suggest that the appearance of synchrony was not affected independently by universal sequences of behavior within our interactions.

A Global Concept of Synchrony?

Even with the apparent success of direct synchrony ratings, one question remains. Earlier we were forced to use a composite rating of synchrony because the three initial rating variables of simultaneous movement, tempo similarity, and coordination were all highly correlated. Analyses were repeated separately for all three, but there were no differences between them. As a consequence, we are left with a vague and undifferentiated concept of synchrony. We know what our judges were not rating, but we could not precisely identify the criterion they were using to make their ratings. We know through our two validation studies that they were sensitive to the entrainment process, but whether they were focusing on simultaneous movement, congruent rhythms, or gestaltilike impressions is unknown.

Are all three hypothesized components of synchrony truly linked? Are they all measures of the same phenomenon? Are humans simply unable to distinguish between them? It is possible that they simply cannot be measured simultaneously by the same rater. The three rating tasks may each require one's undivided attention. To present all at once may exceed subjects' cognitive faculties. In future investigations, it may be beneficial to have three sets of judges rate only one of the three possible components to determine whether there is some kind of halo effect on synchrony judgments.

A Synchrony Finding

We found perceptible differences between the synchrony in related mother–child dyads and unrelated mother–child dyads. Mothers interacting with their own children were more syn-
chronous than unrelated mothers and children, a difference that increased with time. This finding may be due to either of two functions of synchrony. First, mother–infant synchrony plays a central role in the organization of the face-to-face encounter (Beebe et al., 1982). Preverbal infants often have difficulty moderating their stimulus input. In order to avoid overstimulation, mothers may tune in to their child in an attempt to keep pace with their child (Pennebaker, Meares, Baker, & Milgram-Friedman, 1983). Mothers may be more acquainted with, or biologically linked, to their own children’s behavioral patterns than with strangers’ children.

Another possibility involves a motivational function. A state of high synchrony may indicate existing states of or motivation for sociability among participants (Rosenfeld, 1981). The relative differences in synchrony between the related and unrelated pairs may reflect nothing more than the increased rapport a child has with his or her own mother than with that of a stranger.

Another interesting finding was the substantial drop in synchrony within the unrelated mother–child dyad toward the end of their brief encounter. Condon (1982) has studied nonsynchronous states and has suggested that individual differences in synchrony could be diagnostic of psychopathology. In fact, he equated pathology with the absence of synchrony. In our studies, we did not look for individual differences in synchrony, so his hypothesis was not tested; however, we did find a situational effect on synchrony. The extremely low levels of rated synchrony occurred only in the last minute of the unrelated mother–child interaction. It would be incorrect to say that this reflects an absence of synchrony, because an asynchronous condition would suggest no differences between the true synchrony and pseudosynchrony ratings. Because the true synchrony was actually lower than the baseline pseudosynchrony, we can infer something stronger, what we have termed a dis synchronous state.

It has been suggested that synchrony communicates interest and approval (Kendon, 1970). For young children it may be a means by which they regulate their involvement in social interactions. Tronick et al. (1977) speculated that although synchrony may be a way for a child to communicate continue, dis synchrony may communicate stop. For the children in our study, who seemed slightly agitated by the third minute with the female stranger, the message communicated could have been something more affect laden, such as “I’ve had enough, leave me alone!” or “Where’s my Mommy?”

Conclusion

Our results should encourage investigators to use raters to study synchrony. Dyadic interactions of all types (e.g., doctor–patient, teacher–student, interviewer–interviewee, husbandwife, etc.) are amenable to this technique, and ratings of synchrony can be collected quickly and cheaply. It is probably inappropriate to use rating data to estimate absolute magnitudes of synchrony, unless the use of pseudointeraction controls are used. But even by using a relative scale, much can be learned through the widespread use of this efficient method of direct synchrony measurement.

References


Received February 19, 1986
Revision received July 20, 1987
Accepted July 28, 1987